Title: **Predicting the onset of major depression in primary care:** International validation of a risk prediction algorithm from Spain.

Authors: Juan Ángel Bellón[1], *MD, PhD*; Juan de Dios Luna[2], *PhD*; Michael King[3], *MD, PhD*; Berta Moreno-Küstner[4], *PhD*; Irwin Nazareth[5], *MD, PhD*; Carmen Montón-Franco[6], *MD, PhD*; María Josefa GildeGómez-Barragán[7], *MD, PhD*; Marta Sánchez-Celaya[8], *MD, PhD*; Miguel Ángel Díaz-Barreiros[9], *MD, PhD*; Catalina Vicens[10], *MD, PhD*; Jorge A. Cervilla[11], *MD, PhD*; Igor Švab[12], *MD, PhD*; Heidi-Ingrid Maaroos[13], *MD, PhD*; Miguel Xavier[14], *MD, PhD*; Mirjam I Geerlings[15], *PhD*; Sandra Saldivia[16], *PhD*; Blanca Gutiérrez[11], *PhD*; Emma Motrico[17], *PhD*; María Teresa Martínez-Cañavate[18], *MD, PhD*; Bárbara Oliván-Blázquez[19], *PhD*; Ana Vázquez-Medrano[7], *MD, PhD*; María Soledad Sánchez-Artiaga[20], *MD, PhD*; Sebastiá March[21], *PhD*; María del Mar Muñoz-García[22], *PhD*; Patricia Moreno-Peral[23], *PhD*; and Francisco Torres-González[11], *MD, PhD*.

[1] Centro de Salud El Palo. Unidad de Investigación del Distrito de Atención Primaria de Málaga (redIAPP, grupo SAMSERAP). Departamento de Medicina Preventiva. Universidad de Málaga, Spain. [2] Departamento de Bioestadística (redIAPP, grupo SAMSERAP). Universidad de Granada, Spain. [3] Department of Mental Health Sciences, UCL, UK. [4] Fundación IMABIS. Unidad de Investigación del Distrito de Atención Primaria de Málaga (redIAPP, grupo SAMSERAP). Departamento de Personalidad, Evaluación y Tratamiento Psicológico. Universidad de Málaga, Spain. [5] Dept of Primary care and Population Health, University College London and Medical Research Council General Practice Research Framework, UK. [6] Centro de Salud Casablanca. (redIAPP, grupo Aragón) Departamento de Medicina y Psiquiatría. Universidad de Zaragoza, Spain. [7] Unidad Docente de Medicina Familiar y Comunitaria de La Rioja. Servicio Riojano de la Salud, Logroño, La Rioja. Spain. [8] Unidad Docente de Medicina Familiar y Comunitaria. Área I de Atención Primaria, Madrid. Coordinadora de Investigación de la Sociedad Española de Medicina Familiar y Comunitaria. Spain. [9] Centro de Salud Vecindario, Gerencia de Atención Primaria de Gran Canaria, Servicio Canario de Salud, Las Palmas, Spain. [10] Centro de Salud son Serra-La Vileta. Unidad Docente de Medicina Familiar y Comunitaria de Mallorca. Instituto Balear de la Salud *(redIAPP, grupo Baleares).* Palma de Mallorca, Illes Balears. Spain. [11]CIBERSAM. Departamento de Psiquiatría y Medicina legal. Universidad de Granada, Spain. [12] Department of Family Medicine, University of Ljubljana, Slovenia. [13] Faculty of Medicine, University of Tartu, Estonia. [14] Faculdade Ciências Médicas, University of Lisbon, Portugal. [15] University Medical Center, Utrecht, Netherlands. [16] Departamento de Psiquiatría y Salud Mental, Universidad de Concepción, Chile. [17] Fundación IMABIS. Unidad de Investigación del Distrito de Atención Primaria de Málaga (redIAPP, grupo SAMSERAP). Departamento de Psicología Social. Universidad de Málaga, Spain [18] Fundación IAVANTE, Granada, Spain. [19] Unidad de Investigación de Atención Primaria (redIAPP, grupo Aragón). Instituto Aragonés de Ciencias de la Salud, Zaragoza. Spain. [20] Centro de Salud Condes de Barcelona-Boadilla, Área 6 de Atención Primara, Madrid, Spain. [21] Unidad de Investigación de Atención Primaria de Baleares (redIAPP, grupo Baleares). Mallorca. Spain. [22] Departamento de Psiquiatría y Medicina legal. Universidad de Granada, Spain. [23]Fundación IMABIS. Unidad de Investigación del Distrito de Atención Primaria de Málaga (redIAPP, grupo SAMSERAP).

Author for correspondence: Juan Ángel Bellón.

Departamento de Medicina Preventiva
Facultad de Medicina, Universidad de Málaga
Campus de Teatinos
29071 Málaga, Spain
Phone: +34 952 131515
Fax: +34 952 137131
JABELLON@terra.es

Word count: Total: 3,798

ABSTRACT

Context: The different incidence rates of major depression and its associated risk factors suggest the need for specific national rather than supranational risk algorithms.

Objectives: 1) To develop and validate a predictD-Spain-risk-algorithm for the onset of major depression and 2) To compare the performance of the predictD-Spain-risk-algorithm with the predictD-Europe-risk-algorithm in Spanish primary health care.

Design: Prospective cohort study with evaluations at baseline, 6 and 12 months. We measured 39 known risk factors and used multilevel logistic regression and inverse probability weighting to build the predictD-Spain-risk-algorithm. We corrected the model for overfitting and tested it in data from other countries.

Setting: Health Centers in Europe and South-America.

Participants: In Spain (4574), Chile (2133) and other 5 European countries (5184), 11891 non depressed adult primary care attendees formed our at risk population.

Main Outcome Measures: DSM-IV major depression (Composite International Diagnostic Interview).

Results: The predictD-Spain-risk-algorithm was developed in 2787 primary care attendees in Spain and its use validated in Chile (1844) and five other European countries (4075). Six variables were patient characteristics or past events (sex, age, sex*age interaction, education, physical child abuse, and lifetime depression) and six were current status (SF-12-physical-score, SF-12-mental-score, dissatisfaction with unpaid work, number of serious problems in very close persons, dissatisfaction with living together at home, and taking medication for stress, anxiety or depression). Province was the thirteenth factor. The C-index of the predictD-Spain-risk-algorithm was 0.82 (95%CI=0.79-0.84) and in other countries it ranged between 0.70-0.83. Both the test for C-index differences (difference=0.0316; 95%CI=0.0121-0.0530; p<0.0022) and calibration plots showed that the predictD-Spain-risk-algorithm functioned better than the predictD-Europe-risk-algorithm in Spain. However, this did not hold true when applied to other countries in Europe or Chile.

70 Conclusions: The predictD-Spain-risk-algorithm included new variables and offered an improved
71 performance over the predictD-Europe-risk-algorithm for predicting the onset of major depression in
72 Spain.
73

Effective strategies for preventing depression and reducing disease burden are hindered by lack of evidence about whether the risk for major depression can be quantified in the same way as other clinical disorders, such as cardiovascular disease.[1]

The predictD study is a pioneering international study whose main objective was to develop a risk index for the onset of major depression in general practice attendees.[2] From 39 potential risk factors for depression, a risk index of 10 risk factors was drawn up for Europe that has excellent predictive power and good external validity.[3]

Risk models for do not always apply well across countries. For example the incidence of myocardial infarction is relatively low in Spain and other southern European countries and individual cardiovascular risk estimates based on the classic Framingham model have been shown to overestimate the actual individual risk in Spanish persons.[1,4]

Likewise, evidence exists of different prevalence rates of depression across Europe, even after adjusting for likely confounding factors.[5] Moreover, important variations are found across European countries concerning the prevalence of factors associated with depression, including comorbidity of mood and anxiety disorders, living arrangements, unemployment, perception of life events and social support, quality of life, impact of work on mental health, and help seeking from services for mental health problems or misuse of psychotropic drugs.[6-8] Consequently, rather than simply calibrating existing risk prediction tools, new country specific prediction risk scores for the onset of depression are required.

Based on the predictD-International[2-3] and the predictD-Spain study[9] we aimed to develop and to validate a risk algorithm for the onset of major depression in Spanish primary care attendees, and to compare the performance of the European and Spanish risk algorithms in Spanish data.

METHODS

Design

We undertook a prospective cohort study to develop and validate a risk prediction algorithm for the onset of major depression at 12 months in Spanish primary care attendees. We assessed the external validity of the Spanish-risk-algorithm in data from a previous study of patients attending general practices in Chile and 5 other European countries (UK, Netherlands, Estonia, Slovenia and Portugal). The method has been described in detail elsewhere.[2-3,9] The predictD-Spain study was approved by ethical committees in each Spanish province.

Setting

Seven provinces participated with 41 health centers and 231 physicians distributed throughout Spain: Malaga and Granada in southern Spain; Zaragoza and La Rioja in northern Spain; Madrid, capital of Spain, situated in the center; Las Palmas in the Canary Islands; and Majorca in the Balearic Islands. Each health center covers a population of 15,000 to 30,000 inhabitants from a geographically defined area. The physicians in each health center work as a group, with extensive primary care teams. The Spanish National Health Service provides free medical cover to 100% of the population. The health centers taking part cover urban and rural settings in each province.

The external validation study used data collected in the original predictD-International study in 5 other European countries: 25 health centers in the Medical Research Council General Practice Research Framework in the United Kingdom; 74 health centers nationwide in Slovenia; 23 Health Centers

nationwide in Estonia; 7 large Health Centers near Utrecht, the Netherlands; 2 large health centers in Portugal, one in Lisbon and the other in Alentejo; and 78 health centers in Concepción and Talcahuano in the Eighth Region of Chile.

Participants

In the six Spanish provinces systematic random samples from physician appointment lists were taken at regular intervals of between 4 and 6 attendees with random starting points for each day. The study population, aged 18 to 75 years, was recruited between October 2005 and February 2006. The seventh province, Malaga, recruited between October 2003 and February 2004 as it was already participating in the predictD-International study. The external validation data were collected in consecutive attendees aged 18 to 75 years who had been recruited in Europe between April 2003 and September 2004 and in Chile between October 2003 and February 2005. Exclusion criteria for all participant countries were an inability to understand one of the main languages involved, psychosis, dementia, and incapacitating physical illness. In the United Kingdom and the Netherlands patients were recruited in clinics waiting rooms while in the other countries recruitment was conducted in discussion with the family physician. In Chile, attendees were randomly selected stratified by age and sex in each center. Participants who gave informed consent undertook a research interview within two weeks.

Variables

A DSM-IV diagnosis of major depression in the preceding 6 months was made using the depression section of the Composite International Diagnostic Interview (CIDI) at baseline, 6 and 12 months.[10-12] The risk factors selected cover all important areas identified in a systematic review of the literature. For the test-retest analysis, we selected in Spain a random sample of 401 patients stratified by province; 251

completed researcher-administered questionnaires and 150 self-administered questionnaires before the main study began.[9] Test-retest reliability of questions used in the predictD-International study has been reported previously.[2-3] All potential risk factors for depression were measured at baseline:

- Socio-demographic factors: age, sex, marital status, occupation, employment status, ethnicity, nationality, country of birth, educational level, income, owner-occupier of an accommodation, living alone or with others.
- Controls, demands and rewards for unpaid and paid work, using an adapted version of the job content instrument.[13]
- Debt and financial strain.[14]
- Physical and mental well-being, assessed by the SF-12[15-16] and a question on the presence of long-standing illness, disability or infirmity.
- Alcohol misuse, assessed by the Alcohol Use Disorders Identification Test (AUDIT).[17-19]
- A life-time screen for depression based on the first two questions of the CIDI.[20]
- Lifetime use of recreational drugs.[10-12]
- Brief questions on the quality of sexual and emotional relationships with a partner, adapted from a standardized questionnaire.[21]
- Anxiety symptoms using the anxiety section of the Primary Care Evaluation of Mental Disorders (PRIME-MD).[22-23]
- Childhood experiences of physical, emotional or sexual abuse.[24]
- Nature and strength of spiritual beliefs.[25]

- Presence of serious physical, psychological or substance misuse problems, or any serious disability, in persons who were close friends or relations of participants; and difficulty getting on with people and maintaining close relationships, assessed using questions from a social functioning scale.[26]
- Family psychiatric history in first-degree family members, and suicide in first-degree relatives.[27]
- The living environment, including satisfaction with neighborhood and perception of safety inside and outside the home using questions from the Health Surveys for England.[28]
- Recent life-threatening events, using a brief validated checklist.[29]
- Experiences of discrimination on the grounds of sex, age, ethnicity, appearance, disability or sexual orientation using questions from a European study.[30]
- Adequacy, availability and sources of social support from family and friends.[31]

Statistical analyses

In order to develop the predictD-Spain-risk-algorithm we included only patients without major depression at baseline. Participants with missing depression diagnoses at both follow-up points (at 6 and 12 months) were excluded. We also excluded those with missing CIDI data at one follow-up who were not depressed at the other, as we could not conclude whether or not major depression had occurred over the follow-up period. However, we could include patients who were depressed at one follow-up point and missing at the other (at 6 or 12 months), as they met the outcome criterion of depression at some point over the 12 months. We conducted all analyses using STATA, release 10.[32]

Data imputation

Missing data in candidate risk factors were imputed using the method of chained equations, implemented in the *STATA* "Ice" program.[33] We imputed 10 data sets[34] and obtained combined estimates.[35]

Model Building

We performed multilevel logistic regressions to test the hierarchical data structure with the cumulative incidence of depression at 12 months as the dependent variable. The likelihood-ratio test of the null model with health center as a random factor versus usual logistic regression was significant (χ^2=15.20, p<0.001). Nevertheless, the likelihood-ratio test of the null model with health center and doctor as a random factor versus the null model with only health center was not significant (χ^2=1.48, p=0.11). The intraclass correlation coefficients for incidence of depression at 12 months were 0.07 and 0.03 for health center and doctor, respectively. Hence, we used multilevel logistic regression with two levels, patients and health center. We built the risk model at 12 months using all the risk factors described earlier and the province of each participant. We developed these models in the imputed data 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 minimall.[36] We retained age and sex in all regression models because of their well-known associations with the onset of depression.[37] We also retained province because of an a priori assumption of clustering within province, although it had few categories (n=7) that could be considered as random factors.[38] The usefulness of including first-degree interactions was considered. Multivariable fractional polynomial analysis was used to assess possible nonlinear effects of continuous predictors. From the model thus obtained, those variables with p≥ 0.05 were extracted step by step to obtain a more parsimonious model. The variables that modified coefficients by more than 10%, irrespective of the value of p, remained in the model. For each patient the probability of remaining in the follow-up at 12 months was obtained[39] and then inverse probability weighting was applied to the final model to adjust

for a possible selection bias due to participants lost during the follow-up;[40] implemented through the *STATA* "gllamm" program.[41] We repeated the analyses in participants with complete data as a sensitivity analysis.

Internal Validation

The ability to distinguish those who will develop major depression from those who will not, was assessed using the C-index.[42] We used a calculation proposed by Copas[43] to adjust for over fitting of our prediction models. To deal with the overfitting that arises through variable selection, we computed the shrinkage factor based on the initial model including all variables. We calculated effect sizes using the Hedge's g[44] for the difference in log odds of predicted probability between patients who were later found to be depressed and those who were not. We assessed the goodness of fit of the final risk model by grouping individuals into deciles of risk and comparing the observed probability of major depression within these groups with the average risk (*Calibration plots*).

External Validation

We used the C-index, Hedge's g, and calibration plots to evaluate the performance of the predictD-Spain-risk-algorithm (without province) in data from Chile and the other European countries.[2] When the predictD-Spain-risk-algorithm was tested in data from all European countries, we excluded Spanish patients from the European sample. We estimated the same parameters applying the predictD-Europe-risk-algorithm[3] in Spanish data to compare both models, the European and the Spanish. In this last comparison the sample from Malaga was excluded from the sample since it was used to develop the predictD-Europe-risk-algorithm. A test for difference in 2 correlated C-index, estimating error standard by bootstrap, was used.[45] Furthermore, we calculated the *"Integrated Discrimination Improvement"* (IDI)

and the asymptotic test for the null hypothesis of IDI=0.[46] The IDI can be viewed as a difference between improvement in average sensitivity and any potential increase in average '1-specificity'.

RESULTS

In the 7 Spanish provinces 6526 people were asked to take part. Response to recruitment was 83.4%; 5442 were interviewed and 1084 refused to participate at baseline (Figure 1). Of those who refused to participate, 780 gave their consent for their age and sex data to be used in our analysis. A higher proportion of these latter were male, 360 of the 780 (46.1%) versus 1756 of the 5442 (32.3%) patients who provided baseline information, χ^2=18.06 and p<0.001, and those who refused had a lower mean age, 46.9 [95% confidence Interval (CI): 45.7-48.0] versus 48.5 years (95% CI, 48.1- 48.9), *P*=0.018.

At recruitment 5360 participants had full CIDI data to allow a depression diagnosis; of these 4574 were not depressed. The response to follow-up was 70% at 6 months and 66% at 12 months. The analysis of variables associated with non-response has been described elsewhere.[39] In brief, Province and socio-demographic factors were strong predictors of loss to follow-up: those who did not respond were younger, had a lower level of education and income, and were more often male, single, born outside Spain, and less often students than those who responded. Major depression and anxiety had no effect but other psychosocial factors predicted attrition.[39]

In the 6 countries (apart from Spain) that had participated in PredictD, 8567 people took part and their responses to recruitment were, in decreasing order: Chile (97%), Estonia (80%), Slovenia (80%), Portugal (76%), Netherlands (45%) and the UK (44%). In Chile, 2133 participants who were not depressed started the follow-up and 5184 in the 5 other European countries; the respective response to

follow-up at 6 and 12 months was 89% and 82%, and 91% and 88%.[4] The sociodemographic differences between countries are shown in Table 1.

The cumulative 12 months' incidence of *DSM-IV* major depression was 11.5% in Spain, varying between provinces: Las Palmas 17.5%, Malaga 15.2%, Granada 14.9%, Majorca 14%, Zaragoza 7.9%, Madrid 7.2% and La Rioja 5.6%. The incidence in the other countries was: UK 8.8%, Slovenia 4.2%, Portugal 8.5%, Netherlands 5.4%, Estonia 5.9%, and Chile 11.6%.[3]

Missing information in Spanish data was less than 1% for most risk factors; exceptions were ethnicity (2.9%), suicide in brothers or sisters (3.30%), and sexual or emotional relationship with a spouse or partner (18.3%).

The results of reliability analyses have been reported elsewhere.[9] They were good or excellent for practically all the questionnaires and items. However, in Spain the questions on the use of any recreational drugs over the previous 6 months were removed due to poor reliability.[9]

Development of the Spanish model

Eight variables were retained at $p<0.05$ (Table 2), and one more (physical child abuse) was included at $p=0.085$ because coefficients changed more than 10% when this was removed. The interaction age*sex was also included in the final equation because the likelihood ratio test was significant ($\chi^2=4.03$; 1df; $p=0.0447$). Interaction between sex and each of the remaining risk factors in the model was not significant at $p \leq 0.10$. Nor were interactions between age and the other variables in the model significant. Nonlinear transformation of continuous variables did not significantly improve the model fit. Six variables were patient characteristics or past events (sex, age, sex*age interaction, education, physical child abuse and lifetime depression) and six were current status (Short-Form-12-physical-

health-subscale-score, Short-Form-12-mental-health-subscale-score, dissatisfaction with unpaid work, number of serious problems in very close persons, dissatisfaction with living together at home, and taking medication for stress, anxiety or depression); and one concerned Spanish province. The random component (health center) was also significant even after including all variables of the fixed component in the regression models; these coefficients were 0.390 (SE=0.085; p<0.0001) and 0.469 (SE=0.101; p<0.0001) for the models with and without province, respectively. The model derived in participants with complete data (N=2544) and the model derived in the ten imputed datasets (N=2787) were very similar, except for the variables set "age, sex, and age*sex interaction" that was more significant in the model with complete data (Annex 1); nevertheless, there were more differences between the model derived in the ten imputed datasets and the same weighted for the inverse probability of remaining in the follow-up to 12 months (Annex 1).

Internal validation

The average C-index in data sets for predicted probability of depression at 6 or 12 months was 0.82 (95% CI, 0.79-0.84).The average effect size in data sets (Hedge's g) was 1.35 (95% CI, 1.21-1.48). The calibration plot is shown in Figure 2. The predicted probability of depression at 0.113 was associated with estimates of sensitivity, specificity and likelihood ratio (+) of 72.8%, 72.6%, and 2.67, respectively. Examples of the kinds of participants scoring at increasing levels of predicted probability of depression are shown in Annex 2. Predicted probability of major depression over 12 months can be calculated through the predictD-Spain-risk-algorithm at http://www.rediapp.org/predict.php

External validation

The Copas shrinkage factor for the Spanish model was 0.873 including the province and 0.872 without the province. The shrunk regression coefficients are shown in Table 2. The C-index ranged from 0.70 in

Chile to 0.83 in The Netherlands and the Hedge's g from 0.77 in Chile to 1.50 in The Netherlands (Table 3). Calibration plots of the predictD-Spain-risk-algorithm in Chile and the other European countries are shown in Figure 2. The predictD-Spain-risk-algorithm slightly overestimated each decile risk score in Estonia and Slovenia, with the gap increasing as the risk score rose; in The Netherlands, the 6th to 8th deciles were overestimated and in Portugal the 8th to 10th; in the UK the 6th to 9th deciles were underestimated and in Chile the 3rd to 5th; in Europe (without Spain) the predictD-Spain-risk-algorithm slightly overestimated the 6th and 8th to 10th deciles.

When we applied the predictD-Europe-risk-algorithm to the Spanish data (i.e. excluding the sample recruited from Malaga): the C-index was 0.78 (95% CI, 0.73-0.83) and the Hedge's g 1.14 (95% CI, 0.98-1.31) (Table 3); the test for the C-index difference between the predictD-Spain-risk-algorithm and predictD-Europe-risk-algorithm was significant (difference=0.0316; 95%CI=0.0121-0.0530; Z_{exp}=3.10; p<0.0022); the IDI was 0.0461 (SE=0.0095; Z_{exp}=4.84; p<0.0001) due to the increase of sensitivity (0.2710 vs. 0.2463; Z_{exp}=2.62; p<0.0087) and the decrease of "1-specificity" (0.09104 vs. 0.1124; Z_{exp}=14.86; p<0.0001). Calibration plots comparing the predictD-Europe-risk-algorithm and predictD-Spain-risk-algorithm in Spanish data are shown in Figure 2.

DISCUSSION

We have developed a risk score for the development of major depression over 12 months in 2787 general practice attendees in Spain and externally validated in the data from Chile and 5 other European countries. Six risk factors concerned personal characteristics or past events and another six current status that might be open to change. The discriminative power and calibration of the predictD-Spain-risk-algorithm were high when applied to Spanish populations but lower when applied to data from 6 external populations. The predictD-Europe algorithm performed well in the new Spanish data with a C-index very close to that fount during its development.[3] Nevertheless, the predictD-Spain-risk-

algorithm, which includes new variables and removes others, is more accurate than the predictD-Europe-risk-algorithm in prediction of the onset of major depression in Spain.

We recruited a systematic random sample of primary care attendees and we used a criterion of stratification to include urban and rural health centers in each province and included provinces from different geographical areas in both mainland Spain (north, central and south) and the Spanish islands. Although we did not select health centers randomly and our sample could under-represent patients who attend very infrequently,[47] the study population is likely to be fairly representative of primary care attendees in Spain. If the predictD-Spain-risk-algorithm is applied in a different country, or a province other than 1 of the 7 participating provinces, we recommend using the shrunk coefficients of the model without province (Table 2).

We used multilevel regression because of the hierarchical structure of the data. In these cases, this approach improves the accuracy of estimates of coefficients and standard errors.[38] Our large sample size and the number of events (major depression) per variable included in the model (>29) contributed to reducing the risk of selecting unimportant variables and failing to include important ones.[48] The multiple imputation strategy allowed us to gain statistical power and to avoid potentially biased estimates obtained from a reduced complete-case dataset.[35] There were important differences between the Spanish models with and without inverse probability weighting, indicating that loss to follow-up might lead to selection bias and suggesting that this strategy could provide unbiased estimates of coefficients, even in the presence of selection bias.[40] We think follow-up at 12 months is appropriate for the prediction of the onset of depression in primary care because this is sufficient time to develop major depression (11.5% of incident cases). Furthermore, doctors and patients may be more motivated to undertake interventions and behavioral changes when depression is likely to happen sooner rather than later.

In our study we considered a large number of known risk factors for depression, using either validated measures or validating them when necessary.[9] The reliability of the questionnaires was high.[2,9] We evaluated the performance of our model on external datasets in large samples of patients from different cultures and countries, and corrected for over-fitting using shrinkage factors applied to the regression coefficients.

The predictD-Europe and predictD-Spain algorithms share many risk factors in common; however, the new risk factors included in the Spanish equation improved its results for prediction in that country. "*Dissatisfaction with living together at home*" and "*Number of serious problems in very close persons*" are risk factors consistent with the geography of family systems; the center and northern part of Europe, together with North American society, have been characterized by relatively weak family links, whereas the Mediterranean region has strong family ties.[49] Spain belongs to the regions where the family group has traditionally had priority over the individual. Moreover, the association between marital discord, family dysfunction and depression is well known.[50] The inclusion of "*Dissatisfaction with unpaid work*" instead of *"difficulties in paid and unpaid work*" may be due to different ways of measuring these variables. We used two scales in Spain, one for unpaid work and another for paid work, with 7 items each that were valid and reliable,[9] whereas in the predictD-Europe study two items were used to summarize both dimensions together. When we included these 2 items in the predictD-Spain model instead of work scales, it was not significant. However, we cannot rule out the influence of other factors, such as a higher participation of Spanish women in domestic work as compared with other European regions.[51] Although relationships between "*Physical childhood abuse*" and depression are well documented,[52] they are complex, vary between countries, and have cross-cultural differences.[53] Finally, the variable "*Taking medication for anxiety, depression or stress*" might be associated with patients who have suffered previous depressive episodes and were still taking antidepressants. However, the question is phrased with such view that it might also include those taking anxiolytics, often in an

inadequate way, for anxiety comorbidity or even just taking medicines (vitamins, placebos, etc.) for other minor emotional problems. A hypothesis might be that these patients share coping styles such as "Health External Locus" with a tendency to ask their doctors for more psychotropic drugs for emotional problems encountered in everyday life.[54] Spain is also among those European countries that have a higher use of psychotropic drugs.[55]

Our predictD-Spain-risk-algorithm could help physicians with decisions by providing more objective estimates of probability, as a supplement to other relevant clinical information; perhaps in a similar manner to the way cardiovascular risk scores are used to determine the indication for cholesterol lowering. However, trials are needed using our predictD-Spain-risk-algorithm to test different strategies of primary prevention of major depression. Impact studies are also needed to quantify the effect of using our predictD-Spain-risk-algorithm on physicians' behavior, patient outcome, or cost effectiveness of care.

Acknowledgements

The authors thank the Primary Care District of Malaga, particularly Dr. Jose Miguel Morales and Dr. Maximiliano Vilaseca for their support. We also thank Patrick Royston and Christian Bottomley (both UK) for their contribution to the analytical strategy in predictD and predictD-Spain, and Peter Zuithoff (The Netherlands) for his statistic comments of the article. Finally we also thank all interviewers, physicians and patients for their participation in the predictD-Study.

Conflict of interest

The authors all declare they have no competing interests. The funders had no direct role in the design or conduct of the study, interpretation of the data, or review of the manuscript.

Grants

This work was supported in Spain by grants from the Spanish Ministry of Health [grant FIS references: PI041980, PI041771, PI042450, and PI06/1442]; the Andalusian Council of Health [grant references: 05/403, and 06/278] and and the Spanish Ministry of Education and Science [grant reference SAF 2006/07192]; 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 Malaga sample, as part of the predictD-International study, was also co-funded by a grant from The European Commission [reference QL4-CT2002-00683]. The research in Europe was funded by a grant from the European Commission [reference PREDICT-QL4-CT2002-00683]. Funding in Chile was provided by project FONDEF DO2I-1140. Partial support in Europe was from the Estonian Scientific Foundation [grant 5696], and the Slovenian Ministry for Research [grant 4369-1027]. The UK National Health Service Research and Development office provided service support costs in the United Kingdom. The funders had no direct role in the design or conduct of the study, interpretation of the data or review of the manuscript.

Author contributions

JAB is guarantor for the predictD-Spain study. MK and IN had the original idea for the predictD study and its analysis, which formed the basis for the approach used in predictD-Spain. JAB, BMK, FTG, JAC and CMF obtained funding for implementing the study in Spain. JAB coordinated the predictD-Spain study. BMK, FGT, CMF, MJGdGB, MSC, MADB, and CV coordinated the study in each Spanish province. JAC and BG coordinated the predictD-Gene study. MTMC, BOB, AVM, MSSA, SM, EMM, MMMG, and PMP collaborated implementing the study in each province. MK is guarantor for the predictD-Europe study and coordinated this study. IN, MK, IS, HIM, MX, MI coordinated the predictD-Europe in their respective European countries. SS is guarantor for the predictD-Chile study and coordinated this study. JDL collaborated in the design, and JAB and JDL analyzed the data. MK and IN contributed to the approach to the analysis. JAB drafted the paper and all authors discussed and agreed the final version.

References:

1. 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.

2. King M, Weich S, Torres F, et al. Prediction of depression in European general practice attendees: the PREDICT study. BMC Public Health 2006; 6:6.

3. King M, Walker C, Gus L, 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.

4. Marrugat J, Subirana I, Comin E, et al. Validity of an adaptation of the Framingham cardiovascular risk function: the VERIFICA study. J Epidemiol Community Health 2007; 61:40-7.

5. King M, Nazareth I, Levy G, et al. Prevalence of common mental disorders in general practice attendees across Europe. Br J Psichiatry 2008; 192:362-67.

6. De Girolamo G, Alonso J, Vilagut G. The ESEMeD-WMH project: strengthening epidemiological research in Europe through the study of variation in prevalence estimates. Epidemiol Psychiatric Soc 2006;15: 167-73.

7. König HH, Bernert S, Angermeyer MC, et al. Comparison of population health status in six European countries. Results of a representative survey using the EQ-5D questionnaire. Med Care 2009; 47:255-61.

8. European Commission. The state of the mental health in the European Union, 2004.

9. Bellón JA, Moreno B, Montón C, et al. Predicting the onset and persistence of episodes of depression in primary health care. The predictD-Spain study: Methodology. BMC Public Health 2008; 8:256.

10. 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; 45:1069–77.

11. World Health Organization. Composite International Diagnostic Interview (CIDI). Version 2.1. WHO: Geneva; 1997.

12. Rubio-Stipec M, Bravo M. [The Composite International Diagnostic Interview (CIDI): an epidemiologic instrument suitable for using in conjunction with different diagnostic systems in different cultures]. Acta Psiquiat Psicol Amer Lat 1991; 37:191–204. [Article in Spanish]

13. Karasek RA, Theorell T. Healthy work: stress, productivity, and the reconstruction of working life. Basic Books: New York; 1990.

14. Weich S, Lewis G: Poverty, unemployment, and common mental disorders: population based cohort study. BMJ 1998; 317:115-9.

15. 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; 19:179–86.

16. 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. J Clin Epidemiol 1998; 11:1171–8.

17. Barbor TF, de la Fuente JR, Saunders J, Grant M. The alcohol use disorders identification test: Guidelines for the use in primary health care. World Health Organization: Geneva; 1989.

18. 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; 198:11–4. [Article in Spanish]

19. Pérula-de Torres LA, Fernández-García JA, Arias-Vega R, Muriel-Palomino M, Marquez-Rebollo E, Ruiz-Moral R. [Validity of AUDIT test for detection of disorders related with alcohol consumption in women]. Med Clin (Barc) 2005; 125:727–30. [Article in Spanish]

20. Arroll B, Khin N, Kerse N. Screening for depression in primary care with two verbally asked questions: cross sectional study. BMJ 2003; 327:1144–6.

21. 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; 24:231–50.

22. Spitzer RL, Kroenke K, Williams JB. 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; 282:1737-44.

23. Baca E, Saiz J, Agüera L, et al. Validation of the Spanish versión of PRIME-MD: A procedure for diagnosing mental disorders in primary care. Actas Esp Psiquiatr 1999; 27:375–83.[Article in Spanish]

24. Fink LA, Bernstein D, Handelsman L, Foote J, Lovejoy M. Initial reliability and validity of the childhood trauma interview: a new multidimensional measure of childhood interpersonal trauma. Am J Psychiatry 1995; 152:1329–35.

25. King M, Speck P, Thomas A. The Royal Free interview for religious and spiritual beliefs: development and standardization. Psychol Med 1995; 25:1125–34.

26. Tyrer P. Personality disorder and social functioning. In: Peck DF, Shapiro CM, editor. Measuring Human Problems: a Practical Guide. Wiley & Sons, Chichester, New York; 1990. pp. 119–42.

27. Qureshi N, Bethea J, Modell B, et al. Collecting genetic information in primary care: evaluating a new family history tool. Fam Pract 2005; 22:663–9.

28. Sproston K, Primatesta P. Health survey for England 2002: a survey carried out on behalf of the Department of Health. The health of children and young people. Vol. 1. The Stationery Office: London; 2003.

29. Brugha T, Bebbington P, Tennant C, Hurry J. The List of Threatening Experiences: a subset of 12 life event categories with considerable long-term contextual threat. Psychol Med 1985; 15:189–91.

30. Janssen I, Hanssen M, Bak M, et al. Discrimination and delusional ideation. Br J Psychiatry 2003; 182:71–6.

31. Blaxter M. Health and Lifestyles. Routledge, London; 1990.

32. StataCorp. Stata Statistical Software: Release 10. Texas: College Station; 2007.

33. Royston P. Multiple imputation of missing values: update of ice. Stata Journal 2005; 5:527–36.

34. Schafer JL. Multiple imputation: a primer. Stat Methods Med Res 1999; 8:3–15.

35. Little RJA, Rubin DB. Statistical analysis with missing data. 2nd edition. Wiley, New York; 2002.

36. Greenland S. Modeling variables selection in epidemiologic analysis. Am J Public Health 1989; 79:340-9

37. Piccinelli M, Wilkinson G. Gender differences in depression. Br J Psychiatry 2000; 117:486-92.

38. Snijders TAB, Bosker RJ. Multilevel analysis. An introduction to basic and advanced multilevel modelling. London. Sage Publications, 1999.

39. Bellón JA, Luna JD, Moreno 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 2009 Sep 16; doi:10.1136/jech.2008.085290

40. Hernán MA, Hernández-Díaz S, Robins JM. A structural approach to selection bias. Epidemiology 2004; 15:615-25.

41. Rabe-Hesketh S, Skrondal A. Multilevel and longitudinal modelling using STATA. (2nd Ed.), College Station, Texas, STATA Press, 2008.

42. Harrell FE. Regression Modelling Strategies. New York: Springer; 2001.

43. Copas JB. Regression, prediction and shrinkage. J R Stat Soc Ser (Series B) 1983; 45:311–54.

44. Cooper H, Hedges LV. The Handbook of Research Synthesis. New York: Russell Sage Foundation; 1994

45. Pepe MS. The statistical evaluation of medical tests for classification and prediction. New York: Oxford University Press, 2003.

46. Pencina MJ, D'Agostino RB Sr, D'Agostino RB Jr, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. *Stat Med*. 2008;27:157–172

47. Lee ML, Yano EM, Wang MM, et al. What patient population does visit-based sampling in primary care setting represent? *Med Care* 2002;40:761-70

48. Altman DG, Royston P. What do we mean by validating a prognostic model? Statist Med 2000; 19:453-73.

49. Reher DS. Family ties in Western Europe: persistent contrasts. Popul Dev Rev 1998;24:203-34

50. Whisman MA, Uebelacker LA. Prospective associations between marital discord and depressive symptoms in middle-aged and older adults. Psychol Aging 2009; 24:184-9.

51. Drew E, Emerek R, Mahon E. Women, work, and the family in Europe. London: Routledge, 1998.

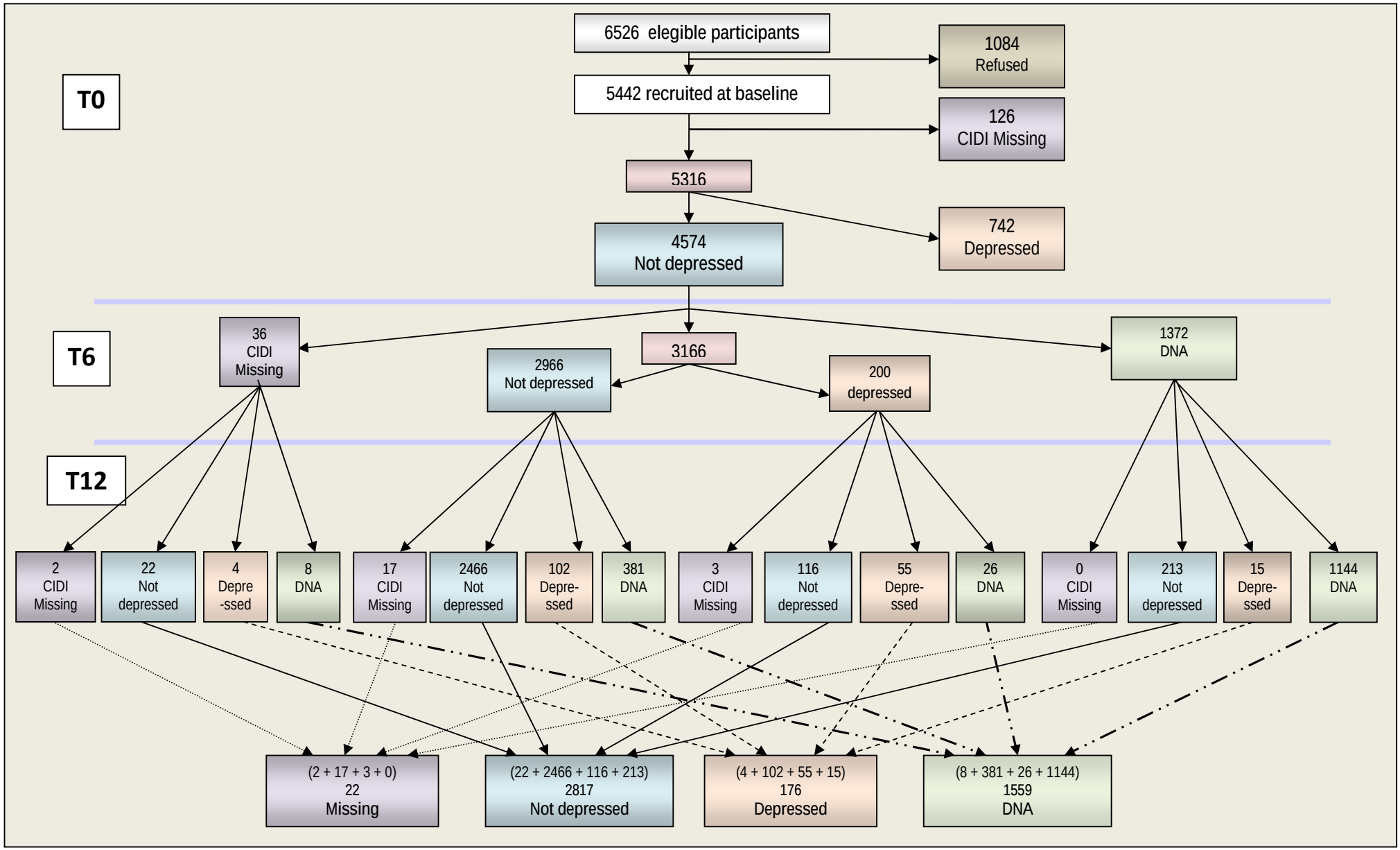
52. Arnow BA. Relationships between childhood maltreatment, adulth health and psychiatric outcomes, and medical utilization. J Clin Psychiatry 2004; 65 Suppl 12:10-15.

53. Sebre S, Sprugevica I, Novotni A et al. Cross-cultural comparisons of child-reported emotional and physical abuse: rates, risk factors and psychosocial symptoms. Child Abuse Negl 2004; 28:113-27.

54. Demyttenaere K, Bonnewyn A, Bruffaerts R et al. Clinical factors influencing the prescription of antidepressants and benzodiazepines: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. J Affect Disord 2008; 110:84-93.

551 55. Alonso J, Angermeyer MC, Bernert R et al. Psychotropic drug utilization in Europe: results from the European
552 Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand 2004; 109 (Suppl.
553 420):55-64.

Figure 1: Flow of patients through the predictD-Spain study and numbers becoming depressed.



CIDI: Composite International Diagnostic Interview; DNA: did not attend. T0, T6 and T12: baseline, 6 and 12 months interview

Table 1: Demographic characteristics and response to follow-up of not depressed participants at baseline.

Characteristic N(%)	Spain	Malaga	Granada	Zaragoza	Madrid	La Rioja	Balearic Isles	Las Palmas	Other European Countries	United Kingdom	Slovenia	Portugal	The Netherlands	Estonia	Chile
Not depressed participants	4574(100)	1030(22.5)	609(13.3)	660(14.4)	647(14.1)	680(14.9)	633(13.8)	315(6.9)	5184(100)	1131(21.8)	1048(20.2)	1005(19.4)	1077(20.8)	923(17.8)	2133
Age,yrs, mean (SD)	49.2(15.8)	49.9(15.5)	50.3(16.7)	47.3(15.4)	50.6(15.8)	49.5(15.8)	49.9(15.7)	43.5(14.5)	48.3(15.5)	52.2(14.7)	48.8(14.5)	50.2(15.4)	48.9(14.9)	41.6(16.0)	47(15.7)
Female	3020(66)	704(68.3)	424(69.6)	399(69.6)	432(66.8)	431(63.4)	413(65.2)	217(68.9)	3392(65.3)	750(66.3)	660(63)	649(64.6)	668(62)	665(72)	1522(71.4)
Marital Status															
Married or living together	3030(66.2)	725(70.4)	408(67)	443(67.1)	440(68)	447(65.7)	397(62.7)	170(54.0)	3783(73)	844(74.6)	732(69.9)	750(74.6)	827(76.8)	630(68.3)	1228(57.6)
Separated or divorced	279(6.1)	51(4.9)	29(4.8)	27(4.1)	42(6.5)	36(5.3)	54(8.5)	40(12.7)	372(7.2)	100(8.8)	56(5.3)	69(6.9)	64(5.9)	83(9)	179(8.4)
Single	955(20.9)	186(18.1)	113(18.6)	160(24.2)	126(19.5)	45(6.6)	129(20.4)	89(28.3)	691(13.3)	121(10.7)	152(14.5)	132(13.1)	121(11.2)	165(17.9)	521(24.4)
Widowed	308(6.7)	68(6.6)	57(9.4)	30(4.5)	39(6)	152(22.4)	53(8.4)	16(5.1)	316(6.1)	65(5.8)	105(10)	53(5.3)	48(4.5)	45(4.9)	205(9.6)
Missing	2(0.1)	0	2(0.3)	0	0	0	0	0	22(0.4)	1(0.1)	3(0.3)	1(0.1)	17(1.6)	0	0
Household Status															
Not living alone	4208(92.0)	971(94.3)	555(91.1)	615(93.2)	594(91.8)	616(90.6)	560(88.5)	297(94.3)	4535(87.5)	981(86.7)	915(87.3)	929(92.4)	894(83)	816(88.4)	2039(95.6)
Living alone	366(8.0)	59(5.7)	54(8.9)	45(6.8)	53(8.2)	64(9.4)	73(11.5)	18(5.7)	649(12.5)	150(13.3)	133(12.7)	76(7.6)	183(17)	107(11.6)	94(4.4)
Missing	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Education															
Higher education	552(12.1)	141(13.7)	75(12.3)	93(14.1)	59(9.1)	112(16.5)	33(5.2)	39(12.4)	1744(33.6)	448(39.6)	181(17.3)	129(12.8)	458(42.5)	528(57.2)	75(3.5)
Secondary	964(21.1)	220(21.4)	103(16.9)	173(26.1)	149(23.0)	140(20.6)	103(16.3)	76(24.1)	1823(35.2)	465(41.1)	385(36.7)	182(18.1)	508(47.2)	283(30.7)	791(37.1)
Primary/no education	2145(46.9)	424(41.2)	227(37.3)	319(48.3)	265(41.0)	382(56.2)	385(60.8)	143(45.4)	1111(21.4)	25(2.2)	235(22.4)	662(65.9)	78(7.2)	111(12)	998(46.8)
Trade/other	910(19.9)	245(23.8)	202(33.2)	75(11.4)	174(26.9)	45(6.6)	112(17.7)	57(18.1)	451(8.7)	171(15.1)	247(23.6)	32(3.2)	0	1(0.1)	265(12.5)
Missing	3(0.1)	0	2(0.3)	0	0	1(0.1)	0	0	55(1.1)	22(1.9)	0	0	33(3.1)	0	2(0.1)
Employment															
Employed/full-time student	2045(44.8)	359(34.9)	230(37.8)	363(55)	308(47.6)	353(51.9)	246(38.9)	186(59)	2907(56.1)	574(50.8)	563(53.7)	486(48.4)	602(55.9)	682(73.9)	749(35.1)
Unemployed	291(6.4)	64(6.2)	47(7.7)	36(5.5)	26(4)	43(6.3)	41(6.5)	34(10.8)	238(4.6)	21(1.9)	53(5.1)	108(10.7)	35(3.2)	21(2.3)	243(11.4)
Unable to work	315(6.9)	102(9.9)	42(6.9)	15(2.3)	29(4.5)	4(0.6)	110(17.4)	13(4.1)	221(4.3)	86(7.6)	16(1.5)	38(3.8)	48(4.5)	33(3.6)	69(3.2)
Retired/looking after family	1906(41.7)	500(48.5)	288(47.3)	244(37)	281(43.4)	279(41.1)	235(37.1)	79(25.1)	1776(34.2)	450(39.8)	409(39)	372(37)	358(33.2)	187(20.3)	1072(50.3)
Missing	17(0.2)	5(0.5)	2(0.3)	2(0.2)	3(0.5)	1(0.1)	1(0.2)	3(1)	42(0.8)	0	7(0.7)	1(0.1)	34(3.2)	0	0
Born in country of residence															
Yes	4327(94.6)	977(94.9)	594(97.5)	632(95.8)	607(93.8)	648(95.3)	586(92.6)	283(89.8)	4700(90.7)	1054(93.2)	834(79.6)	973(96.8)	997(92.6)	842(91.2)	2122(99.5)
Missing	20(0.4)	0	2(0.3)	1(0.2)	5(0.8)	0	7(1.1)	5(1.6)	84(1.6)	3(0.3)	4(0.4)	0	24(2.2)	53(5.7)	4(0.2)
Ethnicity															
White European	4371(95.6)	1017(98.7)	595(97.7)	551(83.5)	631(97.5)	647(95.1)	617(97.5)	313(99.4)	4994(96.2)	1055(93.3)	1042(99.4)	992(98.7)	983(91.3)	922(99.9)	0
Missing	131(2.9)	1(0.1)	6(1.0)	104(15.8)	1(0.2)	14(2.1)	5(0.8)	0	71(1.4)	39(3.4)	2(0.2)	0	30(2.8)	0	0
6-months response	3202(70.0)	810(78.6)	465(76.4)	520(78.8)	403(62.3)	468(68.8)	334(52.8)	202(64.1)	4744(91.4)	987(87.3)	963(91.9)	889(88.5)	1035(96.1)	870(94.3)	1904(89.3)
12-months response	3014(66.0)	743(72.1)	434(71.3)	448(67.9)	403(62.3)	507(74.6)	295(46.6)	184(58.4)	4588(88.5)	965(85.3)	927(88.5)	864(86)	988(91.7)	844(91.4)	1748(82)

Table 2: Spanish PredictD model[1] for predicting the onset of major depression at 12 months.

Prognostic Factors	Levels in factor	Model with "Province"				Model without "Province"			
		Coefficient	Coefficient[2]	SE	P	Coefficient	Coefficient[2]	SE	P
Constant		1.948	1.421	0.764	0.011	2.027	1.482	0.731	0.006
Province	Malaga								
	Granada	0.276	0.241	0.299	0.356				
	Zaragoza	0.166	0.145	0.321	0.606				
	Madrid	- 0.377	- 0.330	0.362	0.297				
	La Rioja	- 0.329	- 0.287	0.398	0.409				
	Balearic Isles	0.337	0.294	0.263	0.200				
	Las Palmas	0.308	0.269	0.191	0.106				
Age	Each year	- 0.032	- 0.027	0.008	< 0.001	- 0.032	- 0.028	0.008	<0.001
Sex	Female								
	Male	- 1.128	- 0.985	0.752	0.134	- 1.111	0.968	0.751	0.139
Sex*Age interaction	Each year	0.022	0.019	0.014	0.118	0.022	0.019	0.014	0.120
Education	Beyond secondary education								
	Secondary education	0.580	0.507	0.304	0.057	0.578	0.504	0.305	0.058
	Primary education	0.839	0.733	0.285	0.003	0.840	0.732	0.288	0.004
	Incomplete primary education or illiterate	1.490	1.302	0.369	< 0.001	1.500	1.307	0.370	<0.001
Physical Childhood Abuse	Never								
	Seldom	0.470	0.411	0.465	0.312	0.453	0.395	0.476	0.341
	Sometimes	0.320	0.279	0.300	0.286	0.345	0.300	0.295	0.243
	Often	0.818	0.715	0.474	0.085	0.820	0.715	0.478	0.087
	Frequently	0.106	0.093	0.459	0.816	0.119	0.104	0.465	0.798
Lifetime depression	No								
	Yes	0.682	0.596	0.199	0.001	0.689	0.601	0.198	0.001
Taking medication for anxiety, depression or stress	No								
	Yes	0.480	0.419	0.151	0.002	0.497	0.433	0.154	0.001

Dissatisfaction with unpaid work scale (Possible range 3-22)	Satisfied (range 3-7) Dissatisfied (range 8-12) Very dissatisfied (range 13-22)	 - 0.029 0.472	 - 0.025 0.412	 0.237 0.235	 0.902 0.045	 - 0.038 0.480	 - 0.033 0.419	 0.240 0.234	 0.873 0.040
Number of serious problems in very close persons. (Alcohol-drugs, psychological, physical, or disability)	None One Two Three Four	 0.049 0.451 0.824 0.766	 0.042 0.394 0.719 0.669	 0.179 0.166 0.288 0.639	 0.786 0.007 0.004 0.231	 0.041 0.435 0.824 0.769	 0.036 0.380 0.718 0.670	 0.179 0.161 0.293 0.644	 0.820 0.007 0.005 0.233
Satisfied with living together at home	Very satisfied Fairly satisfied Neither satisfied nor dissatisfied Fairly dissatisfied Very dissatisfied	 0.091 - 0.256 0.892 0.479	 0.080 - 0.224 0.780 0.418	 0.213 0.262 0.333 0.545	 0.668 0.328 0.008 0.380	 0.077 - 0.270 0.926 0.473	 0.067 - 0.235 0.807 0.412	 0.213 0.267 0.332 0.540	 0.716 0.312 0.005 0.381
Physical Health (SF-12) *Possible range 0-100*	Each point on SF12 subscale score.	- 0.034	- 0.030	0.009	<0.001	- 0.035	- 0.030	0.009	<0.001
Mental Health (SF-12) *Possible range 0-100*	Each point on SF12 subscale score.	- 0.055	- 0.048	0.005	<0.001	- 0.055	- 0.048	0.005	<0.001

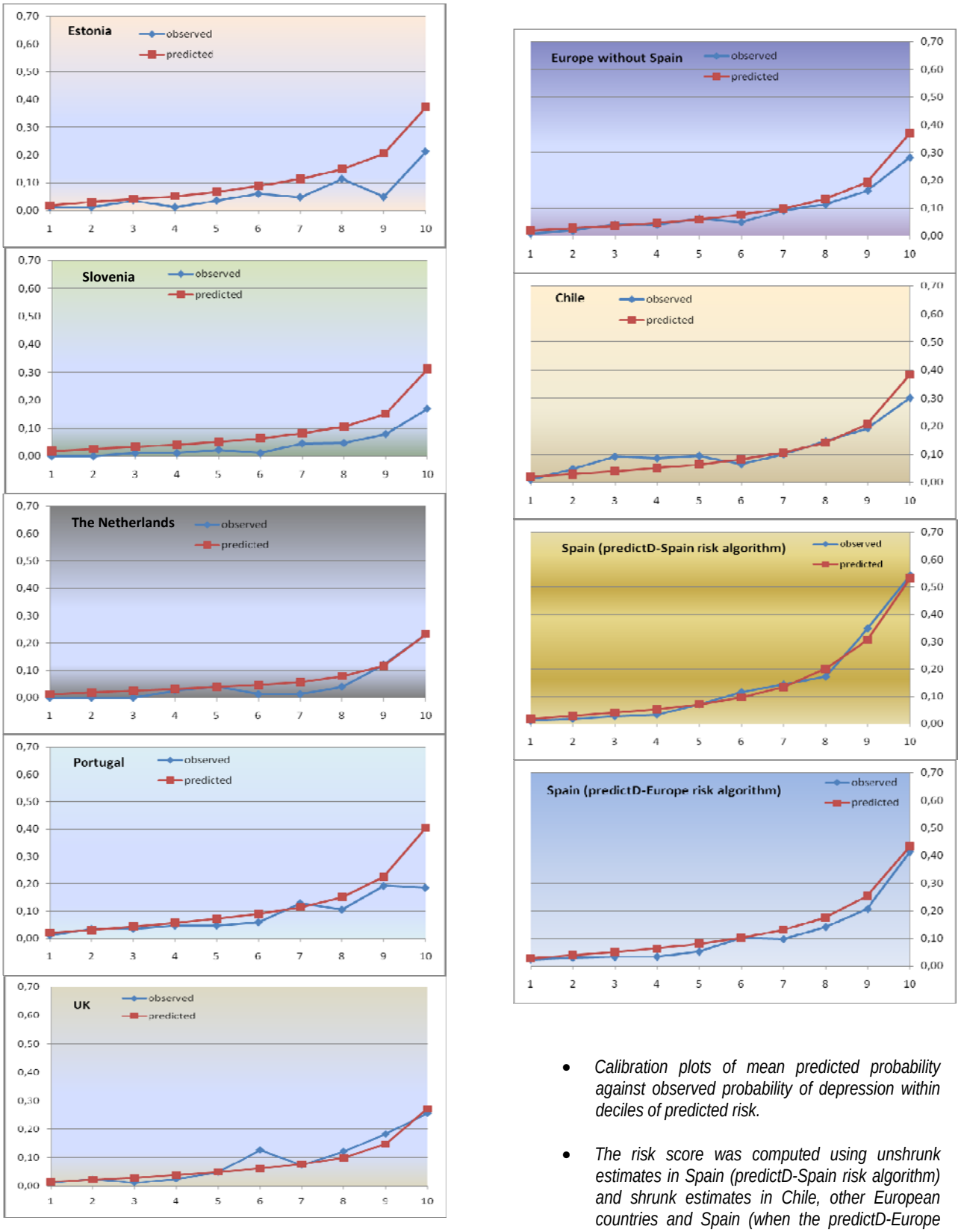
[1]Model derived in the ten imputed datasets weighting for the inverse probability of remaining in the follow-up to 12 months. [2]Coefficient after Copas Shrinkage.

Table 3: C-Index Statistic and Effect Sizes Computed Using Hedge's g.

Country (N)	[1]Spanish predictD risk algorithm		[2]European predictD risk algorithm	
	[®]C- Index (95% CI)	[®]Hedge's g (95% CI)	[®]C- Index (95% CI)	[®]Hedge's g (95% CI)
Spain (2787)	0.82 (0.79 – 0.84)	1.35 (1.21 – 1.48)	0.78 (0.75 – 0.81)	1.15 (1.02 – 1.29)
Spain without Málaga (2045)	0.82 (0.77 – 0.86)	1.36 (1.19 – 1.54)	0.78 (0.73 – 0.83)	1.14 (0.98 – 1.31)
Chile (1844)	0.70 (0.66 – 0.74)	0.77 (0.61 – 0.93)	0.71 (0.67 – 0.74)	0.85 (0.68 – 1.02)
United Kingdom (811)	0.76 (0.71 – 0.82)	1.02 (0.76 – 1.28)	0.76 (0.70 – 0.81)	1.02 (0.78 – 1.27)
Slovenia (866)	0.82 (0.75 – 0.88)	1.38 (0.98 – 1.78)	0.83 (0.77 – 0.89)	1.40 (1.06 – 1.75)
Portugal (844)	0.71 (0.65 – 0.77)	0.78 (0.54 – 1.03)	0.75 (0.69 – 0.80)	0.99 (0.73 – 1.25)
The Netherlands (731)	0.83 (0.77 – 0.90)	1.50 (1.09 – 1.92)	0.85 (0.80 – 0.90)	1.55 (1.25 – 1.85)
Estonia (823)	0.73 (0.66 – 0.80)	0.91 (0.58 – 1.23)	0.76 (0.69 – 0.83)	1.09 (0.76 – 1.42)
[#]Europe without Spain (4075)	0.76 (0.74 – 0.79)	1.04 (0.90 – 1.17)	0.79 (0.77- 0.82)	1.19 (1.05 – 1.32)

[1]The risk score was computed using unshrunk estimates in Spain and shrunk estimates in Chile and other European countries. [2]The risk score was computed using shrunk estimates in Spain, Chile and other European countries. [®]Average C-Index and Hedge's g over 10 imputed data sets.[#] UK + Slovenia + Portugal + The Netherlands + Estonia.

Figure 2: Calibration plots of the predictD-Spain risk algorithm in other countries and itself.



- *Calibration plots of mean predicted probability against observed probability of depression within deciles of predicted risk.*
- *The risk score was computed using unshrunk estimates in Spain (predictD-Spain risk algorithm) and shrunk estimates in Chile, other European countries and Spain (when the predictD-Europe risk algorithm was applied).*

Annex 1: Sensitivity analysis of predictD-Spain risk algorithm. (WEB ONLY)

Prognostic Factors	Levels in factor	Coefficient[1]	SE	Coefficient[2]	SE	Coefficient[3]	SE
Constant		1.763[c]	0.610	1.724[c]	0.591	1.948[b]	0.764
Province	Malaga						
	Granada	0.195	0.243	0.212	0.242	0.276	0.299
	Zaragoza	0.614	0.291	0.034	0.274	0.166	0.321
	Madrid	- 0.642[b]	0.291	- 0.652[b]	0.292	- 0.377	0.362
	La Rioja	- 0.378	0.317	- 0.289	0.300	- 0.329	0.398
	Balearic Isles	0.053	0.283	0.107	0.281	0.337	0.263
	Las Palmas	0.230	0.329	0.324	0.318	0.308	0.191
Age	Each year	- 0.027[d]	0.006	- 0.026 [d]	0.006	- 0.032[d]	0.008
Sex	Female						
	Male	- 1.724[c]	0.642	- 1.081[a]	0.568	- 1.128	0.752
Sex*Age interaction	Each year	0.031[c]	0.012	0.020[a]	0.011	0.022	0.014
Education	Beyond secondary education						
	Secondary education	0.377	0.281	0.297	0.263	0.580[a]	0.304
	Primary education	0.580[b]	0.259	0.461[a]	0.242	0.839[c]	0.285
	Incomplete primary education or illiterate	1.110[d]	0.293	0.981[d]	0.278	1.490[d]	0.369
Physical Childhood Abuse	Never						
	Seldom	0.633	0.440	0.616	0.414	0.470	0.465
	Sometimes	0.570[a]	0.321	0.466	0.311	0.320	0.300
	Often	0.625	0.577	0.932[a]	0.517	0.818[a]	0.474
	Frequently	0.359	0.408	0.273	0.398	0.106	0.459
Lifetime depression	No						
	Yes	0.716[d]	0.151	0.732 [d]	0.146	0.682[c]	0.199
Taking medication for anxiety, depression or stress	No						
	Yes	0.553[d]	0.158	0.590[d]	0.153	0.480[c]	0.151
Dissatisfaction with unpaid work scale (Possible range 3-22)	Satisfied (3-7)						
	Dissatisfied (8-12)	0.212	0.165	0.185	0.162	- 0.029	0.237
	Very dissatisfied (13-22)	0.566[b]	0.241	0.603[b]	0.199	0.472[b]	0.235
Number of serious problems in very close persons. (Alcohol-drugs, psychological, physical, or disability)	None						
	One	0.128	0.168	0.159	0.162	0.049	0.179
	Two	0.497[b]	0.205	0.580[c]	0.199	0.451[c]	0.166
	Three	0.647[b]	0.293	0.680[b]	0.278	0.824[c]	0.288
	Four	0.196	0.560	0.441	0.522	0.766	0.639
Satisfied with living together at home	Very satisfied						
	Fairly satisfied	0.072	0.162	0.070	0.160	0.091	0.213
	Neither satisfied nor dissatisfied	- 0.314	0.229	- 0.217	0.221	- 0.256	0.262
	Fairly dissatisfied	0.636[b]	0.321	0.631[b]	0.306	0.892[c]	0.333
	Very dissatisfied	0.269	0.475	0.438	0.462	0.479	0.545
Physical Health (SF-12) @	Each point on SF12 subscale score.	- 0.036[d]	0.006	- 0.035[d]	0.006	- 0.034[d]	0.009
Mental Health (SF-12) @	Each point on SF12 subscale score.	- 0.050[d]	0.007	- 0.052 [d]	0.006	- 0.055[d]	0.005
Random component (Health Center)		0.248[a]	0.124	0.253[b]	0.113	0.390[d]	0.085

[1]Model derived in participants with complete data (N = 2544); [2]Model derived in ten imputed datasets (N = 2787); [3]Model derived in ten imputed datasets weighting for the inverse probability of remaining in the follow-up to 12 months (N = 2787). [a] p<0.10; [b] p<0.05; [c] p<0.01; [d] p<0.001. @ Possible range 0-100

Annex 2: Box Examples of a range of predicted probabilities of depression at baseline. (WEB ONLY)

- Case 1: Risk score (predicted probability of depression) at baseline 1.4%
  - A man of 47 years living in Granada
  - Education beyond high school
  - No personal history of depression
  - Never suffered physical childhood abuse
  - Does not take medication for anxiety, depression or stress
  - *SF 12 mental score 62
  - *SF12 physical scale score 44
  - No serious problems in very close persons
  - Satisfied with unpaid work
  - Very satisfied with living together at home
- Case 2: Risk score 8.5% (3.1%)
  - A man of 55 years living in Zaragoza
  - Secondary education
  - No personal history of depression
  - Never suffered physical childhood abuse
  - Does not take medication for anxiety, depression or stress
  - SF 12 mental score 30
  - SF12 physical scale score 56
  - One serious problem in very close persons
  - Satisfied with unpaid work
  - Fairly satisfied with living together at home
- Case 3: Risk score 24.4% (12.5%)[1]
  - A woman of 60 years living in Las Palmas
  - Primary education
  - No personal history of depression
  - Never suffered physical childhood abuse
  - Does not take medication for anxiety, depression or stress
  - SF12 mental score 40.6
  - SF12 physical scale score 29.7
  - One serious problem in very close persons
  - Dissatisfied with unpaid work
  - Neither satisfied nor dissatisfied with living together at home
- Case 4: Risk score 39.3% (11.4%)
  - A woman of 62 years living in Balearic Isles
  - Primary education
  - Sometimes suffered physical childhood abuse
  - No personal history of depression
  - Taking medication for anxiety, depression or stress
  - SF12 mental score 45.6
  - SF12 physical scale score 52.8
  - Two serious problems in very close persons
  - Very dissatisfied with unpaid work
  - Fairly dissatisfied with living together at home
- Case 5: Risk score 93.5% (18.4%)[2]
  - A woman of 45 years living in Malaga
  - Incomplete primary education
  - Often suffered physical childhood abuse
  - Personal history of depression
  - Taking medication for anxiety, depression or stress
  - SF12 mental score 29.2
  - SF12 physical scale score 25.7
  - Three serious problems in very close persons
  - Very dissatisfied with unpaid work
  - Very dissatisfied with living together at home

Mean (SD) Short Form 12 (SF-12) mental and physical subscale scores for Spain were 47.1 (12.4) and 43.8 (11.4), respectively. High scores indicate good health/well-being. Scores in parentheses correspond to eliminating dissatisfaction with unpaid work and living together at home, perception of serious problems in close persons and correcting SF-12 physical and mental health scores to the Spanish mean. [1]Perception of serious problems in close persons did not change; [2]fairly dissatisfied with living together at home changed to neither satisfied nor dissatisfied and stopped taking medication for anxiety, depression or stress.