ORIGINAL ARTICLE

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 Ángel Bellón-Saameño, MD, PhD; Berta Moreno, PhD; Igor Švab, MD, PhD; Danica Rotar, MD, MSc; J. Rifel, MD; Heidi-Ingrid Maaroos, MD, PhD; Anu Aluoja, PhD; Ruth Kalda, MD, DrMedSci; Jan Neeleman, MD, PhD; Mirjam I. Geerlings, PhD; Miguel Xavier, MD, PhD; Idalmiro Carraça, MD, MSc; Manuel Gonçalves-Pereira, MD, MSc; Benjamin Vicente, MD, PhD; Sandra Saldivia, 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 5216 European attendees who were not depressed at recruitment and had follow-up data on depression status. It was tested in 1732 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

Author Affiliations are listed at the end of this article.

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,[1] relapse 10 years from first presentation is frequent,[2] and both residual disability and premature mortality are common.[3] Low socioeconomic status[4,5] and female sex[6] are the 2 most consistently identified risk factors. Socioeconomic risk factors include low income and financial strain,[4] unemployment,[4] work stress,[7] social isolation,[8] and poor housing.[5] Other factors, such as family history of depression, play a part.[9] 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.[10] Poor social support, loneliness, and physical disability appear to be particular risks for older adults.[11-13] 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,[14] 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,[15] 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).[16,17] We selected risk factors to cover all important areas identified in a systematic review of the literature.[18] 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.[15] Each instrument or question not available in the relevant languages was translated from English and back-translated by professional translators.[15] 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, and (12) long-standing physical illness.
- (13) Lifetime depression was based on affirmative answers to both of the first 2 questions of the CIDI depression section.[19]
- *Stress in paid and unpaid work in the preceding 6 months using questions from the job content instrument.*[20] *Participants were categorized as feeling in control in (14) paid or (15) unpaid work; (16) experiencing difficulties without support in paid or unpaid work; and (17) experiencing distress without feeling respect for their paid or unpaid work.*
- (18) Financial strain using a question used in UK government social surveys.[4]
- Self-rated (19) physical and (20) mental health were assessed by the Short Form 12.[21] The weights used to calculate scores are from version 1.
- (21) Alcohol use in the preceding 6 months using the Alcohol Use Disorders Identification Test.[22] (22) We asked whether participants had ever had an alcohol problem or treatment for same.
- *(23) Whether participants had ever used recreational drugs using adapted sections of the CIDI.*
- *Questions on the quality of (24) sexual and (25) emotional relationships with partners or spouses.*[23]
- (26) Presence of serious physical, psychological, or substance misuse problems, or any serious disability, in people who were in close relationship to participants.
- *(27) Difficulties in getting on with people and maintaining close relationships.*[24]
- Childhood experiences of (28) physical and/or emotional and (29) sexual abuse.[25]
- (30) Holding religious and/or spiritual beliefs.[26]
- *(31) History of serious psychological problems or (32) suicide in first-degree relatives.*[27]
- (33) Anxiety and (34) panic symptoms in the previous 6 months using relevant sections of the Patient Health Questionnaire (PHQ).[28]
- *(35) Satisfaction with the neighborhood and (36) perceived safety inside/outside of the home using questions from the Health Surveys for England.*[29]
- (37) Major life events in the preceding 6 months using the List of Threatening Life Experiences Questionnaire.[30]
- *(38) Experiences of discrimination in the preceding 6 months on grounds of sex, age, ethnicity, appearance, disability, or sexual orientation using questions from a European study.*[31]
- (39) Adequacy of social support from family and friends.[32]

All participants were reevaluated for *DSM-IV* major depression, our main outcome, after 6 and 12 months using the depression section of the CIDI.

STATISTICAL ANALYSIS

All analyses and data imputation were performed using Stata release 9.[33] 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.[34] We imputed 10 data sets[35] and obtained combined estimates.[36]

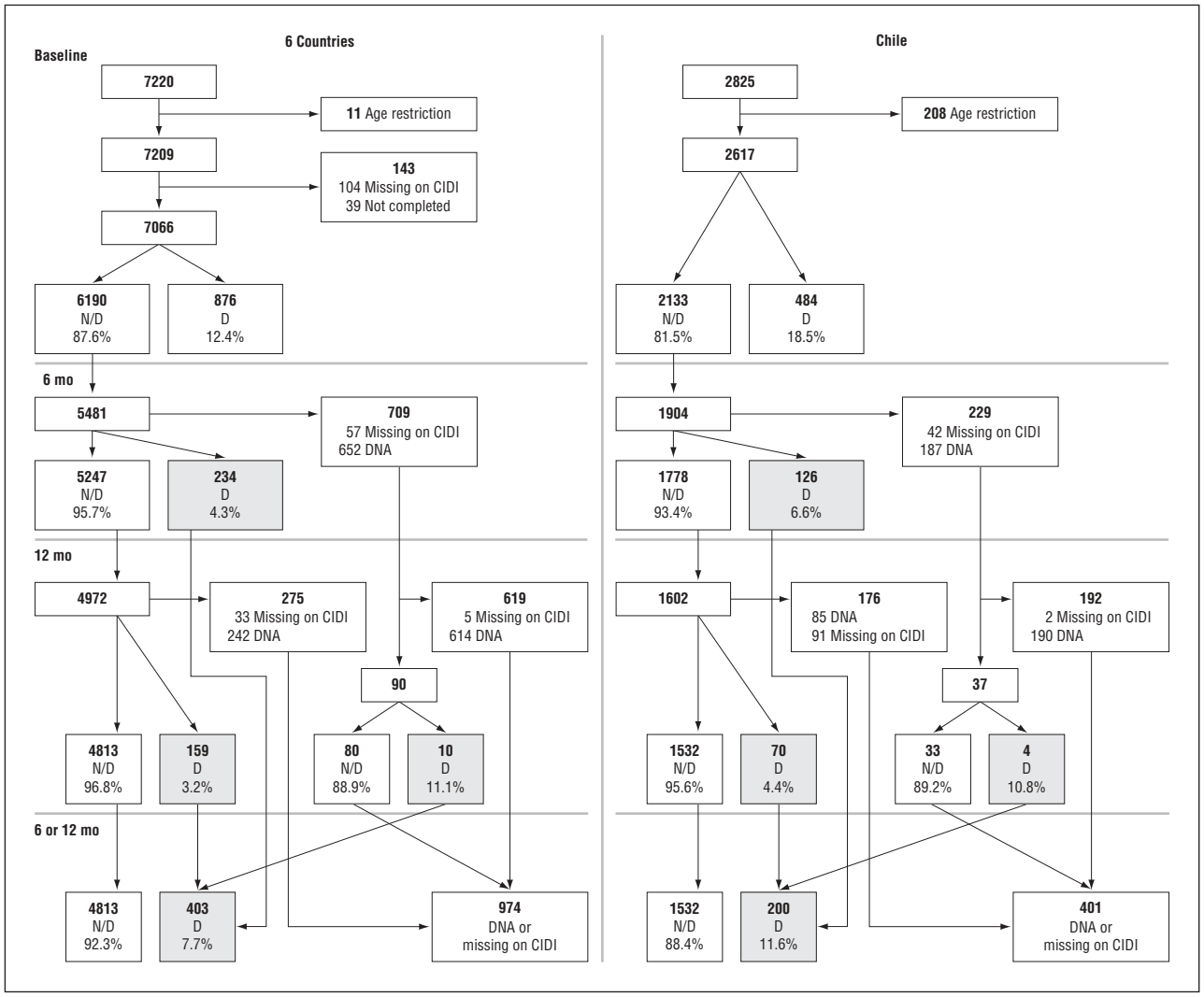

Figure 1. Flow of patients through the study and numbers becoming depressed. CIDI indicates Composite International Diagnostic Interview; DNA, did not attend; N/D, not depressed; and D, depressed.

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 $P < .01$ 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.

RESULTS

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

Table 1. Demographic Characteristics and Response to Follow-up

Characteristic	No. (%)							
	All European Countries	United Kingdom	Spain	Slovenia	Portugal	The Netherlands	Estonia	Chile
European	6190 (100)	1131 (18.3)	1006 (16.3)	1048 (16.9)	1005 (16.2)	1077 (17.4)	923 (14.9)	2133
Age, y, mean (SD)	48.9 (15.5)	52.2 (14.7)	50.8 (15.5)	48.8 (14.5)	50.2 (15.4)	48.9 (14.9)	41.6 (16.0)	47 (15.7)
Female	4081 (65.9)	750 (66.3)	689 (68.5)	660 (63)	649 (64.6)	668 (62)	665 (72)	1522 (71.4)
Marital status								
Married or living together	4491 (72.6)	844 (74.6)	708 (70.4)	732 (69.9)	750 (74.6)	827 (76.8)	630 (68.3)	1228 (57.6)
Separated or divorced	421 (6.8)	100 (8.8)	49 (4.9)	56 (5.3)	69 (6.9)	64 (5.9)	83 (9)	179 (8.4)
Single	872 (14.1)	121 (10.7)	181 (18)	152 (14.5)	132 (13.1)	121 (11.2)	165 (17.9)	521 (24.4)
Widowed	383 (6.2)	65 (5.8)	67 (6.7)	105 (10)	53 (5.3)	48 (4.5)	45 (4.9)	205 (9.6)
Missing	23 (0.4)	1 (0.1)	1 (0.1)	3 (0.3)	1 (0.1)	17 (1.6)	0	0
Household status								
Not living alone	5483 (88.6)	981 (86.7)	948 (94.2)	915 (87.3)	929 (92.4)	894 (83)	816 (88.4)	2039 (95.6)
Living alone	707 (11.4)	150 (13.3)	58 (5.8)	133 (12.7)	76 (7.6)	183 (17)	107 (11.6)	94 (4.4)
Missing	0	0	0	0	0	0	0	0
Education								
Higher education	1879 (30.4)	448 (39.6)	135 (13.4)	181 (17.3)	129 (12.8)	458 (42.5)	528 (57.2)	75 (3.5)
Secondary	2038 (32.9)	465 (41.1)	215 (21.4)	385 (36.7)	182 (18.1)	508 (47.2)	283 (30.7)	791 (37.1)
Primary/no education	1767 (28.6)	25 (2.2)	656 (65.2)	235 (22.4)	662 (65.9)	78 (7.2)	111 (12)	998 (46.8)
Trade/other	451 (7.3)	171 (15.1)	0	247 (23.6)	32 (3.2)	0	1 (0.1)	267 (12.5)
Missing	55 (0.9)	22 (1.9)	0	0	0	33 (3.1)	0	2 (0.1)
Employment								
Employed/full-time student	3256 (52.6)	574 (50.8)	349 (34.7)	563 (53.7)	486 (48.4)	602 (55.9)	682 (73.9)	749 (35.1)
Unemployed	300 (4.8)	21 (1.9)	62 (6.2)	53 (5.1)	108 (10.7)	35 (3.2)	21 (2.3)	243 (11.4)
Unable to work	322 (5.2)	86 (7.6)	101 (10)	16 (1.5)	38 (3.8)	48 (4.5)	33 (3.6)	69 (3.2)
Retired/looking after family	2269 (36.7)	450 (39.8)	493 (49)	409 (39)	372 (37)	358 (33.2)	187 (20.3)	1072 (50.3)
Missing	43 (0.7)	0	1 (0.1)	7 (0.7)	1 (0.1)	34 (3.2)	0	0
Professional status								
Yes	1313 (21.2)	309 (27.3)	107 (10.6)	165 (15.7)	96 (9.6)	373 (34.6)	263 (28.5)	29 (1.4)
Missing	143 (2.3)	27 (2.4)	3 (0.3)	3 (0.3)	0	56 (5.2)	54 (5.8)	3 (0.1)
Born in country of residence								
Yes	5655 (91.4)	1054 (93.2)	955 (94.9)	834 (79.6)	973 (96.8)	997 (92.6)	842 (91.2)	2122 (99.5)
Missing	87 (1.4)	3 (0.3)	3 (0.3)	4 (0.4)	0	24 (2.2)	53 (5.7)	4 (0.2)
Ethnicity								
White European	5988 (96.7)	1055 (93.3)	994 (98.8)	1042 (99.4)	992 (98.7)	983 (91.3)	922 (99.9)	0
Missing	72 (1.2)	39 (3.4)	1 (0.1)	2 (0.2)	0	30 (2.8)	0	0
6-mo Response	5538 (89.5)	987 (87.3)	794 (78.9)	963 (91.9)	889 (88.5)	1035 (96.1)	870 (94.3)	1904 (89.3)
12-mo Response	5319 (85.9)	965 (85.3)	731 (72.7)	927 (88.5)	864 (86)	988 (91.7)	844 (91.4)	1748 (82)

tonia (80%), Slovenia (80%), and Chile (97%) but lower in the United Kingdom (44%) and the Netherlands (45%). Ethical considerations prevented the collection of data on nonresponders at baseline. Across all countries, the response to follow-up was at 89.5% at 6 months and 85.9% at 12 months. Women predominated in each country and prevalence of major depression at baseline was 13.9% in women and 8.5% in men. Seven thousand two hundred nine European participants had full CIDI data at recruitment to allow a depression diagnosis, of whom 6190 were not depressed at recruitment (**Table 1**). Of these, 5216 (84.3%) had full CIDI data for a depression diagnosis at 6 and 12 months' follow-up, and of these, 3972 (76.2%) also had full data on all 39 risk factors. Cumulative 12 months' incidence of *DSM-IV* major depression in the European population was 7.7% (United Kingdom, 8.8%; Spain, 15.1%; Slovenia, 4.2%; Portugal, 8.5%; the Netherlands, 5.4%; and Estonia, 5.9%). Missing information was less than 3% for 38 of the 39 risk factors; however, 12.6% of participants had missing data on their emotional relationship with a spouse or partner (risk factor 25 in the "Major Depression and Known Risk Factors" subsection).

DEVELOPMENT OF THE RISK PREDICTION ALGORITHM IN THE EUROPEAN DATA

In our reliability study prior to recruiting the cohort, all risk factors tested (except discrimination on skin color) produced κ coefficients of 0.59 to 1.00 and percentage of agreement of 67% to 100%. The κ coefficient for agreement on discrimination due to skin color was low because of the small number of nonwhite participants.[15]

The risk algorithm was developed on the 5216 European attendees who were not depressed at recruitment and who had data on our main outcome, *DSM-IV* major depression at 6 and 12 months. Nonlinear transformations of continuous variables did not significantly improve the model fit. Seven variables were retained at $P < .01$ and these were included with country, age, and sex in the regression model (**Table 2**). Five variables in the final model concerned past events or patient char-

Table 2. PredictD Model Derived in the Imputed Data Sets

Prognostic Factor	Levels in Factor	Coefficient	SE	Coefficient After Copas Shrinkage	*P* Value
Constant		1.543	0.439	1.155	<.001
Age	Each year	–0.005	0.005	–0.005	.25
Sex	F				
	M	–0.245	0.138	–0.212	.07
Education	Beyond secondary education				
	Secondary education	0.103	0.128	0.089	.42
	Primary or no education	0.472	0.157	0.409	.003
	Trade/other	0.653	0.210	0.566	.002
Difficulties in paid and unpaid work	No difficulties or often supported				
	Difficulties without support	0.423	0.114	0.366	<.001
Physical health	Each point on SF-12 subscale score; possible range, 0-100	–0.034	0.005	–0.030	<.001
Mental health	Each point on SF-12 subscale score; possible range, 0-100	–0.064	0.005	–0.055	<.001
First-degree relative with emotional problem	No				
	Yes	0.456	0.090	0.395	<.001
Discrimination	None				
	In 1 area	0.186	0.220	0.161	.40
	In >1 area	0.850	0.235	0.736	<.001
Lifetime depression	No				
	Yes	0.565	0.131	0.489	<.001
Country	United Kingdom				
	Spain	0.266	0.205	0.230	.20
	Slovenia	–0.841	0.193	–0.729	<.001
	Estonia	–0.540	0.196	–0.467	.006
	The Netherlands	–0.133	0.220	–0.115	.54
	Portugal	–0.195	0.180	–0.169	.28

Abbreviation: SF-12, Short Form 12.

Table 3. C Index Statistics for Each Country[a]

Country	C Index[b] (95% Confidence Interval)
United Kingdom	0.756 (0.705-0.808)
Spain	0.793 (0.746-0.840)
Slovenia	0.833 (0.775-0.891)
Estonia	0.761 (0.690-0.833)
The Netherlands	0.852 (0.799-0.905)
Portugal	0.747 (0.693-0.800)
Mean over all countries	0.790 (0.767-0.813)

[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.

acteristics (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; $\chi^2_{16} = 19.06$). There was also no evidence for including interactions with age (*P* value = .21; $\chi^2_{16} = 20.19$).

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 of depression on the predictD score. Within each decile, we plotted mean predicted probability vs observed probability of depression (**Figure 2**A). 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

Table 4. Effect Sizes Computed Using Hedges g[a]

Country	Hedges g (95% Confidence Interval)
All European	1.28 (1.17-1.40)
United Kingdom	1.02 (0.78-1.27)
Spain	1.19 (0.97-1.41)
Slovenia	1.40 (1.06-1.75)
Estonia	1.09 (0.76-1.42)
The Netherlands	1.55 (1.25-1.85)
Portugal	0.99 (0.73-1.25)
Chile	0.87 (0.69-1.04)

[a] 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.

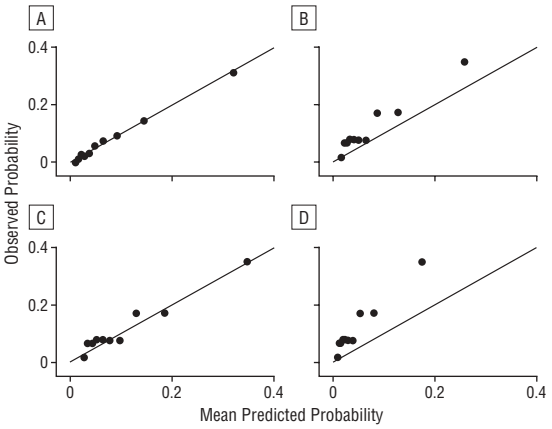

Figure 2. Plots of mean predicted probability against observed probability of depression within deciles of predicted risk. A, Model fitted in the European data using unshrunk coefficients. B-D, European model fitted in the Chilean data using shrunk coefficients. Risk scores are based on the average of 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.

Risk score (predicted probability of depression) at baseline 5%

A man aged 47 years living in Spain
Primary education
No difficulties or supported in paid and unpaid work
SF-12 mental subscale score 62
SF-12 physical subscale score 44
No personal history of depression
Family history of psychological difficulties
No experiences of discrimination

Risk score 10%

A man aged 55 years living in the Netherlands
Education beyond high school
No difficulties or supported in paid or unpaid work
SF-12 mental subscale score 30
SF-12 physical subscale score 56
No personal history of depression
Family history of psychological difficulties
Experience of discrimination in one area

Risk score 15% (8%)

A man aged 60 years living in the United Kingdom
Education beyond high school
Difficulties and unsupported in paid or unpaid work
SF-12 mental subscale score 56.4
SF-12 physical subscale score 29.7
Personal history of depression
No family history of psychological difficulties
Experience of discrimination in more than one area

Risk score 24% (11%)

A woman aged 62 years living in the United Kingdom
Education to high school
Difficulties and unsupported in paid or unpaid work
SF-12 mental subscale score 45.6
SF-12 physical subscale score 52.8
Personal history of depression
Family history of psychological difficulties
Experience of discrimination in more than one area

Risk score 33% (7%)

A man aged 55 years living in Spain
Education beyond high school
Difficulties and unsupported in paid or unpaid work
SF-12 mental subscale score 49.3
SF-12 physical subscale score 16.2
Personal history of depression
No family history of psychological difficulties
Experience of discrimination in more than one area

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).

probability of depression on the predictD score algorithm are shown in **Figure 3**. To demonstrate the potential impact of mutable factors on risk, scores in the last 3 examples in Figure 3 were recalculated after mutable risk factors were reduced or eliminated. Estimates of sensitivity and specificity of the risk score in predicting major depression over 12 months are shown in **Table 5**. Questions in the risk algorithm can be tested at http://www.techflora.com/ucl and a risk score obtained. (The questions, responses, coding, and algorithm for the predictD risk tool are available on request.)

EXTERNAL VALIDATION OF THE RISK PREDICTION ALGORITHM IN THE CHILEAN DATA SET

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 shrunk 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 5. Thresholds for Specificity and Sensitivity in Each Setting

	Predicted Probability of Major Depression, PredictD Risk Score	Specificity	Sensitivity	Likelihood, Ratio
Europe	0.106	0.800	0.645	3.22
	0.130	0.850	0.556	3.71
	0.165	0.900	0.464	4.64
Country				
United Kingdom	0.133	0.800	0.506	2.53
United Kingdom	0.154	0.850	0.458	3.06
United Kingdom	0.183	0.900	0.373	3.72
Spain	0.193	0.800	0.667	3.34
Spain	0.231	0.850	0.565	3.76
Spain	0.292	0.899	0.407	4.05
Slovenia	0.063	0.800	0.632	3.17
Slovenia	0.076	0.849	0.579	3.85
Slovenia	0.097	0.900	0.553	5.51
Estonia	0.088	0.801	0.571	2.86
Estonia	0.100	0.850	0.510	3.41
Estonia	0.130	0.900	0.408	4.09
The Netherlands	0.080	0.800	0.769	3.85
The Netherlands	0.096	0.850	0.731	4.87
The Netherlands	0.111	0.900	0.654	6.51
Portugal	0.117	0.800	0.452	2.26
Portugal	0.147	0.850	0.397	2.65
Portugal	0.185	0.899	0.356	3.54
Chile	0.089	0.800	0.525	2.63
Chile	0.109	0.850	0.410	2.73
Chile	0.138	0.900	0.320	3.20

COMMENT

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 cohorts[44] 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.[14] 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 Bellón-Saameño), and Department of Psychiatry, University of Granada, Granada (Drs Moreno and Torres-Gonzalez), Spain; Department of Family Medicine, University of Ljubljana, Ljubljana, Slovenia (Drs Švab, Rotar, and Rifel); Faculty of Medicine, University of Tartu, Tartu, Estonia (Drs Maaroos, Aluoja, and Kalda); University Medical Center, Utrecht, the Netherlands (Drs Neeleman 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 Saldivia 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 (redIAPP) (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

1. Goldberg DP, Huxley P. *Common Mental Disorders: A Bio-Social Model.* London, England: Tavistock/Routledge; 1992.
2. Thornicroft G, Sartorius N. The course and outcome of depression in different cultures: 10-year follow-up of the WHO Collaborative Study on the Assessment of Depressive Disorders. *Psychol Med.* 1993;23(4):1023-1032.

3. Cassano P, Fava M. Depression and public health: an overview. *J Psychosom Res*. 2002;53(4):849-857.
4. Weich S, Lewis G. Poverty, unemployment, and common mental disorders: population based cohort study. *BMJ*. 1998;317(7151):115-119.
5. Weich S, Lewis G. Material standard of living, social class, and the prevalence of the common mental disorders in Great Britain. *J Epidemiol Community Health*. 1998;52(1):8-14.
6. Weich S, Sloggett A, Lewis G. Social roles and gender difference in the prevalence of common mental disorders. *Br J Psychiatry*. 1998;173:489-493.
7. Stansfeld SA, Fuhrer R, Shipley MJ, Marmot MG. Work characteristics predict psychiatric disorder: prospective results from the Whitehall II Study. *Occup Environ Med*. 1999;56(5):302-307.
8. Bruce ML, Hoff RA. Social and physical health risk factors for first-onset major depressive disorder in a community sample. *Soc Psychiatry Psychiatr Epidemiol*. 1994;29(4):165-171.
9. Angst J, Gamma A, Endrass J. Risk factors for the bipolar and depression spectra. *Acta Psychiatr Scand Suppl*. 2003;(418):15-19.
10. Salokangas RKR, Poutanen O. Risk factors for depression in primary care: findings of the TADEP project. *J Affect Disord*. 1998;48(2-3):171-180.
11. Prince MJ, Harwood RH, Blizard RA, Thomas A, Mann AH. Impairment, disability and handicap as risk factors for depression in old age: the Gospel Oak Project V. *Psychol Med*. 1997;27(2):311-321.
12. Prince MJ, Harwood RH, Blizard RA, Thomas A, Mann AH. Social support deficits, loneliness and life events as risk factors for depression in old age: the Gospel Oak Project VI. *Psychol Med*. 1997;27(2):323-332.
13. Prince MJ, Harwood RH, Thomas A, Mann AH. A prospective population-based cohort study of the effects of disablement and social milieu on the onset and maintenance of late-life depression: the Gospel Oak Project VII. *Psychol Med*. 1998;28(2):337-350.
14. Anderson KM, Wilson PW, Odell PM, Kannel WB. An updated coronary risk profile: a statement for health professionals. *Circulation*. 1991;83(1):356-362.
15. King M, Weich S, Torres-González F, Svab I, Maaroos HI, Neeleman J, Xavier M, Morris R, Walker C, Bellón-Saameño JA, Moreno-Küstner B, Rotar D, Rifel J, Aluoja A, Kalda R, Geerlings MI, Carraça I, de Almeida MC, Vicente B, Saldivia S, Rioseco P, Nazareth I. Prediction of depression in European general practice attendees: the PREDICT study. *BMC Public Health*. 2006;6(1):6.
16. Robins LN, Wing J, Wittchen HU, Helzer JE, Babor TF, Burke J, Farmer A, Jablenski A, Pickens R, Regier DA. The Composite International Diagnostic Interview: an epidemiologic instrument suitable for use in conjunction with different diagnostic systems and in different cultures. *Arch Gen Psychiatry*. 1988;45(12):1069-1077.
17. World Health Organization. *Composite International Diagnostic Interview (CIDI). Version 2.1*. Geneva, Switzerland: WHO; 1997.
18. Weich S. *Risk Factors for the Common Mental Disorders in Primary Care*. Cambridge, England: University of Cambridge; 2001.
19. Arroll B, Khin N, Kerse N. Screening for depression in primary care with two verbally asked questions: cross sectional study. *BMJ*. 2003;327(7424):1144-1146.
20. Karasek RA, Theorell T. *Healthy Work: Stress, Productivity, and the Reconstruction of Working Life*. New York, NY: Basic Books; 1990.
21. Jenkinson C, Layte R, Jenkinson D, Lawrence K, Petersen S, Paice C, Stradling J. A shorter form health survey: can the SF-12 replicate results from the SF-36 in longitudinal studies? *J Public Health Med*. 1997;19(2):179-186.
22. Barbor TF, de la Fuente JR, Saunders J, Grant M. *The Alcohol Use Disorders Identification Test: Guidelines for the Use in Primary Health Care*. Geneva, Switzerland: World Health Organization; 1989.
23. Taylor JF, Rosen RC, Leiblum SR. Self-report assessment of female sexual function: psychometric evaluation of the Brief Index of Sexual Functioning for Women. *Arch Sex Behav*. 1994;23(6):627-643.
24. Tyrer P. Personality disorder and social functioning. In: Peck DF, Shapiro CM, eds. *Measuring Human Problems: a Practical Guide*. Chichester, NY: Wiley & Sons; 1990:119-142.
25. 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(9):1329-1335.
26. King M, Speck P, Thomas A. The Royal Free interview for religious and spiritual beliefs: development and standardization. *Psychol Med*. 1995;25(6):1125-1134.
27. Qureshi N, Bethea J, Modell B, Brennan P, Papageorgiou A, Raeburn S, Hapgood R, Modell M. Collecting genetic information in primary care: evaluating a new family history tool [published online ahead of print July 29, 2005]. *Fam Pract*. 2005;22(6):663-669.
28. 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(18):1737-1744.
29. Sproston K, Primatesta P. *Health Survey for England 2002: a Survey Carried out on Behalf of the Department of Health. Volume 1: The Health of Children and Young People*. London, England: The Stationery Office; 2003.
30. 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(1):189-194.
31. Janssen I, Hanssen M, Bak M, Bijl RV, de Graaf R, Vollebergh W, McKenzie K, van Os J. Discrimination and delusional ideation. *Br J Psychiatry*. 2003;182:71-76.
32. Blaxter M. *Health and Lifestyles*. London, England: Routledge; 1990.
33. *Stata* [computer program]. Release 9. College Station, TX: StataCorp; 2007.
34. Royston P. Multiple imputation of missing values: update of ice. *Stata Journal*. 2005;5(4):527-536.
35. Schafer JL. Multiple imputation: a primer. *Stat Methods Med Res*. 1999;8(1):3-15.
36. Rubin DB. *Multiple Imputation for Non-Response in Surveys*. New York, NY: John Wiley & Sons; 1987.
37. Piccinelli M, Wilkinson G. Gender differences in depression: critical review. *Br J Psychiatry*. 2000;177(6):486-492.
38. Weissman MM, Bland RC, Canino GJ, Faravelli C, Greenwald S, Hwu HG, Joyce PR, Karam EG, Lee CK, Lellouch J, Lépine JP, Newman SC, Rubio-Stipec M, Wells JE, Wickramaratne PJ, Wittchen H, Yeh EK. Cross-national epidemiology of major depression and bipolar disorder. *JAMA*. 1996;276(4):293-299.
39. Harrell FE. *Regression Modelling Strategies*. New York, NY: Springer; 2001.
40. Copas JB. Regression, prediction and shrinkage. *J R Stat Soc Ser B*. 1983;45:311-354.
41. Cooper H, Hedges LV. *The Handbook of Research Synthesis*. New York, NY: Russell Sage Foundation; 1994.
42. Rubenstein LV, Rayburn NR, Keeler EB, Ford DE, Rost KM, Sherbourne CD. Predicting outcomes of primary care patients with major depression: development of a depression prognosis index. *Psychiatr Serv*. 2007;58(8):1049-1056.
43. Pepe MS, Janes H, Longton G, Leisenring W, Newcomb P. Limitations of the odds ratio in gauging the performance of a diagnostic, prognostic, or screening marker. *Am J Epidemiol*. 2004;159(9):882-890.
44. Conroy RM, Pyörälä K, Fitzgerald AP, Sans S, Menotti A, De Backer G, De Bacquer D, Ducimetière P, Jousilahti P, Keil U, Njølstad I, Oganov RG, Thomsen T, Tunstall-Pedoe H, Tverdal A, Wedel H, Whincup P, Wilhelmsen L, Graham IM; SCORE project group. Estimation of ten-year risk of fatal cardiovascular disease in Europe: the SCORE project. *Eur Heart J*. 2003;24(11):987-1003.
45. Altman DG, Royston P. What do we mean by validating a prognostic model? *Stat Med*. 2000;19(4):453-473.
46. Moons KG, Donders AR, Steyerberg EW, Harrell FE. Penalized maximum likelihood estimation to directly adjust diagnostic and prognostic prediction models for overoptimism: a clinical example. *J Clin Epidemiol*. 2004;57(12):1262-1270.

48. Ongür D, Drevets WC, Price JL. Glial reduction in the subgenual prefrontal cortex in mood disorders. *Proc Natl Acad Sci U S A*. 1998;95(22):13290-13295.
49. Rajkowska G. Cell pathology in bipolar disorder. *Bipolar Disord*. 2002;4(2):105-116.
50. Tkachev D, Mimmack ML, Ryan MM, Wayland M, Freeman T, Jones PB, Starkey M, Webster MJ, Yolken RH, Bahn S. Oligodendrocyte dysfunction in schizophrenia and bipolar disorder. *Lancet*. 2003;362(9386):798-805.
51. Sequeira A, Turecki G. Genome wide gene expression studies in mood disorders. *OMICS*. 2006;10(4):444-454.
52. Sokolov BP. Oligodendroglial abnormalities in schizophrenia, mood disorders and substance abuse: comorbidity, shared traits or molecular phenocopies? *Int J Neuropsychopharmacol*. 2007;10(4):547-555.
53. Davis KL, Stewart DG, Friedman JI, Buchsbaum M, Harvey PD, Hof PR, Buxbaum J, Haroutunian V. White matter changes in schizophrenia: evidence for myelin-related dysfunction. *Arch Gen Psychiatry*. 2003;60(5):443-456.
54. McDonald C, Bullmore ET, Sham PC, Chitnis X, Wickham H, Bramon E, Murray RM. Association of genetic risks for schizophrenia and bipolar disorder with specific and generic brain structural endophenotypes. *Arch Gen Psychiatry*. 2004; 61(10):974-984.
55. Carter CJ. EIF2B and oligodendrocyte survival: where nature and nurture meet in bipolar disorder and schizophrenia? [published online ahead of print February 27, 2007]. *Schizophr Bull*. 2007;33(6):1343-1353.
56. Hakak Y, Walker JR, Li C, Wong WH, Davis KL, Buxbaum JD, Haroutunian V, Fienberg AA. Genome-wide expression analysis reveals dysregulation of myelination-related genes in chronic schizophrenia. *Proc Natl Acad Sci U S A*. 2001; 98(8):4746-4751.
57. Vostrikov VM, Uranova NA, Orlovskaya DD. Deficit of perineuronal oligodendrocytes in the prefrontal cortex in schizophrenia and mood disorders. *Schizophr Res*. 2007;94(1-3):273-280.
58. Segal D, Koschnick JR, Slegers LHA, Hof PR. Oligodendrocyte pathophysiology: a new view of schizophrenia. *Int J Neuropsychopharmacol*. 2007;10(4): 503-511.
59. Haznedar MM, Roversi F, Pallanti S, Baldini-Rossi N, Schnur DB, Licalzi EM, Tang C, Hof PR, Hollander E, Buchsbaum MS. Fronto-thalamo-striatal gray and white matter volumes and anisotropy of their connections in bipolar spectrum illnesses. *Biol Psychiatry*. 2005;57(7):733-742.
60. Chen G, Zeng WZ, Yuan PX, Huang LD, Jiang YM, Zhao ZH, Manji HK. The mood-stabilizing agents lithium and valproate robustl increase the levels of the neuroprotective protein bcl-2 in the CNS. *J Neurochem*. 1999;72(2):879-882.
61. Moore GJ, Bebchuk JM, Hasanat K, Chen G, Seraji-Bozorgzad N, Wilds IB, Faulk MW, Koch S, Glitz DA, Jolkovsky L, Manji HK. Lithium increases N-acetyl-aspartate in the human brain: in vivo evidence in support of bcl-2's neurotrophic effects? *Biol Psychiatry*. 2000;48(1):1-8.
62. Koo MS, Levitt JJ, Salisbury DF, Nakamura M, Shenton ME, McCarley RW. A cross-sectional and longitudinal magnetic resonance imaging study of cingulate gyrus gray matter volume abnormalities in first-episode schizophrenia and first-episode affective psychosis. *Arch Gen Psychiatry*. 2008;65(7):746-760.
63. Nakamura M, Salisbury DF, Hirayasu Y, Bouix S, Pohl KM, Yoshida T, Koo MS, Shenton ME, McCarley RW. Neocortical gray matter volume in first-episode schizophrenia and first-episode affective psychosis: a cross-sectional and longitudinal MRI study. *Biol Psychiatry*. 2007;62(7):773-783.
64. Moorhead TWJ, McKirdy J, Sussman JED, Hall J, Lawrie SM, Johnstone EC, McIntosh AM. Progressive gray matter loss in patients with bipolar disorder. *Biol Psychiatry*. 2007;62(8):894-900.

Correction

Error in Text. 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.