Development and Validation of an International Risk Prediction Algorithm for Episodes of Major Depression in General Practice Attendees

The PredictD Study

Michael King, MD, PhD; Carl Walker, BSc, PhD; Gus Levy, MSc; Christian Bottomley, PhD; Patrick Royston, DSc; Scott Weich, MBBS, DM; Juan Angel Bellon-Saameno, MD, PhD; Berta Moreno, PhD; Igor Svab, MD, PhD; Danica Rotar, MD, MSc; J. Rifel, MD; Heidi-Ingrid Maaroos, MD, PhD; Anu Aluoja, PhD; Ruth Kaldal, MD, DrMedSci; Jan Neeleman, MD, PhD; Miriam I. Geerlings, PhD; Miguel Xavier, MD, PhD; Idalmiro Carraça, MD, MSc; Manuel Gonçalves-Pereira, MD, MSc; Benjamin Vicente, MD, PhD; Sandra Saldívar, PhD; Roberto Melipillan, MSc; Francisco Torres-Gonzalez, MD, PhD; Irwin Nazareth, MBBS, PhD

Context: Strategies for prevention of depression are hindered by lack of evidence about the combined predictive effect of known risk factors.

Objectives: To develop a risk algorithm for onset of major depression.

Design: Cohort of adult general practice attendees followed up at 6 and 12 months. We measured 39 known risk factors to construct a risk model for onset of major depression using stepwise logistic regression. We corrected the model for overfitting and tested it in an external population.

Setting: General practices in 6 European countries and in Chile.

Participants: In Europe and Chile, 10,045 attendees were recruited April 2003 to February 2005. The algorithm was developed in 5,216 European attendees who were not depressed at recruitment and had follow-up data on depression status. It was tested in 1,732 patients in Chile who were not depressed at recruitment.

Main Outcome Measure: DSM-IV major depression.

Results: Sixty-six percent of people approached participated, of whom 89.5% participated again at 6 months and 85.9%, at 12 months. Nine of the 10 factors in the risk algorithm were age, sex, educational level achieved, results of lifetime screen for depression, family history of psychological difficulties, physical health and mental health subscale scores on the Short Form 12, unsupported difficulties in paid or unpaid work, and experiences of discrimination. Country was the tenth factor. The algorithm’s average C index across countries was 0.790 (95% confidence interval [CI], 0.767-0.813). Effect size for difference in predicted log odds of depression between European attendees who became depressed and those who did not was 1.28 (95% CI, 1.17-1.40). Application of the algorithm in Chilean attendees resulted in a C index of 0.710 (95% CI, 0.670-0.749).

Conclusion: This first risk algorithm for onset of major depression functions as well as similar risk algorithms for cardiovascular events and may be useful in prevention of depression.

Arch Gen Psychiatry. 2008;65(12):1368-1376

Reducing the prevalence of depression is a public health challenge for the 21st century. Depression occurs in up to a quarter of general practice attendees, relapse 10 years from first presentation is frequent, and both residual disability and premature mortality are common. Low socioeconomic status and female sex are the 2 most consistently identified risk factors. Socioeconomic risk factors include low income and financial strain, unemployment, work stress, social isolation, and poor housing. Other factors, such as a family history of depression, play a part. Additional risk factors identified in adult general practice populations are negative life events, poor physical health, poor marital or other interpersonal relationships, a partner or spouse’s poor health, and problems with alcohol. Poor social support, loneliness, and physical disability appear to be particular risks for older adults. Estimating overall risk across a range of likely risk factors is essential in efforts to prevent depression. However, effective strategies for prevention are hindered by lack of evidence about the combined effect of known risk factors. Our objectives were...
to develop a risk algorithm for the onset of major depression in European general practice attendees and test its predictive power in a non-European setting. We modeled our approach on risk indexes for cardiovascular disease, which provide a percentage risk estimate over a given period.

## METHODS

### STUDY SETTING AND DESIGN

We undertook a prospective study to develop a quantitative risk prediction algorithm for the onset of major depression over 12 months in general practice attendees. Given the relapsing and remitting nature of major depression, 12 months was considered a useful period for prediction in this setting. The method, described in detail elsewhere, was approved by ethical committees in each country. The study was conducted in 6 European centers: (1) 25 general practices in the Medical Research Council General Practice Research Framework in the United Kingdom; (2) 9 large primary care centers in Andalucía, Spain; (3) 74 general practices nationwide in Slovenia; (4) 23 general practices nationwide in Estonia; (5) 7 large general practice centers near Utrecht, the Netherlands; and (6) 2 large primary care centers in the Lisbon area of Portugal. We assessed the external validity of the risk algorithm in patients attending 78 general practices in Concepción and Talcahuano in the Eighth Region of Chile. General practices covered urban and rural populations with considerable socioeconomic variation.

### STUDY PARTICIPANTS

Consecutive attendees aged 18 to 75 years were recruited in Europe between April 2003 and September 2004 and in Chile between October 2003 and February 2005. Exclusion criteria were an inability to understand one of the main languages involved, psychosis, dementia, and incapacitating physical illness. Recruitment differed slightly in each country because of local service preferences. In the United Kingdom and the Netherlands, researchers spoke to patients waiting to see practice staff. In the remaining European countries, physicians introduced the study before contact with researchers. In Chile, attendees were stratified on age and sex according to figures for the populations served by each health center and participants selected randomly within each stratum. Participants gave informed consent and undertook a research evaluation within 2 weeks.

### MAJOR DEPRESSION AND KNOWN RISK FACTORS

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). We selected risk factors to cover all important areas identified in a systematic review of the literature. Where possible, we used standardized self-report measures. Questions adapted from standardized questionnaires or developed for the study were evaluated for test-retest reliability in 285 general practice attendees evenly recruited across the European countries before the main study began. Each instrument or question not available in the relevant languages was translated from English and back-translated by professional translators. The 39 candidate risk factors are numbered, and those subjected to test-retest reliability are italicized.

1. Age
2. Sex
3. Occupation
4. Educational level
5. Marital status
6. Employment status
7. Ethnicity
8. Owner-occupier accommodation
9. Living alone or with others
10. Born in country of residence or abroad
11. Satisfaction with living conditions
12. Long-standing physical illness
13. Lifetime depression
14. Feeling in control
15. Paid or unpaid work
16. Experiencing difficulties
17. Support from family and friends
18. Financial strain
19. Physical and mental health
20. Alcohol problems
21. Alcohol use
22. Alcohol use disorders
23. Drug use
24. Drug use disorders
25. Substance misuse
26. Substance misuse problems
27. Serious psychological problems
28. Psychotic disorders
29. Dysthymia
30. Anxiety symptoms
31. Panic symptoms
32. Suicide in first-degree relatives
33. Anxiety
34. Panic symptoms
35. History of serious psychological problems
36. Suicide in first-degree relatives
37. Major life events
38. Historical events
39. Major life events
40. Historical events
41. Life events
42. Historical events
43. Life events
44. Historical events
45. Life events
46. Historical events
47. Life events
48. Historical events
49. Life events
50. Historical events
51. Life events
52. Historical events
53. Life events
54. Historical events
55. Life events
56. Historical events
57. Life events
58. Historical events
59. Life events
60. Historical events
61. Life events
62. Historical events
63. Life events
64. Historical events
65. Life events
66. Historical events
67. Life events
68. Historical events
69. Life events
70. Historical events
71. Life events
72. Historical events
73. Life events
74. Historical events
75. Life events
76. Historical events
77. Life events
78. Historical events
79. Life events
80. Historical events
81. Life events
82. Historical events
83. Life events
84. Historical events
85. Life events
86. Historical events
87. Life events
88. Historical events
89. Life events
90. Historical events
91. Life events
92. Historical events
93. Life events
94. Historical events
95. Life events
96. Historical events
97. Life events
98. Historical events
99. Life events
100. Historical events

### STATISTICAL ANALYSIS

All analyses and data imputation were performed using Stata release 9.3. We included only patients without major depression at baseline. Participants with missing depression diagnoses at any point were excluded as this outcome was central to our risk estimation.

### Data Imputation

Missing data in candidate risk factors were imputed using the method of chained equations, implemented in the Stata command ice. We imputed 10 data sets and obtained combined estimates.
Model Building

We built a risk model using the 39 risk factors described earlier and country of residence of each participant. We developed this model in the imputed data using stepwise logistic regression with robust standard errors to adjust for general practice clustering. We used a conservative threshold for inclusion of \( P < 0.01 \) to produce a stable model and minimize the degree of overfitting. We retained age and sex in all regression models because of their well-known associations with onset of depression.\(^{37,38}\) We also retained country because of an a priori assumption of clustering within country. Multivariable fractional polynomial analysis was used to assess possible nonlinear effects of continuous predictors. The resulting risk score provides a predicted probability of depression over 12 months.

Internal Validation

We calculated the C index\(^{39}\) to estimate the discriminative power of the final model in each European country and all European countries combined. We used a calculation proposed by Copas\(^{40}\) to adjust for overfitting of our prediction model. This involves computing a shrinkage factor that is applied to the model coefficients to provide more accurate predictions when the risk algorithm is applied in new settings. To deal with the overfitting that arises through variable selection, we computed the shrinkage factor based on the initial model including all 39 variables. 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. We calculated effect sizes using the Hedges g\(^{41}\) for the difference in log odds of predicted probability between patients who were later observed to be depressed and those who were not. Finally, we report the threshold values of risk score, and the associated sensitivity, for a range of specificity that would be practical (minimizing false positives) when using the instrument in a clinical setting.

External Validation

We used the C index, Hedges g, and a comparison of predicted vs observed probability of depression to evaluate the performance of the predictD model in the Chilean data.

In the 7 countries, 10 045 people took part (Figure 1). Response to recruitment was high in Portugal (76%), Es-

---

**RESULTS**

In the 7 countries, 10 045 people took part (Figure 1). Response to recruitment was high in Portugal (76%), Es-

---

©2008 American Medical Association. All rights reserved.

Downloaded From: by a Non-Human Traffic (NHT) User on 10/26/2018
Table 1. Demographic Characteristics and Response to Follow-up

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>All European Countries</th>
<th>United Kingdom</th>
<th>Spain</th>
<th>Slovenia</th>
<th>Portugal</th>
<th>The Netherlands</th>
<th>Estonia</th>
<th>Chile</th>
</tr>
</thead>
<tbody>
<tr>
<td>European</td>
<td>6190 (100)</td>
<td>1131 (18.3)</td>
<td>1006 (16.3)</td>
<td>1048 (16.9)</td>
<td>1005 (16.2)</td>
<td>1077 (17.4)</td>
<td>923 (14.9)</td>
<td>2133</td>
</tr>
<tr>
<td>Age, y, mean (SD)</td>
<td>48.9 (15.5)</td>
<td>52.2 (14.7)</td>
<td>50.8 (15.5)</td>
<td>48.8 (14.5)</td>
<td>50.2 (15.4)</td>
<td>48.9 (14.9)</td>
<td>41.6 (16.0)</td>
<td>47 (15.7)</td>
</tr>
<tr>
<td>Female</td>
<td>4081 (65.9)</td>
<td>750 (66.3)</td>
<td>689 (68.5)</td>
<td>660 (63)</td>
<td>649 (64.6)</td>
<td>668 (62)</td>
<td>665 (72)</td>
<td>1522 (71.4)</td>
</tr>
<tr>
<td>Marital status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married or living together</td>
<td>4491 (72.6)</td>
<td>844 (74.6)</td>
<td>708 (70.4)</td>
<td>732 (69.9)</td>
<td>750 (74.6)</td>
<td>827 (76.8)</td>
<td>630 (68.3)</td>
<td>1228 (57.6)</td>
</tr>
<tr>
<td>Separated or divorced</td>
<td>421 (6.8)</td>
<td>100 (8.8)</td>
<td>49 (4.9)</td>
<td>56 (5.3)</td>
<td>69 (6.9)</td>
<td>64 (5.9)</td>
<td>83 (9)</td>
<td>179 (8.4)</td>
</tr>
<tr>
<td>Single</td>
<td>872 (14.1)</td>
<td>121 (10.7)</td>
<td>181 (18)</td>
<td>152 (14.5)</td>
<td>132 (13.1)</td>
<td>121 (11.2)</td>
<td>165 (17.9)</td>
<td>521 (24.4)</td>
</tr>
<tr>
<td>Widowed</td>
<td>383 (6.2)</td>
<td>65 (5.8)</td>
<td>67 (6.7)</td>
<td>105 (10)</td>
<td>53 (5.3)</td>
<td>48 (4.5)</td>
<td>45 (4.9)</td>
<td>205 (9.6)</td>
</tr>
<tr>
<td>Missing</td>
<td>23 (0.4)</td>
<td>1 (0.1)</td>
<td>1 (0.1)</td>
<td>3 (0.3)</td>
<td>3 (0.3)</td>
<td>1 (0.1)</td>
<td>1 (0.1)</td>
<td>0</td>
</tr>
<tr>
<td>Household status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not living alone</td>
<td>5483 (88.6)</td>
<td>981 (86.7)</td>
<td>948 (94.2)</td>
<td>915 (87.3)</td>
<td>929 (92.4)</td>
<td>894 (83)</td>
<td>816 (88.4)</td>
<td>2039 (95.6)</td>
</tr>
<tr>
<td>Living alone</td>
<td>707 (11.4)</td>
<td>150 (13.3)</td>
<td>58 (5.8)</td>
<td>133 (12.7)</td>
<td>76 (7.6)</td>
<td>183 (17)</td>
<td>107 (11.6)</td>
<td>94 (4.4)</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Higher education</td>
<td>1879 (30.4)</td>
<td>448 (39.6)</td>
<td>135 (13.4)</td>
<td>181 (17.3)</td>
<td>129 (12.8)</td>
<td>458 (42.5)</td>
<td>528 (57.2)</td>
<td>75 (3.5)</td>
</tr>
<tr>
<td>Secondary</td>
<td>2038 (32.9)</td>
<td>465 (41.1)</td>
<td>215 (21.4)</td>
<td>385 (36.7)</td>
<td>182 (18.1)</td>
<td>508 (47.2)</td>
<td>283 (30.7)</td>
<td>791 (37.1)</td>
</tr>
<tr>
<td>Primary/no education</td>
<td>1767 (28.6)</td>
<td>25 (2.2)</td>
<td>656 (65.2)</td>
<td>235 (23.6)</td>
<td>32 (3.2)</td>
<td>0</td>
<td>1 (0.1)</td>
<td>267 (12.5)</td>
</tr>
<tr>
<td>Trade/other</td>
<td>451 (7.3)</td>
<td>171 (15.1)</td>
<td>0</td>
<td>247 (23.6)</td>
<td>32 (3.2)</td>
<td>0</td>
<td>1 (0.1)</td>
<td>267 (12.5)</td>
</tr>
<tr>
<td>Missing</td>
<td>55 (0.9)</td>
<td>22 (1.9)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>33 (3.1)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Employment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed/full-time student</td>
<td>3256 (52.6)</td>
<td>574 (50.8)</td>
<td>349 (34.7)</td>
<td>563 (53.7)</td>
<td>486 (48.4)</td>
<td>602 (55.9)</td>
<td>682 (73.9)</td>
<td>749 (35.1)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>1313 (21.2)</td>
<td>309 (27.3)</td>
<td>107 (10.6)</td>
<td>165 (15.7)</td>
<td>96 (9.6)</td>
<td>373 (34.6)</td>
<td>263 (28.5)</td>
<td>29 (1.4)</td>
</tr>
<tr>
<td>Unable to work</td>
<td>143 (2.3)</td>
<td>27 (2.4)</td>
<td>3 (0.3)</td>
<td>3 (0.3)</td>
<td>0</td>
<td>56 (5.2)</td>
<td>54 (5.8)</td>
<td>3 (0.1)</td>
</tr>
<tr>
<td>Retired/looking after family</td>
<td>2669 (43.7)</td>
<td>450 (39.8)</td>
<td>493 (49)</td>
<td>409 (39)</td>
<td>372 (37)</td>
<td>358 (33.2)</td>
<td>187 (20.3)</td>
<td>1072 (50.3)</td>
</tr>
<tr>
<td>Missing</td>
<td>43 (0.7)</td>
<td>0</td>
<td>1 (0.1)</td>
<td>7 (0.7)</td>
<td>1 (0.1)</td>
<td>34 (3.2)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Professional status</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1313 (21.2)</td>
<td>309 (27.3)</td>
<td>107 (10.6)</td>
<td>165 (15.7)</td>
<td>96 (9.6)</td>
<td>373 (34.6)</td>
<td>263 (28.5)</td>
<td>29 (1.4)</td>
</tr>
<tr>
<td>Missing</td>
<td>143 (2.3)</td>
<td>27 (2.4)</td>
<td>3 (0.3)</td>
<td>3 (0.3)</td>
<td>0</td>
<td>56 (5.2)</td>
<td>54 (5.8)</td>
<td>3 (0.1)</td>
</tr>
<tr>
<td>Born in country of residence</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>5655 (91.4)</td>
<td>1054 (93.2)</td>
<td>955 (94.9)</td>
<td>834 (79.6)</td>
<td>973 (96.8)</td>
<td>997 (92.6)</td>
<td>842 (91.2)</td>
<td>2122 (99.5)</td>
</tr>
<tr>
<td>Missing</td>
<td>87 (1.4)</td>
<td>3 (0.3)</td>
<td>3 (0.3)</td>
<td>4 (0.4)</td>
<td>0</td>
<td>24 (2.2)</td>
<td>53 (5.7)</td>
<td>4 (0.2)</td>
</tr>
<tr>
<td>Ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White/European</td>
<td>5988 (96.7)</td>
<td>1055 (93.3)</td>
<td>994 (98.8)</td>
<td>1042 (99.4)</td>
<td>992 (98.7)</td>
<td>983 (91.3)</td>
<td>922 (99.9)</td>
<td>0</td>
</tr>
<tr>
<td>Missing</td>
<td>72 (1.2)</td>
<td>39 (3.4)</td>
<td>1 (0.1)</td>
<td>2 (0.2)</td>
<td>0</td>
<td>30 (2.8)</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>6-mo Response</td>
<td>5538 (89.5)</td>
<td>987 (87.3)</td>
<td>794 (78.9)</td>
<td>963 (91.9)</td>
<td>889 (88.5)</td>
<td>1035 (96.1)</td>
<td>870 (94.3)</td>
<td>1904 (89.3)</td>
</tr>
<tr>
<td>12-mo Response</td>
<td>5319 (85.9)</td>
<td>965 (85.3)</td>
<td>731 (72.7)</td>
<td>927 (88.5)</td>
<td>864 (86)</td>
<td>988 (91.7)</td>
<td>844 (91.4)</td>
<td>1748 (82)</td>
</tr>
</tbody>
</table>

©2008 American Medical Association. All rights reserved.
characteristics (sex, age, education, results of lifetime depression screen, family history of psychological difficulties); 4, current status (Short Form 12 physical health subscale score, Short Form 12 mental health subscale score, unsupported difficulties in paid and/or unpaid work, and discrimination) (Table 2); and 1 concerned country. Examination of the risk model derived in each of the 10 imputed data sets revealed that it was stable in terms of the variables selected. Besides country, age, and sex, 5 variables (results of lifetime depression screen, family history of psychological difficulties, Short Form 12 physical health subscale score, Short Form 12 mental health subscale score, and unsupported difficulties in paid and/or unpaid work) were consistently selected in each of the imputed data sets. Discrimination was selected in 7 data sets and education, in 4 data sets. Three other variables that did not reach the full model were also selected in a number of imputed data sets. These were PHQ panic syndrome (6 sets), childhood sexual abuse (1 set), and PHQ anxiety syndrome (1 set).

We compared a model with interactions between sex and the remaining risk factors to the model with no interactions. A Wald test provided no evidence to suggest that including interaction terms improves the model fit (P value = .27; \( \chi^2 \) = 19.06). There was also no evidence for including interactions with age (P value = .21; \( \chi^2 \) = 20.19).

The average C index across countries for predicted probability of depression at 6 or 12 months in all 6 European countries was 0.790 (Table 3). The model was most predictive in the Netherlands (0.852) and least predictive in Portugal (0.747). The effect size for the difference in log odds of predicted probability between attendees in Europe who subsequently became depressed and those who did not was 1.28 (95% confidence interval [CI], 1.17-1.40) (Table 4). Again, the model discriminated best in the Netherlands (1.55) and least well in Portugal (0.99). To examine the fit of the model, we divided the European sample into deciles of predicted probability on the predictD score. Within each decile, we plotted mean predicted probability vs observed probability of depression (Figure 2A). Figure 2A shows that the incidence of major depression in the highest decile of risk score in Europe was more than 30% in contrast to the overall incidence of 7.7%. Examples of the kinds of participants scoring at increasing levels of predicted depression.

### Table 2. PredictD Model Derived in the Imputed Data Sets

<table>
<thead>
<tr>
<th>Prognostic Factor</th>
<th>Levels in Factor</th>
<th>Coefficient</th>
<th>SE</th>
<th>Coefficient After Copas Shrinkage</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td></td>
<td>1.543</td>
<td>0.439</td>
<td>1.155</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Age</td>
<td>Each year</td>
<td>-0.005</td>
<td>0.005</td>
<td>-0.005</td>
<td>.25</td>
</tr>
<tr>
<td>Sex</td>
<td>M</td>
<td>-0.245</td>
<td>0.138</td>
<td>-0.212</td>
<td>.07</td>
</tr>
<tr>
<td>Education</td>
<td>Beyond secondary education</td>
<td>0.103</td>
<td>0.128</td>
<td>0.089</td>
<td>.42</td>
</tr>
<tr>
<td></td>
<td>Secondary education</td>
<td>0.472</td>
<td>0.157</td>
<td>0.409</td>
<td>.003</td>
</tr>
<tr>
<td></td>
<td>Primary or no education</td>
<td>0.653</td>
<td>0.210</td>
<td>0.566</td>
<td>.002</td>
</tr>
<tr>
<td>Difficulties in paid and unpaid work</td>
<td>No difficulties or often supported</td>
<td>0.423</td>
<td>0.114</td>
<td>0.366</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Physical health</td>
<td>Each point on SF-12 subscale score; possible range, 0-100</td>
<td>-0.034</td>
<td>0.005</td>
<td>-0.030</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Mental health</td>
<td>Each point on SF-12 subscale score; possible range, 0-100</td>
<td>-0.064</td>
<td>0.005</td>
<td>-0.055</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>First-degree relative with emotional problem</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Discrimination</td>
<td>None</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>In 1 area</td>
<td>0.186</td>
<td>0.220</td>
<td>0.161</td>
<td>.40</td>
</tr>
<tr>
<td></td>
<td>In &gt;1 area</td>
<td>0.850</td>
<td>0.235</td>
<td>0.736</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Lifetime depression</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>United Kingdom</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Spain</td>
<td>0.266</td>
<td>0.205</td>
<td>0.230</td>
<td>.20</td>
</tr>
<tr>
<td></td>
<td>Slovenia</td>
<td>-0.841</td>
<td>0.193</td>
<td>-0.729</td>
<td>&lt;.001</td>
</tr>
<tr>
<td></td>
<td>Estonia</td>
<td>-0.540</td>
<td>0.196</td>
<td>-0.467</td>
<td>.006</td>
</tr>
<tr>
<td></td>
<td>The Netherlands</td>
<td>-0.133</td>
<td>0.220</td>
<td>-0.115</td>
<td>.54</td>
</tr>
<tr>
<td></td>
<td>Portugal</td>
<td>-0.195</td>
<td>0.180</td>
<td>-0.169</td>
<td>.28</td>
</tr>
</tbody>
</table>

Abbreviation: SF-12, Short Form 12.

### Table 3. C Index Statistics for Each Country

<table>
<thead>
<tr>
<th>Country</th>
<th>C Indexa (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>United Kingdom</td>
<td>0.756 (0.705-0.808)</td>
</tr>
<tr>
<td>Spain</td>
<td>0.793 (0.746-0.840)</td>
</tr>
<tr>
<td>Slovenia</td>
<td>0.833 (0.775-0.891)</td>
</tr>
<tr>
<td>Estonia</td>
<td>0.761 (0.690-0.833)</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>0.852 (0.799-0.905)</td>
</tr>
<tr>
<td>Portugal</td>
<td>0.747 (0.693-0.800)</td>
</tr>
<tr>
<td>Mean over all countries</td>
<td>0.790 (0.767-0.813)</td>
</tr>
</tbody>
</table>

a The C index is also known as the area under the relative operating characteristic curve of sensitivity against 1 - specificity. A perfect test has a C index of 1.00 while a test that performs no better than chance has a C index of 0.5.41

b Average C index over 10 imputed data sets.
Cumulative 12-month incidence of major depression in Chilean general practice attendees was 11.6%. There were no missing data in Chile on any of the 10 risk factors of the final European model. The model was validated using data on 1732 attendees who were not depressed at recruitment (Figure 1). The Copas shrinkage factor for the European model was 0.866, suggesting a degree of overfitting. We evaluated the prediction algorithm’s external validity in the Chilean data using the shrinked regression coefficients derived in the European data and comparing predicted with observed probability. Because country is included in the model, it was necessary to base risk scores in Chile on an assumed country effect. Using the coefficient for Spain gave the best concordance between predicted and observed probability of major depression in Chile (Figure 2C and D) and reflects the prevalence of depression in Chile being more similar to Spain than Slovenia. The C index for the risk algorithm in Chile was 0.710 (95% CI, 0.670-0.749). This lower degree of discrimination can also be seen in the estimates of specificity and sensitivity in Chile (Table 5).

**Table 4. Effect Sizes Computed Using Hedges g**

<table>
<thead>
<tr>
<th>Country</th>
<th>Hedges g (95% Confidence Interval)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All European</td>
<td>1.28 (0.97-1.41)</td>
</tr>
<tr>
<td>United Kingdom</td>
<td>1.02 (0.69-1.04)</td>
</tr>
<tr>
<td>Spain</td>
<td>1.09 (0.97-1.15)</td>
</tr>
<tr>
<td>Slovenia</td>
<td>1.06 (0.76-1.23)</td>
</tr>
<tr>
<td>Estonia</td>
<td>1.02 (0.57-1.48)</td>
</tr>
<tr>
<td>The Netherlands</td>
<td>0.99 (0.73-1.23)</td>
</tr>
<tr>
<td>Portugal</td>
<td>0.87 (0.69-1.04)</td>
</tr>
<tr>
<td>Chile</td>
<td>0.87 (0.69-1.04)</td>
</tr>
</tbody>
</table>

*Predicted probabilities were logarithmically transformed and compared between depressed and nondepressed individuals over the subsequent 12 months. The Hedges g is preferred to the Cohen d where the sizes of each group (depressed/nondepressed) are markedly unequal. The risk score was computed using unshrunk estimates in Europe and shrunk estimates in Chile.*

**Figure 2.** Plots of mean predicted probability against observed probability of depression within deciles of predicted risk. A, Model fitted in the European data using unshrunk coefficients. B-D, European model fitted in the Chilean data using shrunk coefficients. Risk scores are based on the average of the 6 European country coefficients (the UK coefficient is 0). (B) and the coefficients for Spain and Slovenia, respectively (C and D). Each point on the graphs represents a decile of risk.

**Figure 3.** Examples of a range of predicted probabilities of depression at baseline. Mean (SD) Short Form 12 (SF-12) mental and physical subscale scores for Europe were 48.5 (10.6) and 44.2 (11.0), respectively. High scores indicate good health/well-being. Scores in parentheses correspond to eliminating discrimination and work difficulties and correcting SF-12 physical and mental health scores to the European mean (see text).
We have developed a risk score from recognized risk factors for major depression over 12 months in 5216 general practice attendees in Europe and validated its use in 1732 attendees in Chile. To our knowledge, this is the first risk algorithm to be developed simultaneously in a number of cultures in one continent for prediction of new episodes of major depression in a general medical setting and validated in another continent. This is arguably the most rigorous test that can be applied to a prediction tool. We emphasize that our study was not about recognition of current depression, nor was it about a search for new risk factors; these are well known. Nor was it about developing a prognostic tool for outcome of depression, which has been achieved recently.42 Our aim was to determine the key factors in a valid clinical prediction algorithm. Five risk factors are immutable (age, sex, educational level achieved, results of lifetime screen for depression, and family history of depression) and 4 are mutable factors relating to current status (Short Form 12 physical health and mental health subscale scores, unsupported difficulties in paid and/or unpaid work, and experiences of discrimination). The C index provides a standardized way of comparing the discriminative power of tests that use different measurement units in different settings.43 The predictD risk score compares favorably with a risk index for cardiovascular events developed in 12 European cohorts44 that reported C indexes between 0.71 and 0.82.

Our calculation of a shrinkage factor provides a measure of overfitting in the European data and allows for its adjustment in predicting risk of depression in new settings. External validation and shrinkage for overfitting are often not undertaken.45,46

When the algorithm is applied in a country outside of the 6 participating European countries, we recommend that either the average country coefficient be used (Figure 2) or the coefficient for the European country that most closely matches the annual incidence of depression (if known) in the new setting.

Despite the advantages of a cross-national study and an external population in which to validate the risk algorithm, there are limitations to our study. Lower recruitment rates occurred in the United Kingdom and the Netherlands, possibly because the study was not so obviously endorsed by physicians. However, response to follow-up in all countries was high. There were differences in the geographical distribution of general practices in each country, which reflected the varying networks available to the centers. Follow-up was relatively short but in keeping with what would be acceptable for prediction of depression in general practice. People from nonwhite ethnic minorities were relatively underrepresented. Although our risk factors are based on self-report, we used standardized instruments, and nonstandardized questions were tested for reliability. Our data imputation retained power and reduced bias. Although 24% of European participants had missing data on at least 1 risk factor, as we reported, missing data were less than 3.0% on 38 of the 39 factors. Finally, we stress that our study did not aim to provide insights into pathways to depression. Rather, we aimed to develop a predictive tool for the detection of DSM-IV major depressive disorders prior to onset. Such an instrument could then be used for prevention of depression in a manner similar to an existing instrument used in cardiovascular prevention in...
family practice settings. Some of our risk factors in the predictD algorithm may be mediators on the pathway to depression. For example, childhood experiences of emotional abuse may make depression at an early age more likely, but once it has occurred, this will show up most parsimoniously in the algorithm as lifetime history.

Our study does not address how the risk algorithm for depression might best be implemented in general practice. However, the questions making up the algorithm are brief and easy to complete, and thus it has potential as a clinical tool for prediction of future episodes of depression in this setting (http://www.techflora.com/ucl). Our results expressed by the C index and effect sizes demonstrate a clear difference in risk between participants who became depressed and those who did not do so. In suggesting useful thresholds of sensitivity and specificity (Table 5), we have erred on the side of maximizing specificity at the cost of reduced sensitivity to minimize the workload for family physicians engaging with false positives. We would recommend setting specificity at 80% to 85% (risk score, ≥10.6%) to contain the workload of the physician, albeit at the cost of missing a proportion of future major depressive episodes.

Patients identified as being at risk on screening can be flagged on practice computers to alert physicians when they consult. Recognition of those at risk may be helpful when it leads to watchful waiting or active support, such as restarting treatment in patients with a history of depression. Advising patients on the nature of depression or on brief cognitive behavior strategies they might undertake to reduce their risk could also be envisaged. The application of such strategies to the prevention of depression in primary care would benefit from further evaluation. Four of the 10 factors were open to intervention/change and the impact of such change is shown in Figure 3. Efforts to reduce the incidence of depression might usefully address these factors through a combination of physical, psychological, and medical interventions. However, this implies that the risk model has a causal interpretation, something that our study cannot demonstrate. It also does not mean that when immutable factors predominate in any particular individual there can be no recourse to prevention. The introduction of brief cognitive behavior skills might be a preventive strategy regardless of the risk factors implicated. The same is true for starting or restarting antidepressant medication use.

CONCLUSIONS

This risk algorithm for major depression compares favorably with risk algorithms for prediction of cardiovascular events and may be useful in prevention of depression in general medical settings.

Submitted for Publication: December 14, 2007; final revision received April 1, 2008; accepted May 12, 2008.

Author Affiliations: Departments of Mental Health Sciences (Drs King and Walker and Mr Levy) and Primary Care and Population Sciences (Drs Bottomley and Nazareth), University College London, Medical Research Council General Practice Research Framework (Mr Levy and Dr Nazareth), and Medical Research Council Clinical Trials Unit (Dr Royston), London, and Health Sciences Research Institute, University of Warwick, Coventry (Dr Weich), England; Department of Preventive Medicine, El Palo Health Centre, Malaga (Dr Bello-Saño), and Department of Psychiatry, University of Granada, Granada (Drs Moreno and Torres-Gonzalez), Spain; Department of Family Medicine, University of Ljubljana, Ljubljana, Slovenia (Drs Svab, Rotar, and Rifel); Faculty of Medicine, University of Tartu, Tartu, Estonia (Drs Maaroos, Aluoja, and Kalda); University Medical Center, Utrecht, the Netherlands (Drs Neelen and Geerlings); Faculdade Ciências Médicas, University of Lisbon (Drs Xavier and Gonçalves-Pereira), and Encarnação Health Centre (Dr Carraça), Lisbon, Portugal; and Departamento de Psiquiatría y Salud Mental, Universidad de Concepción, Concepción, Chile (Drs Vicente and Saldívar and Mr Melipillan).

Correspondence: Michael King, MD, PhD, Department of Mental Health Sciences, University College London Medical School, Royal Free Campus, Rowland Hill Street, London NW3 2PF, England (m.king@medsch.ucl.ac.uk).

Author Contributions: Dr King had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Financial Disclosure: None reported.

Funding/Support: 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), the Slovenian Ministry for Research (grant 4369-1027), the Spanish Ministry of Health (grant field-initiated studies program references PI041980, PI041771, and PI042450), the Spanish Network of Primary Care Research (redAPP) (ISCIII-RETIC RD06/0018), and SAMSERAP group. The UK National Health Service Research and Development office provided service support costs in the United Kingdom.

Disclaimer: The funders had no direct role in the design or conduct of the study, interpretation of the data, or review of the manuscript.

Additional Contributions: The European Office at University College London provided administrative assistance at the coordinating centre and Kevin McCarthy, project scientific officer, European Commission, Brussels, Belgium, provided helpful support and guidance. We thank all patients and general practice staff who took part; the UK Medical Research Council General Practice Research Framework (MRC GPRF); Louise Letley, MSc, from the MRC GPRF; the general practitioners of the Utrecht General Practitioners’ Network; and the Camden and Islington Mental Health and Social Care Trust.

REFERENCES


Correction

In the Original Article by King et al titled “Development and Validation of an International Risk Prediction Algorithm for Episodes of Major Depression in General Practice Attendees: The PredictD Study,” published in the December issue of the *Archives* (2008;65[12]:1368-1376), an incorrect URL was given in the “Results” and “Comment” sections for the predictD algorithm. The algorithm can be found at http://www.ucl.ac.uk/predict-depression.