Research Article

DEVELOPMENT AND VALIDATION OF A PREDICTION ALGORITHM FOR USE BY HEALTH PROFESSIONALS IN PREDICTION OF RECURRENCE OF MAJOR DEPRESSION

Jian Li Wang, Ph.D.,[1,2]* Scott Patten, M.D., Ph.D.,[1,2] Jitender Sareen, M.D.,[3] James Bolton, M.D.,[3] Norbert Schmitz, Ph.D.,[4] and Glenda MacQueen, M.D.[1]

Background: *There exists very little evidence to guide clinical management for preventing recurrence of major depression. The objective of this study was to develop and validate a prediction algorithm for recurrence of major depression.* **Methods:** *Wave 1 and wave 2 longitudinal data from the U.S. National Epidemiological Survey on Alcohol and Related Condition (2001/2002–2003/2004) were used. Participants with a major depressive episode at baseline and who had visited health professionals for depression were included in this analysis (n = 2,711). Mental disorders were assessed based on the DSM-IV criteria.* **Results:** *With the development data (n = 1,518), a prediction model with 19 unique factors had a C statistics of 0.7504 and excellent calibration (P = .23). The model had a C statistics of 0.7195 in external validation data (n = 1,195) and 0.7365 in combined data. The algorithm calibrated very well in validation data. In the combined data, the 3-year observed and predicted risk of recurrence was 25.40% (95% CI: 23.76%, 27.04%) and 25.34% (95% CI: 24.73%, 25.95%), respectively. The predicted risk in the 1st and 10th decile risk group was 5.68% and 60.21%, respectively.* **Conclusions:** *The developed prediction model for recurrence of major depression has acceptable discrimination and excellent calibration, and is feasible to be used by physicians. The prognostic model may assist physicians and patients in quantifying the probability of recurrence so that physicians can develop specific treatment plans for those who are at high risk of recurrence, leading to personalized treatment and better use of resources. Depression and Anxiety 31:451–457, 2014.*

Key words: *prediction algorithm; recurrence; major depression; development; validation*

INTRODUCTION

Major depression has a high rate of recurrence after recovery.[1–3] For physicians who treat patients with major depression, once the depressive episode is remitted, the goal of the next treatment stage is to prevent recurrence of a new episode.[4] This treatment stage is very challenging and there exists very little evidence to guide clinical management. In the NICE guidelines for depression, number of previous episodes of depression, residual

[1]Department of Psychiatry, Faculty of Medicine, University of Calgary, Calgary, Canada
[2]Department of Community Health Sciences, Faculty of Medicine, University of Calgary, Calgary, Canada
[3]Department of Psychiatry, Faculty of Medicine, University of Manitoba, Winnipeg, Canada
[4]Department of Psychiatry, Faculty of Medicine, McGill University, Montreal, Canada

*Correspondence to: JianLi Wang, Department of Psychiatry, Faculty of Medicine, University of Calgary, Room 4D69, TRW Building, 3280 Hospital Drive NW, Calgary, AB Canada T2N 4Z6. E-mail: jlwang@ucalgary.ca

Received for publication 16 September 2013; Revised 16 October 2013; Accepted 18 October 2013

DOI 10.1002/da.22215
Published online 13 November 2013 in Wiley Online Library (wileyonlinelibrary.com).

symptoms and concurrent physical health problems, and psychosocial difficulties are considered as risk factors for recurrence of depression.[5] The Canadian CANMAT guidelines for depression specify three or more episodes as a risk factor for recurrence.[6] In community-based studies, younger age, younger age of onset, family history of depression, previous episodes, duration of the episode, residual symptoms, smoking, comorbid anxiety disorders, negative youth experiences, and ongoing difficulties have been found to be associated with recurrence.[6–13] Although previous research has significantly enhanced our understanding about potential factors for depression recurrence and the clinical practice guidelines have clearly laid out the factors to be considered at maintenance stage, these have not provided physicians with sufficient ability to predict and quantify the risk of recurrence. Recurrence of major depressive episode (MDE) is not determined by a single factor, rather by the combined effect of multiple risk factors. To accurately predict and quantify the risk of recurrence, a prediction model that considers the combined effect of a key set of prognostic factor is needed.[14]

A prediction or prognostic algorithm is not about searching for new risk factors; it is a combination of multiple known predictors from which risk of a specific endpoint can be calculated for individual patients.[14] Such models are embedded in clinicians' daily practice as the primary tool to estimate individual risk of future disease. Well-known examples include the Framingham risk prediction algorithms for cardiovascular disease[15] and prediction algorithms for cancer risk.[16,17]

Although many studies established the associations between factors described previously and recurrence of major depression, using statistical modeling approaches (either logistic regression or Cox model), the purpose of these models was to estimate relative risks (e.g., exposed vs. nonexposed), and they are different from prediction algorithms that provide absolute risk (probability of outcome at endpoint). They should have acceptable discriminating power and calibration not only in the population from which the models are developed, but also in distinct but related populations. Such models to predict recurrence of depressive episodes in specific individuals are not available, despite the fact that they could have clinical utility. The selection of maintenance therapy in remitted patients is guided in large part by the physicians' clinical judgment. Improved ability to identify recurrence risk for specific patients could, therefore, improve clinicians' decisions regarding the need for ongoing or maintenance therapy. The objective of this study was to develop and validate a prognostic model for predicting recurrence of major depression using data from a population-based, nationally representative cohort.

METHODS

We used the longitudinal data from the U.S. National Epidemiological Survey on Alcohol and Related Conditions (NESARC). The NESARC is a nationally representative survey of the U.S. population.[18] Wave 1 of the NESARC[19] was collected between 2001 and 2002, and included 43,093 respondents aged 18 years and older. Wave 2 of the NESARC was collected between 2004 and 2005, about 3 years after the wave 1, and completed interviews in 34,653 participants of the original wave 1 sample.[20] Individuals who were institutionalized, who joined in military, and who went to prisons after the wave 1 were not interviewed in wave 2. The response rate in eligible participants was 86.7%.[21] A detailed description of the design and field procedures of the NESARC can be found in previous publications.[19,20]

Assessment of mental disorders: Major depression and other Axis-I and Axis-II mental disorders were diagnosed using the Alcohol Use Disorder and Associated Disabilities Interview Schedule – DSM-IV version (AUDADIS-IV),[22] a fully structured diagnostic interview appropriate for use by trained lay interviewers and clinicians. Lifetime and past-year diagnoses were assessed at wave 1. At wave 2, diagnoses since wave 1 were assessed.

Participants included in this analysis: For this study, we included the wave 1 (baseline) participants who reported (1) current or lifetime MDE; (2) no lifetime manic episode or hypomanic episode; (3) having remitted from the recent depressive episode for at least 2 months; and (4) that they went to health professionals (councilors and/or medical doctors) for help to improve mood, were hospitalized for depression, or went to emergency room because of depression. In the NESARC, after participants were inquired about the presence of a MDE, they were asked: "Since this MOST RECENT time BEGAN, have there been at least 2 months when your mood was much improved or back to normal AND when you DIDN'T have ANY of the OTHER experiences you mentioned?" The NESARC used the answer (yes or no) to this question to determine whether remission had occurred. We included eligible participants from South and West region (regions 3 and 4) in the development data to develop the model ($n = 1{,}518$). Eligible participants from Northeast and Midwest region (regions 1 and 2) were kept in validation data ($n = 1{,}195$).

The outcome variable of this study was meeting the DSM-IV diagnostic criteria for MDE since wave 1 interview.

Risk factors for recurrence: We initially selected and examined the associations between the following potential risk factors and recurrence. We also examined the associations between other Axis-I and Axis-II disorders at baseline and the risk of recurrence.

Demographic and socioeconomic characteristics: Sex, age, marital status, annual personal income levels, educational levels, working status, race, living arrangement.

Clinical variables related to previous MDE: As part of the AUDADIS-IV module for major depression, the following were available: number of MDEs, age of first onset, duration of longest MDE, presenting depressive symptoms in an episode, presence of major depressive disorder (MDD) in the past year and prior to the past year.

Family history: As part of the AUDADIS-IV, participants were asked about whether their biological parents and siblings ever had depression (six items).[22]

Physical conditions: Self-reported general health and physician-diagnosed or healthcare professional diagnosed health conditions in the past year.[23]

Stressful life events in the past 12 months included the occurrence of 14 stressful life events during the 12 months prior to the interview.[22]

Past month health-related quality of life: Two derived variables were based on the Medical Outcomes study – Short Form (SF-12):[24] norm-based physical disability scores and norm-based mental health disability scores.

Childhood adversities: 29 questions from the Conflict Tactics Scale[25] and the Childhood Trauma Questionnaire.[26]

Discrimination: The Experiences with Discrimination scales[27] (54 items) were used to assess experience of discrimination in physical disability, race–ethnicity, gender, sexual orientation, religion, and overweight. These questions were asked at wave 2 and accommodated

two time periods: the past 12 months, and prior to the past 12 months. We assumed that people's experience of discrimination (especially related to race and religion) did not change significantly over a short time (2 years). Therefore, we used participants' answers about experience of discrimination prior to the past 12 months as an indicator for discrimination.

At the time of the wave 1 interview, missing data were imputed by the NESARC team for background variables deemed critical for analysis, for example, age, sex, race, education, income, and so forth.[28] Imputation was done by means of a hot-deck procedure defined by combinations of relevant characteristics.[28]

Model development and validation: All analyses were conducted using STATA 12.[29] To develop the model, we first examined the bivariate relations between selected variables and the risk of recurrence in the development data. Variables that were statistically significant at the *P*-value level of .005 were retained as candidates for modeling. Although some variables were not statistically significant at the level of .005, they have been found to be strong risk factors for depression. These variables were later examined regarding their added values for model performance.

The prediction algorithm was developed using logistic regression modeling. To develop the algorithm, we used combined procedures of forward and backward selection as each individual approach has its own limitations.[30,31] We first included age (continuous variable), sex, and number of previous depressive episodes in the model. We used two ways to determine whether an additional variable should be included in the model. First, by adding another variable from specific domains (demographic, psychosocial, and clinical), we examined whether it would improve the model's discriminative power and calibration with data. This was done by comparing the difference between the *C* statistics of the models with and without the variable. Once the variables from a specific domain had been examined, we used backward deletion to examine whether removing a variable would affect the performance of the model. The variable with the weakest association was first examined. After the backward deletion was completed, we repeated the forward selection by adding a variable from another domain. The combined forward and backward selection were repeated until all variables were examined. Second, we used the method of Net Reclassification Improvement (NRI)[32,33] to examine whether adding this particular variable could correctly reclassify participants into appropriate categories. When participants with and without the outcome are considered separately, any upward movement in disease categories for subjects with the outcome implies improved classification, and any downward movement indicates worse reclassification. The interpretation is opposite for subjects without the outcome. The NRI was quantified as the sum of differences in proportions of individuals moving up minus the proportion moving down for those with the outcome, and the proportion of individuals moving down minus the proportion moving up for those without the outcome.[34] The *P*-value (.05) of statistical test for comparing the difference was used to determine if the improvement is significant. We also examined the interactions between the included predictors. Interaction terms that significantly contributed to *C* statistics and NRI were retained in the prediction model.

Internal validation: The model's performance was assessed by discrimination and calibration. Discrimination is the ability of a prediction model to separate those who experienced the outcome events from those who did not. We quantified this by calculating the *C* statistic, which is identical to the area under a receiver operating characteristic curve when the outcome variable is binary. Calibration measures how closely predicted outcomes agree with actual outcomes (or accuracy). For this, we used the Hosmer–Lemeshow test (H–L test) to compare the differences between mean predicted and actual event rates; large *P*-value (>.05) indicates good calibration.

Shrinkage: The model developed using the development data may be vulnerable to overfitting. Therefore, we shrunk the coefficients with a heuristic shrinkage factor, which was estimated as follows:[35]

$$S_{\text{heur}} = [\text{model } \chi^2 - (df - 1)]/\text{model } \chi^2,$$

where *df* indicates the degrees of freedom of the model and model χ^2 is calculated on the log-likelihood scale.

External validation: We applied the shrunken model in the validation data ($n = 1{,}195$) and calculated *C* statistics (for discrimination) and the H–L χ^2 statistics (for calibration/accuracy). The H–L χ^2 is a test for overall calibration, but it provides no information about specific areas (risk groups) where the predicted risk agrees, over estimates, or under estimates the true risk. To examine the agreement between predicted and observed risk by specific risk groups, we combined the development and validation data and applied the model in the combined data. Participants were classified into 10 (decile) risk groups. The predicted versus observed risk of recurrence were visually compared. Additionally, we calculated the *C* statistics and the H–L statistics of the model in the combined data.

RESULTS

The characteristics of the participants in the development and validation datasets and in the combined dataset are in Table 1. The participants in the development and validation data resembled each other, except that participants in the validation data were more likely to be men, white, and to have visited an emergency room for depression than those in the development data. In the development data ($n = 1{,}518$), 382 (25.16%) developed recurrent major depression since wave 1. In the validation data ($n = 1{,}195$), 307 (25.69%) developed recurrent major depression over the follow-up period. The participants had complete data about selected variables, except that two participants had missing data about the SF-12 scores.

In bivariate analysis, many variables including family history, duration of depressive episode, suicidal behaviors, and ongoing life events were associated with recurrence of major depression. However, these variables did not add to the prediction of recurrence by examination of their impacts on *C* statistics and NRI. The final model based on the development data contains 19 unique predictors and 4 interaction terms (Table 2). The *C* statistics of this model was 0.7504. The model had excellent calibration in the development data (H–L $\chi^2(8) = 10.48$, $P = .23$).

The shrunken coefficients of the model are in Table 2. We applied the model in the validation data. The data showed that the model had a *C* statistics of 0.7195 and excellent calibration in the validation data (H–L $\chi^2(8) = 3.51$, $P = .90$). We applied the algorithm in the combined development and validation data. With the combined data, the model had a *C* statistics of 0.7365 and excellent calibration (H–L $\chi^2(8) = 6.22$, $P = .62$). The observed risk of recurrence over 3 years was 25.40% (95% CI: 23.76%, 27.04%); the mean predicted risk of recurrence based on the model was 25.34% (95% CI: 24.73%, 25.95%). We visually compared the predicted versus the observed risk of recurrence by decile risk groups (Fig. 1).

TABLE 1. Demographic, socioeconomic, and clinical characteristics of the participants in development and validation datasets

Variables	Devel. data $N = 1,518$ *n*, %	Validation data $N = 1,195$ *n*, %	Combined $N = 2,713$ *n*, %	Nonservice users $N = 1,704$ *n*, %
Men	343, 22.6%	300, 25.1%	643, 23.70%	604, 35.45%
Women	1,175, 77.4%	895, 74.9%	2,070, 76.30%	1,100, 64.55%
Age (mean, *SE*)	45.38, 0.37	45.37, 0.41	45.37, 0.28	43.70, 0.40
Married/CL	735, 48.42%	583, 48.79%	1,318, 48.58%	802, 47.07%
Never married	112, 7.38%	86, 7.2%	198, 7.30%	156, 9.15%
D/S/W	671, 44.20%	526, 44.01%	1,197, 44.12%	746, 43.78%
Household income				
$60,000+	465, 30.64%	352, 29.46%	817, 30.11%	417, 24.47%
$35,000–$59,999	399, 26.28%	331, 27.70%	730, 26.91%	438, 25.70%
$20,000–$34,999	302, 19.89%	246, 20.58%	548, 20.20%	397, 23.30%
<$19,999	352, 23.19%	266, 22.26%	618, 22.78%	452, 26.53%
White	1,020, 67.19%	937, 78.41%	1,957, 72.13%	1,080, 63.62%
Non-white	498, 32.81%	258, 21.59%	756, 27.87%	624, 36.62%
Being hospitalized due to depression	244, 16.07%	191, 15.98%	435, 16.03%	
Went to emergency Room for depression	172, 11.33%	170, 14.23%	342, 12.61%	

SE, standard error; CL, common-law relationship; D/S/W; divorced/separated/widowed; Devel., development.

In general, the predicted risk agreed with the observed risk very well. There was about 3% overprediction in group 7 and 4% underprediction in group 8. The predicted and observed risks were nearly identical in the highest risk group.

DISCUSSION

The purpose of this prognostic model is not to identify new risk factors for recurrent depression, nor is it to estimate the relative risks of the predictors. Instead, it provides a statistical formula based on which individuals' absolute risks of developing recurrent depressive episode in the future can be calculated. The developed prediction algorithm provides a statistical formula based on which individuals' absolute risks of developing recurrent depressive episode in the future can be calculated. This type of tool can be used collaboratively by both physicians and patients when deciding on ongoing treatment and monitoring after remission of their depressive episode. The model developed in this study based on the NESARC data had acceptable discrimination and excellent calibration. Unlike most prediction algorithms that did not have external validation when they were first developed, this prediction model was validated in independent samples.

Consistent with current clinical practice guidelines, number of previous episodes, residual symptoms, concurrent physical health problems, and psychosocial difficulties are import prognostic factors in this model. It should be noted that accurate prediction of recurrence can only be made when these factors are considered simultaneously, rather than by focusing on one or a few prognostic factors. Number of previous depressive episodes is emphasized in both NICE and CANMAT guidelines.[5,6] In the NESARC data, when only number of previous episodes was used to predict recurrence, the model had a *C* statistic of 0.5897, which was similar to a model that included gender and age only ($C = 0.5794$). This highlights that using number of previous depressive episodes alone would not make accurate prediction of recurrence risk.

In addition to what clinical practice guidelines have previously noted, this model showed that demographic characteristics (gender, age, marital status, race) and comorbid psychiatric disorders (lifetime generalized anxiety disorder, specific phobia, and avoidant personality disorder) contributed to the prediction of recurrence. Previous studies are inconsistent in whether these factors are associated with recurrence of major depression. In a clinical study, Grilo et al. reported that personality disorders, specifically borderline and obsessive-compulsive disorders, were strong predictors for recurrence of major depression.[36] In another clinical study, avoidant personality disorder was strongly associated with relapse of MDD.[37] In an analysis of the NESARC data, Skodol et al. found that none of the personality disorders was associated with recurrence of MDD.[21] The discrepancy between current analysis and Skodol et al.'s study may be due to the fact that they included all participants who used and did not use mental health services for depression, whereas the data we used to develop and validate the algorithms included only those who used mental health services for depression. Gender, age, and negative life events were not associated with recurrence in the Netherlands Study of Depression and Anxiety.[8] However, in the Netherlands Mental Health Survey and Incidence Study, younger age, negative youth experiences,

TABLE 2. Prediction algorithm for recurrence/relapse of major depression

Predictors	Coefficients In development data	Shrunken coefficients (SF = 0.8731)
Female sex	1.008454	0.88048119
Age (continuous)	−0.0064933	−0.0056693
Married/common-law	0.524479	0.45792261
Divorced/separated/single	1.035842	0.90439365
White	0.3511896	0.30662364
Had MDD last year	0.2255779	0.19695206
2 Depressive episodes	0.2256867	0.19704706
3+ Depressive episodes	0.2663524	0.23255228
Lifetime GAD or specific phobia	0.5212917	0.45513978
Avoidant personality disorder (*avoid*)	1.521121	1.3280907
Depressive symptoms in MDE		
Difficulties in concentration	0.5412085	0.47252914
Wanted to eat more	0.2653747	0.23169865
Felt guilty (*guilty*)	0.8180665	0.71425386
Took medication for low mood	0.2690194	0.23488084
SF-12 physical disability scores		
53.9–57.8	0.5816857	0.50786978
43.3–53.8	1.438105	1.2556095
0–43.2	1.619849	1.4142902
SF-12 mental disability scores		
48.4–54.5	0.484173	0.42273145
37.7–48.3	0.6907672	0.60310884
0–37.6	0.9452507	0.82529839
Experience of racial discrimination	0.1294431	0.11301677
Ever physically attacked/beaten/injured	1.520656	1.3276848
By spouse, partner, or anyone else (*abuse*)		
Experience of sexual assault	0.2462544	0.21500472
Before 18, parents/caregiver swear, insult, or say hurtful things to you		
Almost never/sometimes	0.1181686	0.103173
Fairly often/very often	0.6303988	0.55040119
Before 10, being left alone/unsupervised by parents/care givers		
Almost never/sometimes	0.1497462	0.13074341
Fairly often/very often	0.4963712	0.43338169
Interaction terms		
Sex × SF-physical	−0.299002	−0.26105865
Marital × Abuse	−0.7820871	−0.68284025
Race × Avoid	−1.113202	−0.97193667
SF-physical × Guilty	−0.270001	−0.23573787
Constant	−5.536562	−5.536562

Predicted vs. observed risk of recurrence/relapse of major depression in combined data by decile risk groups.
MDD, major depressive disorder; MDE, major depressive episode; GAD, generalized anxiety disorder.

and the presence of ongoing difficulties were associated with recurrence.[9] There may be various explanations for the inconsistencies, including different study populations, and different measurements of prognostic factors and outcome variables. This highlights the importance that, in prognostic research of major depression, diagnostic instruments and the definitions of prognostic factors should be standardized. For some clinically important factors with inconsistent findings in relation to recurrence of major depression, their prognostic values (not statistical significance) should be examined in multiple studies.

Although the developed model contained 19 unique prognostic factors, it is feasible to be used by health professionals for the following reasons. Although the predictors were selected from a battery of questionnaires and instruments, this does not mean that these questionnaires and instruments should be administered completely in order to use the risk calculator. Rather the users just need to ask the questions and enter the data related to these 19 predictors. When patients with depression are treated by physicians, a thorough psychiatric assessment and psychosocial inquiry are often conducted. Most of the data related to the predictors in model are available at this stage and will remain constant and do not fluctuate with the course of a current episode. The data that may change over the course are age, marital status, and presence or absence of residual physical and mental health symptoms (the two prognostic factors in the model) that are assessed by the SF-12. This

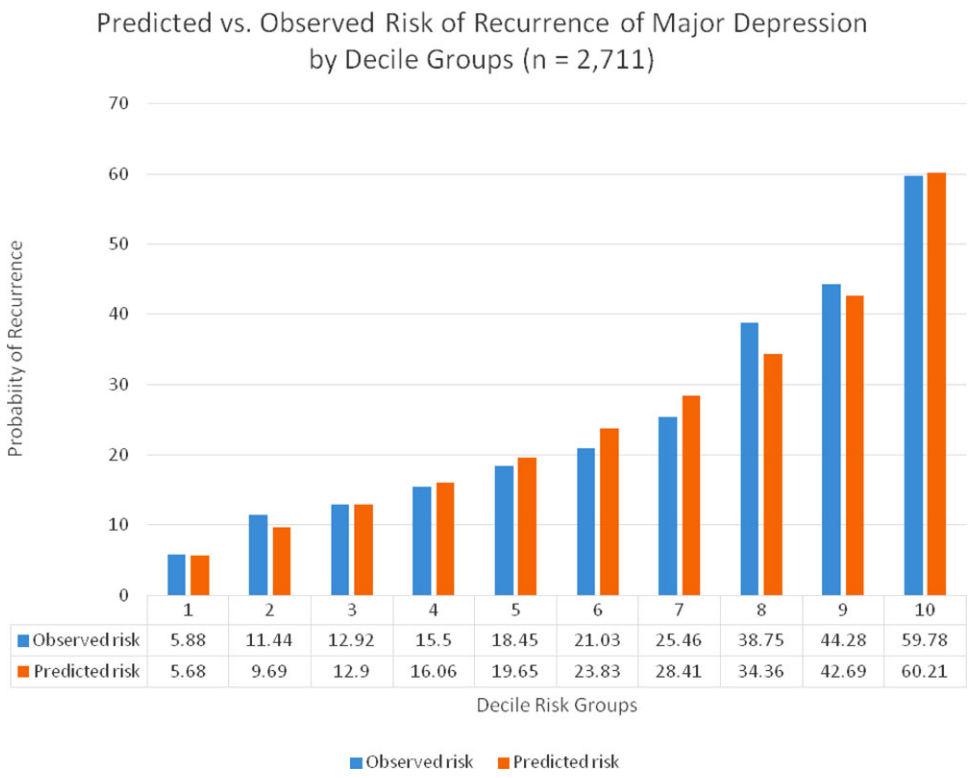

Figure 1. Perceived versus observed risk of recurrence of major depression by decile groups ($n = 2,711$).

makes frequent assessments in clinical settings feasible. Health professionals would not be expected to manually calculate the predicted risk using the coefficients presented in Table 2, rather a computerized or web-based version of this algorithm using these coefficients would calculate the predicted risk instantly upon input of the relevant variables.[38] The screen shots of the web version are in Supporting Information Fig. S1. With frequent assessment, a chart that depicts the changes of risk over time can be constructed (see Supporting Information Fig. S1). This algorithm could also be programmed into an electronic health record, smartphones, or tablet computers.

The strength of this study is that the data were population-based and the sample size was large. Furthermore, the developed prognostic model was externally validated. This study also has limitations, including the fact that the NESARC relied on self-report, possibly increasing the risk of reporting and recall biases. Such biases may also contribute to the inconsistencies in the predictive power of some prognostic factors in different regions/populations. Another potential limitation is that both participants with current and lifetime major depression at wave 1 were included in this study. Restriction to participants who had recently remitted MDE would lead to reduced samples for algorithm development and validation and imprecise prediction. Nevertheless, in our additional analysis, we applied the algorithm in those who had MDE in the past 12 months at wave 1 ($n = 419$), and found that the predicted and observed risks were the same (33.41%). Although the algorithm was validated in participants from different census regions, the performance of the model has not been validated in populations outside of the United States. This should be a direction for future research.

Reducing recurrence of depressive episodes is the priority and challenge of maintenance treatment. The prediction algorithm we developed and validated is a tool that can assist physicians in accurately quantifying the absolute risk of recurrence in the future. This will enable physicians to plan personalized treatment strategies, leading to better outcome of patients with major depression.

Acknowledgments. This analysis was conducted in Calgary, Canada, and was supported by an operating grant from the Canadian Institutes of Health Research (CIHR). However, the CIHR has no role in design, analysis, and interpretation of this study.

REFERENCES

1. Frank E, Kupfer DJ, Perel JM, et al. Three-year outcomes for maintenance therapies in recurrent depression. Arch Gen Psychiatry 1990;47:1093–1099.
2. Keller MB, Lavori PW, Lewis CE, et al. Predictors of relapse in major depressive disorder. JAMA 1983;250:3299–3304.

3. Mueller TI, Leon AC, Keller MB, et al. Recurrence after recovery from major depressive disorder during 15 years of observational follow-up. Am J Psychiatry 1999;156:1000–1006.
4. Patten SB, Kennedy SH, Lam RW, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) clinical guidelines for the management of major depressive disorder in adults. I. Classification, burden and principles of management. J Affect Disord 2009;117(Suppl 1):S5–S14.
5. National Institute for Health and Clinical Excellence. Depression in adults: the treatment and management of depression in adults. 2009.
6. Lam RW, Kennedy SH, Grigoriadis S, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) clinical guidelines for the management of major depressive disorder in adults. III. Pharmacotherapy. J Affect Disord 2009;117(Suppl 1):S26–S43.
7. Colman I, Naicker K, Zeng Y, et al. Predictors of long-term prognosis of depression. CMAJ 2011;183:1969–1976.
8. Hardeveld F, Spijker J, de Graaf R, et al. Recurrence of major depressive disorder across different treatment settings: results from the NESDA study. J Affect Disord 2012.
9. Hardeveld F, Spijker J, de Graaf R, et al. Recurrence of major depressive disorder and its predictors in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). Psychol Med 2013;43:39–48.
10. Wang JL, Patten SB, Currie S, et al. Predictors of 1-year outcomes of major depressive disorder among individuals with a lifetime diagnosis: a population-based study. Psychol Med 2012;42:327–334.
11. Patten SB, Williams JV, Lavorato DH, et al. Depressive episode characteristics and subsequent recurrence risk. J Affect Disord 2012;140:277–284.
12. Fava GA, Ruini C, Belaise C. The concept of recovery in major depression. Psychol Med 2007;37:307–317.
13. Gilman SE, Trinh NH, Smoller JW, et al. Psychosocial stressors and the prognosis of major depression: a test of Axis IV. Psychol Med 2013;43:303–316.
14. Steyerberg EW, Moons KG, van der Windt DA, et al. Prognosis Research Strategy (PROGRESS) 3: prognostic model research. PLoS Med 2013;10:e1001381.
15. Anderson KM, Wilson PW, Odell PM, et al. An updated coronary risk profile. A statement for health professionals. Circulation 1991;83:356–362.
16. Chen J, Pee D, Ayyagari R, et al. Projecting absolute invasive breast cancer risk in white women with a model that includes mammographic density. J Natl Cancer Inst 2006;98:1215–1226.
17. Park Y, Freedman AN, Gail MH, et al. Validation of a colorectal cancer risk prediction model among white patients age 50 years and older. J Clin Oncol 2009;27:694–698.
18. Compton WM, Grant BF, Colliver JD, et al. Prevalence of marijuana use disorders in the United States: 1991–1992 and 2001–2002. JAMA 2004;291:2114–2121.
19. Hasin DS, Goodwin RD, Stinson FS, et al. Epidemiology of major depressive disorder: results from the National Epidemiologic Survey on Alcoholism and Related Conditions. Arch Gen Psychiatry 2005;62:1097–1106.
20. Grant BF, Goldstein RB, Chou SP, et al. Sociodemographic and psychopathologic predictors of first incidence of DSM-IV substance use, mood and anxiety disorders: results from the Wave 2 National Epidemiologic Survey on Alcohol and Related Conditions. Mol Psychiatry 2009;14:1051–1066.
21. Skodol AE, Grilo CM, Keyes KM, et al. Relationship of personality disorders to the course of major depressive disorder in a nationally representative sample. Am J Psychiatry 2011;168:257–264.
22. Ruan WJ, Goldstein RB, Chou SP, et al. The alcohol use disorder and associated disabilities interview schedule-IV (AUDADIS-IV): reliability of new psychiatric diagnostic modules and risk factors in a general population sample. Drug Alcohol Depend 2008;92:27–36.
23. El-Gabalawy R, Katz LY, Sareen J. Comorbidity and associated severity of borderline personality disorder and physical health conditions in a nationally representative sample. Psychosom Med 2010;72:641–647.
24. Ware J, Jr., Kosinski M, Keller SD. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. Med Care 1996;34:220–233.
25. Straus MA. Measuring intrafamily conflict and violence: the Conflict Tactics (CT) scales. In: Straus MA, Gelles RJ, editors. Physical Violence in American Families: Risk Factors and Adaptations to Violence in 8,125 Families. New Brunswick, NJ: Transaction; 1990:29–47.
26. Bernstein DP, Fink L, Handelsman L, et al. Initial reliability and validity of a new retrospective measure of child abuse and neglect. Am J Psychiatry 1994;151:1132–1136.
27. Krieger N, Smith K, Naishadham D, et al. Experiences of discrimination: validity and reliability of a self-report measure for population health research on racism and health. Soc Sci Med 2005;61:1576–1596.
28. NESARC. Wave 2 NESARC data notes. 2008.
29. StataCorp. Stata Statistical Software: Release 12.0. College Station, TX: Stata Corporation; 2012.
30. Royston P, Moons KG, Altman DG, et al. Prognosis and prognostic research: developing a prognostic model. BMJ 2009;338:b604.
31. Steyerberg EW. Clinical Prediction Models: A Practical Approach to Development, Validation and Updating. New York: Springer; 2009.
32. Cook NR. Statistical evaluation of prognostic versus diagnostic models: beyond the ROC curve. Clin Chem 2008;54:17–23.
33. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology 2010;21:128–138.
34. Pencina MJ, D'Agostino RB, Sr., D'Agostino RB, Jr., et al. 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.
35. van Houwelingen JC, le Cessie S. Predictive value of statistical models. Stat Med 1990;9:1303–1325.
36. Grilo CM, Stout RL, Markowitz JC, et al. Personality disorders predict relapse after remission from an episode of major depressive disorder: a 6-year prospective study. J Clin Psychiatry 2010;71:1629–1635.
37. Mulder RT, Frampton CM, Luty SE, et al. Eighteen months of drug treatment for depression: predicting relapse and recovery. J Affect Disord 2009;114:263–270.
38. Wang JL, Manuel D, Williams J, et al. Development and validation of prediction algorithms for major depressive episode in the general population. J Affect Disord 2013;151:39–45.