

# Research Article

## MAJOR DEPRESSIVE DISORDER SUBTYPES TO PREDICT LONG-TERM COURSE

Hanna M. van Loo, M.D.,<sup>1</sup> Tianxi Cai, Sc.D.,<sup>2</sup> Michael J. Gruber, M.S.,<sup>3</sup> Junlong Li, Ph.D.,<sup>2</sup> Peter de Jonge, Ph.D.,<sup>1</sup> Maria Petukhova, Ph.D.,<sup>3</sup> Sherri Rose, Ph.D.,<sup>3</sup> Nancy A. Sampson, B.A.,<sup>3</sup> Robert A. Schoevers, M.D., Ph.D.,<sup>1</sup> Klaas J. Wardenaar, Ph.D.,<sup>1</sup> Marsha A. Wilcox, Ed.D., Sc.D.,<sup>4</sup> Ali Obaid Al-Hamzawi, M.D., F.I.C.M.S.,<sup>5</sup> Laura Helena Andrade, M.D., Ph.D.,<sup>6</sup> Evelyn J. Bromet, Ph.D.,<sup>7</sup> Brendan Bunting, Ph.D.,<sup>8</sup> John Fayyad, M.D.,<sup>9</sup> Silvia E. Florescu, M.D., Ph.D.,<sup>10</sup> Oye Gureje, M.D., Ph.D., D.Sc., F.R.C.Psych.,<sup>11</sup> Chiyi Hu, M.D., Ph.D.,<sup>12</sup> Yueqin Huang, M.D., M.P.H., Ph.D.,<sup>13</sup> Daphna Levinson, Ph.D.,<sup>14</sup> Maria Elena Medina-Mora, Ph.D.,<sup>15</sup> Yoshibumi Nakane, M.D., Ph.D.,<sup>16</sup> Jose Posada-Villa, M.D.,<sup>17</sup> Kate M. Scott, Ph.D.,<sup>18</sup> Miguel Xavier, M.D., Ph.D.,<sup>19</sup> Zahari Zarkov, M.D.,<sup>20</sup> and Ronald C. Kessler, Ph.D.<sup>3\*</sup>

**Background:** *Variation in the course of major depressive disorder (MDD) is not strongly predicted by existing subtype distinctions. A new subtyping approach is considered here. Methods:* *Two data mining techniques, ensemble recursive partitioning and Lasso generalized linear models (GLMs), followed by k-means cluster analysis are used to search for subtypes based on index episode symptoms predicting subsequent MDD course in the World Mental Health (WMH) surveys. The WMH surveys are community surveys in 16 countries. Lifetime DSM-IV MDD was reported by 8,261 respondents. Retrospectively reported outcomes included measures of persistence (number of years with an episode, number of years with an episode lasting most of the year) and severity (hospitalization for MDD, disability due to MDD). Results:* *Recursive partitioning found*

<sup>1</sup>Department of Psychiatry, University Medical Center Groningen, University of Groningen, Groningen, The Netherlands

<sup>2</sup>Department of Biostatistics, Harvard School of Public Health, Boston, Massachusetts

<sup>3</sup>Department of Health Care Policy, Harvard Medical School, Boston, Massachusetts

<sup>4</sup>Johnson and Johnson Pharmaceutical Research and Development, Titusville, New Jersey

<sup>5</sup>Department of Psychiatry, Al-Qadisia University College of Medicine, Diwania, Iraq

<sup>6</sup>Section of Psychiatric Epidemiology—LIM 23, Department and Institute of Psychiatry, University of São Paulo Medical School, São Paulo, Brazil

<sup>7</sup>State University of New York at Stony Brook, Stony Brook, New York

<sup>8</sup>Psychology Research Institute, University of Ulster, Londonderry, United Kingdom

<sup>9</sup>Institute for Development Research, Advocacy, and Applied Care, St. George Hospital University Medical Center, Beirut, Lebanon

<sup>10</sup>National School of Public Health, Management and Professional Development, Bucharest, Romania

<sup>11</sup>Department of Psychiatry, University College Hospital, Ibadan, Nigeria

<sup>12</sup>Shenzhen Institute of Mental Health, Shenzhen Kangning Hospital, Guangdong, People's Republic of China

<sup>13</sup>Institute of Mental Health, Peking University, Beijing, People's Republic of China

<sup>14</sup>Department of Research and Planning, Mental Health Services, Ministry of Health, Jerusalem, Israel

<sup>15</sup>Instituto Nacional de Psiquiatria Ramon de la Fuente, Mexico City, Mexico

<sup>16</sup>Department of Social Work, Faculty of Human Sociology, Nagasaki International University, Nagasaki, Japan

<sup>17</sup>Universidad Colegio Mayor de Cundinamarca, Bogota, Colombia

<sup>18</sup>Department of Psychological Medicine, Otago University, Dunedin, New Zealand

<sup>19</sup>Department of Mental Health, Universidade Nova de Lisboa, Lisbon, Portugal

<sup>20</sup>Department of Mental Health, National Center of Public Health and Analyses, Sofia, Bulgaria

\*Correspondence to: Ronald C. Kessler, Department of Health Care Policy, Harvard Medical School, 180 Longwood Avenue, Boston, MA 02115. E-mail: NCS@hcp.med.harvard.edu

Received for publication 19 November 2013; Revised 05 December 2013; Accepted 07 December 2013

DOI 10.1002/da.22233

Published online in Wiley Online Library (wileyonlinelibrary.com).

*significant clusters defined by the conjunctions of early onset, suicidality, and anxiety (irritability, panic, nervousness–worry–anxiety) during the index episode. GLMs found additional associations involving a number of individual symptoms. Predicted values of the four outcomes were strongly correlated. Cluster analysis of these predicted values found three clusters having consistently high, intermediate, or low predicted scores across all outcomes. The high-risk cluster (30.0% of respondents) accounted for 52.9–69.7% of high persistence and severity, and it was most strongly predicted by index episode severe dysphoria, suicidality, anxiety, and early onset. A total symptom count, in comparison, was not a significant predictor. Conclusions: Despite being based on retrospective reports, results suggest that useful MDD subtyping distinctions can be made using data mining methods. Further studies are needed to test and expand these results with prospective data. Depression and Anxiety 1:1–13, 2014. © 2014 Wiley Periodicals, Inc.*

**Key words:** *epidemiology; depression; anxiety/anxiety disorders; suicide/self-harm; panic attacks*

## INTRODUCTION

Patients with major depressive disorder (MDD) vary substantially in treatment response and illness course. Recognition of this variation has led researchers to search for depression subtypes defined either by presumed causes (e.g., postnatal depression),<sup>[1,2]</sup> clinical presentation (e.g., atypical or melancholic depression,<sup>[3,4]</sup>) or empirically derived symptom profiles using cluster analysis,<sup>[5]</sup> factor analysis,<sup>[6]</sup> or latent class analysis,<sup>[7]</sup> in hopes that patients in subtypes would be sufficiently similar in psychopathological processes to help identify underlying molecular etiologies or predict treatment response.<sup>[7–9]</sup> However, subtyping distinctions up to now have not lived up to these expectations,<sup>[8,10]</sup> although some commentators suggest that subtyping using endophenotypes or intermediate phenotypes might hold more promise.<sup>[11,12]</sup>

Another potentially useful approach to subtyping, given the goal of prediction, would be to define subtypes using recursive partitioning<sup>[13,14]</sup> and related data mining methods<sup>[15,16]</sup> that search for synergistic associations of predictors with illness course. Such methods have been used in other areas of medicine<sup>[17,18]</sup> and relatively simple applications have been used in psychiatry to predict depression treatment response<sup>[19–23]</sup> and suicidality.<sup>[24–26]</sup>

This report presents results of preliminary analyses designed to find symptom-based subtypes predicting course of MDD using more complex data mining methods than in previous studies. The analysis is preliminary because it uses retrospective data on depression course collected in cross-sectional population epidemiological surveys rather than longitudinal clinical studies. Results are nonetheless useful in providing a proof of concept of the approach in a large and diverse sample of subjects who were asked about potentially important subtyping variables in their index episodes and assessed for multi-

ple indicators of subsequent depression persistence and severity.

## MATERIALS AND METHODS

### SAMPLE

Data come from the World Health Organization World Mental Health (WMH) surveys ([www.hcp.med.harvard.edu/wmh](http://www.hcp.med.harvard.edu/wmh)), a series of well-characterized community epidemiological surveys<sup>[27–30]</sup> administered in six countries classified by the World Bank as high income (Israel, Japan, New Zealand, Northern Ireland, Portugal, United States), five upper-middle income (Brazil, Bulgaria, Lebanon, Mexico, Romania), and five low/lower-middle income (Colombia, Iraq, Nigeria, People's Republic of China, Ukraine).<sup>[31]</sup> Most surveys feature nationally representative household samples, whereas two (Colombia, Mexico) represent all urban areas in the country, one selected states (Nigeria), and three selected metropolitan areas (Brazil, Japan, People's Republic of China); (Table 1). A total of 93,167 adults (age 18+ years) participated, 8,261 of whom met lifetime DSM-IV criteria for MDD. Sample sizes range from 2,357 (Romania) to 12,790 (New Zealand). The average weighted response rate was 73.7% (range: 55.1–95.2%). Weights adjusted for differential probabilities of selection and discrepancies with population sociodemographic/geographic distributions. Further details about WMH sampling and weighting are available elsewhere.<sup>[32]</sup>

### MEASURES

**Interview Procedures.** Translation, back-translation, and harmonization of the interview schedule used standardized procedures.<sup>[33]</sup> Interviews were fully structured and administered face-to-face in the homes of respondents by trained lay interviewers. Rigorous interviewer training and quality-control procedures were employed.<sup>[34]</sup> The research presented here is in compliance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). The institutional review board of the organization that coordinated the survey in each country approved and monitored compliance with procedures for obtaining informed consent and protecting human subjects.

**MDD.** DSM-IV MDD was assessed with the Composite International Diagnostic Interview (CIDI), version 3.0,<sup>[35]</sup> a fully structured diagnostic interview designed for administration by trained

**TABLE 1. WMH sample characteristics by World Bank income categories<sup>a</sup>**

Country by income category	Survey <sup>b</sup>	Sample characteristics <sup>c</sup>	Field dates	Age range (years)	Response rate <sup>d</sup>	Sample size	
						Total	With MDD
<b>I. High-income countries</b>							
Israel	NHS	Nationally representative.	2002–4	21–98	72.6	4,859	284
Japan	WMHJ 2002–2006	11 metropolitan areas.	2002–6	20–98	55.1	4,129	219
New Zealand <sup>f</sup>	NZMHS	Nationally representative.	2003–4	18–98	73.3	12,790	1,908
N. Ireland	NISHS	Nationally representative.	2004–7	18–97	68.4	4,340	423
Portugal	NMHS	Nationally representative.	2008–9	18–81	57.3	3,849	379
United States	NCS-R	Nationally representative.	2002–3	18–99	70.9	9,282	1,562
Total					<b>67.9</b>	<b>(39,249)</b>	<b>(4,775)</b>
<b>II. Upper-middle-income countries</b>							
Brazil—São Paulo	São Paulo Megacity	São Paulo metropolitan area.	2005–7	18–93	81.3	5,037	408
Bulgaria	NSHS	Nationally representative.	2003–7	18–98	72.0	5,318	283
Lebanon	L.E.B.A.N.O.N	Nationally representative.	2002–3	18–94	70.0	2,857	267
Mexico	M-NCS	All urban areas of the country (approximately 75% of the total national population).	2001–2	18–65	76.6	5,782	397
Romania	RMHS	Nationally representative.	2005–6	18–96	70.9	2,357	54
Total					<b>74.8</b>	<b>(21,351)</b>	<b>(1,409)</b>
<b>III. Low-income and lower-middle-income countries</b>							
Colombia	NSMH	All urban areas of the country (approximately 73% of the total national population)	2003	18–65	87.7	4,426	476
Iraq	IMHS	Nationally representative.	2006–7	18–96	95.2	4,332	193
Nigeria	NSMHW	21 of the 36 states in the country, representing 57% of the national population. The surveys were conducted in Yoruba, Igbo, Hausa, and Efik languages.	2002–3	18–100	79.3	6,752	176
PRC*—Beijing/Shanghai	B-WMH/S-WMH	Beijing and Shanghai metropolitan areas.	2002–3	18–70	74.7	5,201	151
PRC*—Shenzhen <sup>f</sup>	Shenzhen	Shenzhen metropolitan area.	2006–7	18–88	80.0	7,132	452
Ukraine <sup>f</sup>	CMDPSD	Included temporary residents as well as household residents.	2002	18–91	78.3	4,724	629
Total					<b>81.4</b>	<b>(32,567)</b>	<b>(2,077)</b>

<sup>a</sup>The World Bank (2012). Data. Accessed June 5, 2012 at: <http://data.worldbank.org/country>.

<sup>b</sup>NSMH, Colombian National Study of Mental Health; IMHS, Iraq Mental Health Survey; NSMHW, Nigerian Survey of Mental Health and Wellbeing; B-WMH, Beijing World Mental Health Survey; S-WMH, Shanghai World Mental Health Survey; CMDPSD, Comorbid Mental Disorders during Periods of Social Disruption; NSHS, Bulgaria National Survey of Health and Stress; L.E.B.A.N.O.N., Lebanese Evaluation of the Burden of Ailments and Needs of the Nation; M-NCS, Mexico National Comorbidity Survey; RMHS, Romania Mental Health Survey; NHS, Israel National Health Survey; WMHJ 2002–2006, World Mental Health Japan Survey; NZMHS, New Zealand Mental Health Survey; NISHS, Northern Ireland Study of Health and Stress; NMHS, Portugal National Mental Health Survey; NCS-R, U.S. National Comorbidity Survey Replication.

<sup>c</sup>Most WMH surveys are based on stratified multistage clustered area probability household samples in which samples of areas equivalent to counties or municipalities in the United States were selected in the first stage followed by one or more subsequent stages of geographic sampling (e.g., towns within counties, blocks within towns, households within blocks) to arrive at a sample of households, in each of which a listing of household members was created and one or two people were selected from this listing to be interviewed. No substitution was allowed when the originally sampled household resident could not be interviewed. These household samples were selected from Census area data in all countries. The Japanese sample is the only totally unclustered sample, with households randomly selected in each of the 11 metropolitan areas and one random respondent selected in each sample household. Ten of the sixteen surveys are based on nationally representative household samples.

<sup>d</sup>The response rate is calculated as the ratio of the number of households in which an interview was completed to the number of households originally sampled, excluding from the denominator households known not to be eligible either because of being vacant at the time of initial contact or because the residents were unable to speak the designated languages of the survey. The weighted average response rate is 73.7%.

<sup>e</sup>People's Republic of China.

<sup>f</sup>For the purposes of cross-national comparisons, we limit the sample to those 18+.

lay interviewers. The CIDI translation, back-translation, and harmonization protocol required culturally competent bilingual clinicians to review, modify, and approve key phrases describing symptoms. Clinical reappraisal studies conducted in several WMH countries found good concordance between lifetime DSM-IV/CIDI diagnoses of major depression and independent diagnoses based on blinded Structured Clinical Interview for DSM-IV clinical reappraisal interviews,<sup>[36]</sup> with area under the ROC curve (AUC) averaging .75 and likelihood ratio positive averaging 8.8 (a level close to the threshold considered definitive for ruling in a clinical diagnosis from a screen).<sup>[37]</sup>

Respondents with lifetime DSM-IV/CIDI MDD were asked retrospective questions about age of onset (AOO), whether their first lifetime depressive episode “was brought on by some stressful experience” or happened “out of the blue,” all DSM-IV Criterion A–D symptoms of major depressive episode for the index episode (including separate questions about weight loss and weight gain, insomnia and hypersomnia, psychomotor agitation and retardation, and thoughts of death, suicide ideation, suicide plans, and suicide gestures–attempts), ICD-10 (International Classification of Diseases) severity specifiers, questions to operationalize diagnostic hierarchy rule exclusions, and questions about symptoms during the index episode that might be markers of (i) dysthymia (inability to cope, social withdrawal), (ii) mixed episodes (sleep much less than usual and still not feel tired, racing thoughts), and (iii) anxious depression (feeling irritable, nervous–anxious–worried, having sudden attacks of intense fear or panic).

Four retrospective questions were asked about subsequent lifetime MDD course: number of years since AOO when the respondent had an episode (i) lasting 2 weeks or longer or (ii) lasting *most days* throughout the year; (iii) a dichotomous measure of whether the depression was ever so severe that the respondent was hospitalized overnight (and, if so, age of first hospitalization); and (iv) a dichotomous measure of whether the respondent was currently disabled (at least 50% limitation in ability to perform paid work) because of depression. These are the four outcomes considered here. The two measures of years in episode were divided by number of years between age at interview (AAI) and AAO+1 to create continuous outcomes in the range 0–100%.

**Other Predictors.** In addition to the information described above about the index episode, additional predictors included discretized information about the respondent’s AOO in eight nested age categories selected for sensitivity in the age range with most onsets (less than or equal to ages 12, 15, 19, 24, 29, 34, 39, and 59 years), similarly nested and discretized information about AAI–AOO, and a binary variable for respondent Family History Research Diagnostic Criteria Interview<sup>[38]</sup> reports for whether respondents’ parents had a history of major depression.

## ANALYSIS METHODS

Analysis of the deidentified WMH master dataset was approved by the Institutional Review Board of Harvard Medical School, the site of the WMH Data Coordination Center. An ensemble of 100 classification trees was used to find important interactions among predictors of the outcomes. The ensemble approach (i.e., combining results across a large number of replicates, each replicate estimated in a different simulated pseudosample) was used to reduce risk of overfitting.<sup>[13–15]</sup> The recursive partitioning R package *rpart*<sup>[39]</sup> was used for this purpose. The minimum number of observations in a node for further splitting was set at 20 and the threshold complexity parameter (cp) at 0.01. The models to predict years in episode, which used a Poisson link function, were estimated among respondents where AAI–AOO was either 10+ years (years with episodes lasting most of the year) or 15+ years (years with any episode) based on preliminary inspection showing that outcome scores stabilized after these cut points. Proportional hazards survival models were used to predict age at first hospitalization for depression among respondents who were not hospitalized for depres-

sion at AOO. Logistic regression models were used to predict current disability in the total sample.

Each tree in the ensemble was built in a randomly selected bootstrap sample drawn without replacement from the sample and cross-validated among the remaining respondents to determine appropriate tree depth. Inspection of summary frequencies of unique terminal nodes (i.e., subgroups of respondents defined by the conjunction of the dichotomous predictors selected to optimize prediction of the outcome) across the 100 trees was used to select the interactions to retain in a second step of analysis. This second step fitted a separate generalized linear model (GLM) for the multivariate associations of all predictors with each outcome. Included here were additive associations of the individual predictors with the outcome, the interactions found to occur repeatedly in the tree models, and nested dichotomies to describe the total number of symptoms endorsed. The inclusion of the latter predictors was important to distinguish differential predictive effects of especially important symptoms from predictive effects of an overall symptom count.

As some of the predictors in the GLM models were highly correlated, conventional regression methods yielded unstable results. Stepwise regression,<sup>[40]</sup> which is often used to address this problem, overfits and performs poorly in new samples.<sup>[41]</sup> A number of data mining methods have been developed to improve on stepwise regression. We used one such method, the Lasso,<sup>[42]</sup> to address this problem. The Lasso is one of several *penalized regression* methods that trades off bias to increase the efficiency of estimation by constraining the sum of variance of nonzero values of standardized regression coefficients with coefficient shrinkage parameters. We selected Lasso instead of alternatives, as this penalty handles high correlations among predictors by yielding a sparse model (i.e., forces coefficients of weak predictors to zero).<sup>[43]</sup> The R-package *glmnet*<sup>[44]</sup> was used to estimate the Lasso GLMs using the same link functions as in the regression tree models. Coefficients from the Lasso models were exponentiated to create incidence density ratios (IDRs) to predict proportion of years in episode, hazard ratios (HRs) to predict hospitalization, and odds ratios (ORs) to predict disability. No confidence intervals were generated, as standard errors in such models are biased.

The best-fitting Lasso coefficients were then used to generate predicted values of each outcome for all respondents. Based on evidence of strong correlations among these predicted values across outcomes, *k*-means cluster analysis was used to partition the sample into subtypes with similar multivariate profiles of predicted scores across the four outcomes using the R-package *stats*<sup>[45]</sup> and using 100 random starts for each number of clusters. Inspection of observed (as opposed to predicted) mean dichotomized outcome scores (percentages of respondents with high persistence and chronicity, hospitalization, and disability) and calculation of AUC (adjusted appropriately for the survival outcome<sup>[46]</sup>) were used to select an optimal number of clusters. Associations of cluster membership with dichotomized versions of outcomes were then examined by calculating relative risk of the adverse outcomes in the high-risk versus other clusters, positive predictive value (PPV; the proportion of high-risk cluster respondents that experienced the adverse outcomes), and sensitivity (SN; the proportion of all adverse outcomes that occurred in the high-risk cluster).

## RESULTS

### DISTRIBUTIONS OF THE OUTCOMES

The mean, median, and interquartile range (25th–75th percentiles) percentages of years after AOO when respondents in the analysis sample reported having a depressive episode lasting 2 weeks or longer were 25.8, 13.0, and 6.2–29.4%, respectively. The comparable



**TABLE 2. Lasso GLM coefficients to predict subsequent course of DSM-IV MDD based on characteristics of the incident episode<sup>a</sup>**

	Percentage of years in episode		Hospitalized HR <sup>b</sup>	Disabled OR <sup>b</sup>
	Any episode IDR <sup>b</sup>	Episode lasting most of year IDR <sup>b</sup>		
<b>I. Criterion A symptoms of major depression</b>				
Severe dysphoria <sup>c</sup> (ICD-10 severity specifier)	1.1	1.4	1.7	1.2
Anhedonia				1.1
Weight loss	0.9			
Weight gain		1.1		0.8
Insomnia			1.3	
Hypersomnia			1.5	
Psychomotor agitation			1.2	
Psychomotor retardation				1.2
Suicidality		1.1	1.6	1.5
<b>II. Symptoms of dysthymia</b>				
Inability to cope			1.9	1.4
<b>III. Symptoms of anxiety</b>				
Irritability	1.1		0.8	1.2
Panic		1.1	1.3	1.5
<b>IV. Symptoms of mixed episode</b>				
Racing thoughts		0.8		
High energy			1.2	
<b>V. Multivariate symptom profiles</b>				
AOO <19 and suicidality				1.3
AOO <19 and anxiety	1.1	1.3	1.2	
AOO ≥35 and suicidality and anxiety		1.2		
<b>VI. Other predictors<sup>d</sup></b>				
Endogenous				0.7
Parental history of depression		1.2	1.2	1.2
<i>N</i>	(2,869)	(3,958)	(6,465)	(8,261)

<sup>a</sup>Based on Lasso GLM penalized regression models, with the size of penalty determined by 10-fold cross-validation to select the penalty yielding cross-validating results with minimum mean squared prediction error. No confidence intervals are reported because standard errors of such simulated models are biased. See the text for a discussion of differences in link functions and sample sizes.

<sup>b</sup>IDR, incidence density ratio; HR, hazard ratio; OR, odds ratio.

<sup>c</sup>This is not the DSM-IV Criterion A symptom of dysphoria but the ICD-10 symptom for somatic depression that the dysphoria is so severe that the patient has a lack of emotional reaction to events or activities that normally produce an emotional response. The DSM-IV symptom of dysphoria, in comparison, was not a significant predictor in any of the models.

<sup>d</sup>An additional 12 predictors were included in the Lasso GLM models that had coefficients of either zero or near zero across all outcomes. These predictors are dysphoria, fatigue/loss of energy, worthlessness or excessive guilt, diminished ability to concentrate or indecisiveness, social withdrawal, nervousness-worry-anxiety, multivariate symptoms profiles of childhood (before age 13 years) onset with anxiety and/or suicidality, multivariate symptom profiles of AOO before 19 years with anxiety and suicidality, other multivariate symptom profiles of AOO either before 13 years or before 19 years or after 34 years with either anxiety and/or suicidality, little need for sleep, total number of symptoms, age of onset, and time between onset and age at interview.

percentages for years having a depressive episode lasting *most days* throughout the year were 9.5, 0.0, and 0.0–9.3%. Lifetime hospitalization for a depressive episode was reported by 4.3% of respondents and current disability due to depression was reported by 1.6% of respondents.

**RECURSIVE PARTITIONING**

The terminal nodes repeatedly predicting outcomes in recursive partitioning all involved two-way or three-way interactions between child-adolescent (before age 19) AOO, suicidality, and anxiety (nervous-anxious-worried, irritable, attacks of fear-panic) during index de-

pressive episodes. The conjunction of later AOO (age 35+ years) with anxiety and suicidality also predicted chronicity. The cells defined by the conjunction of early onset, suicidality, and anxiety had either the highest or, in one case (disability), second-highest scores on all outcomes across cells of the table defined by these predictors (detailed results are available on request). Based on these results, all two-way and three-way interactions among AOO, anxiety, and suicidality were included in the Lasso GLMs.

**LASSO GLMS**

Four predictors of persistence, eight of chronicity, and eleven each of hospitalization and disability were

retained in the GLMs with Lasso coefficients meaningfully different from zero (Table 2). The vast majority (85%) of these coefficients was positive. The positive IDRs for years in episode were in the range 1.1–1.4. The positive HRs for hospitalization and ORs for disability were in the range 1.1–1.9. Only one predictor, severe dysphoria, was retained in all four models. Severe dysphoria was also the strongest predictor of chronicity (IDR = 1.4) and one of the strongest predictors of hospitalization (OR = 1.7). Four other predictors with consistently positive coefficients retained in three of the four models included suicidality (1.1–1.6), panic attacks (1.1–1.5), the multivariate profile of pediatric onset and anxiety (either nervousness–anxiety–worry or panic, 1.1–1.3), and parental history of major depression (1.2). One of these four, suicidality, was also among the strongest predictors of hospitalization (HR = 1.6) and disability (OR = 1.5), whereas panic was one of the strongest predictors of disability (OR = 1.5). Other strong predictors of hospitalization included inability to cope (HR = 1.9) and hypersomnia (HR = 1.5), whereas inability to cope was also one of the strongest predictors of disability (OR = 1.4). Early-AOO suicidality also predicted disability, while later-AOO (older than age 34 years) suicidality predicted chronicity. The latter represented nonlinearity in the effect of the multivariate AOO anxiety–suicidality profile.

## CLUSTER ANALYSIS

Predicted values of each outcome were calculated for each respondent based on the GLM model coefficients. Spearman's rank-order correlations among these predicted values were in the range .76–.89. Principal axis exploratory factor analysis showed that the correlations were consistent with the existence of a single underlying factor (factor loadings in the range .89–.94). Based on these results, *k*-means cluster analysis of transformed (to percentiles) predicted outcome scores searched for multivariate clusters defining differential risk of the outcomes.

Inspection of mean percentile scores for solutions between three and eight clusters showed all solutions defined one class with the highest mean scores on all outcomes, a second class with lowest mean scores on all outcomes, and other classes with consistently intermediate mean scores on all outcomes (Fig. 1A–1F). Based on this observation, alternative three-cluster solutions were constructed from the original four- through eight-cluster solutions by collapsing the intermediate clusters. AUC was then compared across these solutions to predict dichotomous versions of the measures of years in episodes (distinguishing the 5–10 top percentiles of respondents with highest scores), hospitalization, and disability to see if classifications of high-risk or low-risk clusters were refined in solutions with more than three clusters. None of the collapsed solutions had higher AUCs than the original three-cluster solution (.64 for years in episode, .61 for years in episodes lasting more

than half the year, .70 for hospitalization, and .72 for disability).

The distribution of membership in the three-cluster solution was 30.7% high-risk, 35.6% intermediate-risk, and 33.7% low-risk. Respondents in the high-risk cluster were 2.1–5.1 times as likely as others and 2.5–11.3 times as likely as respondents in the low-risk cluster to have high levels of long-term MDD persistence and severity (Table 3). Respondents in the high-risk cluster included 52.9–69.7% of all those with high levels of long-term MDD persistence and severity, and 68.4–71.1% of those with two or more such adverse outcomes.

Cluster membership was strongly associated (Cramer's *V* greater than .50) with only one baseline predictor, suicidality (*V* = .54), and moderately associated (Cramer's *V* in the range .30–.50) with eight others, including one Criterion A depressive symptom (worthlessness/excessive guilt, *V* = .34), the ICD-10 severe dysphoria marker (*V* = .47), one symptom of dysthymia (inability to cope, *V* = .50), two of the three symptoms of anxiety (irritability, panic attacks, *V* = .30–.44), and the early-AOO multivariate symptom profiles retained in the Lasso GLMs (early AOO with either suicidality or anxiety, *V* = .35–.46; Table 4). Scores on these variables were consistently higher in the high-risk than intermediate-risk cluster and in the intermediate-risk than the low-risk cluster. However, proportional high-risk versus intermediate-risk differences were relatively modest in most cases (1.1–1.4 risk ratios) other than for panic (1.7) and the early-AOO multivariate symptoms profiles (2.0–3.2), while proportional intermediate-risk versus low-risk differences were consistently larger, with the highest risk-ratios for panic (2.8), inability to cope (2.5), suicidality (2.0), and the multivariate symptoms profiles (2.4–7.1).

## DISCUSSION

The above results are limited by being based on retrospective data collected in fully structured interviews excluding information on such potentially important predictors as temporally primary comorbid disorders and treatment status. Sample biases could also have been introduced by differential response related to predictors or predictor effects or differential mortality. The limitations involving use of a fully structured interview and restricted predictors almost certainly led to downward bias in the estimated strength of associations, but the other limitations could have introduced either conservative or anticonservative biases. Results should be considered only exploratory because of these limitations, although the results have value both as a proof of concept and as a source of ideas about prediction patterns that warrant analysis in future studies.

Within the context of these limitations, three results emerged that could serve as a starting point for future prospective clinical studies. First, the recursive partitioning found an early-onset anxious–suicidal subtype associated with all four outcomes (persistence, chronicity,

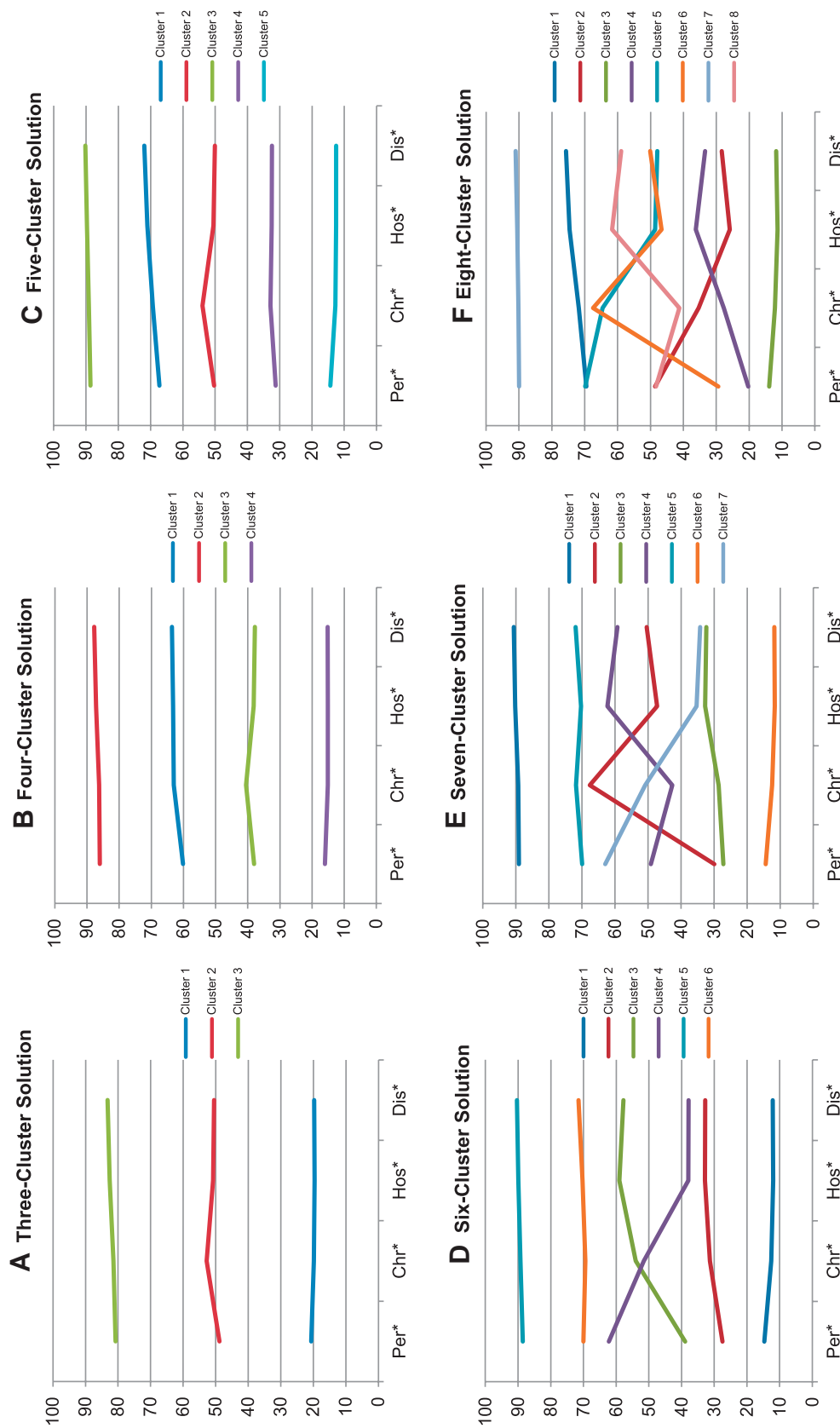


Figure 1. Mean predicted outcome scores in the three-cluster through eight-cluster *k*-means.<sup>1</sup>

**TABLE 3. Associations of cluster membership with positive screening characteristics**

	Relative risk <sup>a</sup> in the high-risk cluster vs.				Positive predictive value <sup>c</sup>		Sensitivity <sup>c</sup>	
	All others <sup>b</sup>		Those in the low-risk cluster		%	(SE)	%	(SE)
	Est	(95% CI)	Est	(95% CI)				
Percentage of years in any episode								
Top 5 percentile	2.7	(1.7–3.7)	3.3	(1.7–4.9)	7.9	(0.8)	60.0	(4.3)
Top 10 percentile	2.5	(1.9–3.1)	3.1	(1.9–4.2)	16.4	(1.1)	58.0	(3.0)
Percentage of years in episodes lasting most of the year								
Top 5 percentile	2.7	(1.8–3.5)	4.0	(1.9–6.1)	8.2	(0.8)	59.3	(3.5)
Top 10 percentile	2.1	(1.6–2.5)	2.5	(1.7–3.3)	14.4	(0.9)	52.9	(2.5)
Hospitalized	5.1	(3.4–6.7)	10.4	(4.3–16.6)	9.6	(0.8)	69.7	(3.5)
Disabled	4.6	(2.5–6.7)	11.3	(2.9–19.8)	3.4	(0.4)	67.1	(4.9)
Summary outcomes using top 5 percentile								
Any <sup>d</sup>	3.2	(2.4–3.9)	4.5	(3.0–6.0)	25.4	(1.6)	64.2	(3.0)
Multiple <sup>e</sup>	4.3	(1.8–6.9)	6.5	(0.9–12.1)	5.6	(0.8)	71.1	(5.6)
Summary outcomes using top 10 percentile								
Any <sup>d</sup>	2.6	(2.1–3.1)	3.1	(2.3–4.0)	33.0	(1.9)	59.3	(2.6)
Multiple <sup>e</sup>	3.8	(2.3–5.3)	4.8	(1.9–7.7)	11.2	(1.1)	68.4	(4.3)

<sup>a</sup>Relative risk is the ratio of the percent of respondents in the high-risk cluster that experienced the adverse outcome compared to the percent in the other clusters or in the low-risk cluster.

<sup>b</sup>Others = respondents in either the intermediate-risk or low-risk clusters.

<sup>c</sup>Positive predictive value is the percent of respondents in the high-risk cluster that experienced the adverse outcome; sensitivity is the percent of observed adverse outcomes that occurred in the high-risk cluster.

<sup>d</sup>These are dichotomous variables that differentiate respondents who had one or more of the following four adverse outcomes: in the top 5 percentile (or 10 percentile) of years with episodes, in the top 5 percentile (or 10 percentile) of years with episodes lasting most of the year, hospitalized, or disabled.

<sup>e</sup>These are dichotomous variables that differentiate respondents who had two or more of the four adverse outcomes.

hospitalization, disability) and a late-onset anxious-suicidal subtype associated with chronicity. Second, the GLMs found that a number of index episode symptoms were significant predictors of all outcomes. The most consistent and powerful of these was severe dysphoria, while others included parental history of major depression, suicidality, panic attacks, and multivariate profiles of pediatric onset with anxiety and/or suicidality. Third, strong clustering was found in these predicted values across the outcomes, with the roughly 30% of respondents in the high-risk cluster accounting for more than two-thirds of cases with multiple indicators of high long-term persistence, chronicity, and severity.

Several previous epidemiological studies examined baseline predictors of long-term course either in treatment<sup>[47,48]</sup> or community<sup>[49–51]</sup> samples, but did not attempt to search for depression subtypes. Although these studies found several replicated predictors, including co-occurring anxiety, pain-physical comorbidity, and family history of depression,<sup>[50,52–54]</sup> no attempt was made in those studies to examine synergistic effects of predictor clusters other than for summary measures of overall depression symptom number. Importantly, we included a total count of depressive symptoms in our GLMs but this measure was not significant.

As noted in the Introduction, subtyping analyses more similar to those reported here have been done to pre-

dict treatment response<sup>[19,20]</sup> and naturalistic patterns of remission among patients<sup>[23]</sup> or in the placebo control group of a depression clinical trial.<sup>[21]</sup> A number of recent clinical studies have also used methods similar to ours either to predict suicidality during<sup>[22,25,26]</sup> or after termination of<sup>[24]</sup> treatment. However, none of those analyses used ensemble methods or combined recursive partitioning with GLM to assess both synergistic and additive predictor effects.

In considering the possibility of future extensions to prospective studies, it is important to note that although we found an early-onset anxious-suicidal depression subtype that predicts all the outcomes (suicidality being the critical element in predicting disability and anxiety in predicting the other outcomes), we failed to find recursive partitioning profiles associated with a larger set of predictors despite the sample being much bigger than in existing prospective studies (i.e., affording good statistical power to detect synergistic symptom profiles if they existed) and the symptoms considered being quite broad. Taken together with the results of a recent secondary analysis that failed to find stable symptom-based MDD subtypes defined by internal consistency,<sup>[10]</sup> our failure to find more elaborate subtypes argues against the existence of complex MDD subtypes defined *exclusively* on the basis of synergistic associations among index depressive episode symptoms other than the early-onset anxious-suicidal subtype.



**TABLE 4. Symptoms associated with the high-risk, intermediate-risk, and low-risk clusters**

	High-risk		Intermediate-risk		Low-risk		Total		$\chi^2_2$	Cramer's V
	%	(SE)	%	(SE)	%	(SE)	%	(SE)		
I. Criterion A symptoms of major depression										
Dysphoria	99.6	(0.2)	98.9	(0.2)	96.1	(0.4)	98.2	(0.2)	57.0 <sup>a</sup>	.11
Severe dysphoria	97.8	(0.4)	91.0	(0.7)	56.0	(1.1)	81.3	(0.5)	713.6 <sup>a</sup>	.47
Anhedonia	95.0	(0.5)	88.7	(0.7)	79.0	(0.9)	87.3	(0.5)	205.1 <sup>a</sup>	.20
Weight loss	74.9	(1.0)	74.8	(1.0)	76.4	(1.0)	75.4	(0.6)	1.8	.02
Weight gain	18.0	(0.8)	14.8	(0.8)	12.0	(0.7)	14.8	(0.5)	28.5 <sup>a</sup>	.07
Insomnia	85.3	(0.9)	83.1	(0.8)	78.5	(0.9)	82.2	(0.5)	28.4 <sup>a</sup>	.07
Hypersomnia	10.5	(0.8)	9.0	(0.6)	9.7	(0.6)	9.7	(0.4)	2.4	.02
Psychomotor agitation	16.9	(0.8)	17.7	(0.9)	14.8	(0.8)	16.5	(0.5)	6.4 <sup>a</sup>	.03
Psychomotor retardation	67.7	(1.0)	55.0	(1.0)	41.0	(1.1)	54.2	(0.6)	247.9 <sup>a</sup>	.22
Fatigue/loss of energy	89.4	(0.8)	86.6	(0.8)	81.2	(0.8)	85.7	(0.5)	54.4 <sup>a</sup>	.10
Fatigue/loss or excessive guilt	98.2	(0.3)	88.4	(0.7)	68.3	(1.0)	84.6	(0.4)	724.0 <sup>a</sup>	.34
Diminished ability to concentrate/indisiveness	96.2	(0.4)	91.7	(0.6)	84.5	(0.7)	90.7	(0.4)	167.9 <sup>a</sup>	.16
Suicidality	98.8	(0.2)	76.9	(1.0)	38.2	(1.1)	70.6	(0.6)	1,449.6 <sup>a</sup>	.54
II. Symptoms of dysthymia										
Inability to cope	86.0	(0.8)	59.5	(1.0)	24.0	(1.0)	55.7	(0.7)	1,296.2 <sup>a</sup>	.50
Social withdrawal/tearfulness	99.2	(0.2)	94.4	(0.5)	88.0	(0.8)	93.7	(0.3)	261.3 <sup>a</sup>	.19
III. Symptoms of anxiety										
Irritability	77.7	(1.0)	62.9	(1.0)	41.1	(1.1)	60.1	(0.6)	541.3 <sup>a</sup>	.30
Nervousness-worry-anxiety	88.4	(0.9)	76.2	(1.0)	56.6	(1.1)	73.3	(0.6)	369.6 <sup>a</sup>	.29
Panic	66.4	(1.1)	38.1	(1.1)	13.5	(0.7)	38.5	(0.6)	842.1 <sup>a</sup>	.44
IV. Symptoms of mixed episode										
Little need for sleep	47.1	(1.2)	47.0	(1.2)	41.7	(1.1)	45.2	(0.7)	14.2 <sup>a</sup>	.05
Racing thoughts	10.6	(0.8)	12.4	(0.7)	15.7	(0.9)	12.9	(0.4)	16.6 <sup>a</sup>	.06
High energy	3.2	(0.4)	3.0	(0.4)	3.3	(0.4)	3.1	(0.2)	0.3	.01
V. Multivariate symptom profiles										
AOO <19 and suicidality	48.6	(1.2)	17.7	(0.9)	4.9	(0.5)	22.9	(0.6)	853.0 <sup>a</sup>	.43
AOO <19 and anxiety	48.9	(1.2)	24.4	(1.1)	9.9	(0.7)	27.1	(0.6)	658.4 <sup>a</sup>	.35
AOO <19 and suicidality and anxiety	47.8	(1.2)	14.9	(0.9)	2.1	(0.3)	20.7	(0.5)	919.8 <sup>a</sup>	.46
AOO ≥35 and suicidality and anxiety	16.4	(0.8)	21.9	(0.9)	9.1	(0.6)	15.9	(0.4)	179.2 <sup>a</sup>	.15
VI. Other predictors										
Endogenous	17.0	(0.9)	16.2	(0.9)	14.7	(0.8)	15.9	(0.6)	4.3	.03
Parental history of depression	9.9	(0.7)	5.8	(0.6)	3.2	(0.4)	6.2	(0.4)	63.8 <sup>a</sup>	.11
N	(2,520)		(2,899)		(2,842)		(8,261)			

<sup>a</sup>Significant at the 0.05 level, two-sided test.

It is important to note that broader MDD predictive subtypes not defined exclusively by index episode symptoms might be found in either of two other ways. One possibility would involve expanding the search for subtypes beyond symptoms of an index episode. Included here, for example, could be information about temporally primary comorbid mental disorders (e.g., early-onset distress, fear-circuitry, or impulse-control disorders), physical disorders (e.g., metabolic syndrome), sociodemographics, and (neuro)biological factors to define subtypes. We purposefully did not include such expansions here, as we wanted to focus on subtypes defined by index episode symptoms, but future analysis should do so to broaden the search for subtypes to include these other predictors. It would be interesting for future research to examine the possibility that the significant association found here between later-AOO anxious-suicidal depression in the index episode with later chronicity but no other outcome might reflect the importance of a late-onset depression subtype that might occur in conjunction with a physical comorbidity, such as cardiometabolic illness<sup>[55]</sup> associated with episodes of long duration but not high persistence or severity.

Such a possibility can only be examined by broadening the search for subtypes to include comorbid physical disorders. The potential value of expanding the search for subtypes to include information about biomarkers is illustrated in recent studies showing that the course of atypical and melancholic depression is differentially predicted by hypothalamic-pituitary-adrenal axis, metabolic syndrome, and inflammatory parameters<sup>[56]</sup> and that inflammatory dysregulation is associated with the onset of “mixed state depression.”<sup>[57]</sup> Such analyses have the potential to discover clinically meaningful and biologically valid disease clusters across a range of clinically relevant outcomes, an approach consistent with the recent call for what has been referred to as a *stratified medicine* approach<sup>[58]</sup> that bypasses the search for a gold standard and focuses instead on the discovery of subtypes associated with a range of clinically meaningful outcomes.

A second possibility would be to look more closely within the high-risk cluster found in our analysis to search for embedded subtypes. To understand this suggestion, it is important to recognize that the clusters we discovered cannot themselves be thought of as subtypes in the classical sense because they were discovered by clustering predicted outcome scores rather than the predictors themselves. A great many different combinations of predictors could yield the same predicted outcome scores. This means that further effort is needed to define subtypes within the high-risk cluster by considering multivariate profiles among the predictors that determine cluster membership so as to take into consideration the differential importance of these predictors within and across outcomes. No attempt was made to do this here, but it is clearly something that warrants future investigation in future studies based on the analysis of a more complete set of predictors.

It would also be useful, finally, if future studies expanded the range of outcomes considered here. The four outcomes in our analysis were selected purely based on availability. Given the discovery that predicted values are strongly correlated across these outcomes, it would be useful to develop an understanding of the range of outcomes over which this consistency occurs. Such an investigation could be carried out informally using the simple correlational methods used here, or a more formal approach might be conceived along the lines of the canonical regression models used to study latent mediators in the development of comorbidity among mental disorders.<sup>59–61</sup> Or it might be possible to address this issue by adapting the data mining methods developed to discover what have been called master regulators<sup>62</sup> in molecular genetic studies of physical disorders.<sup>63–65</sup> Regardless of method, though, the discovery of common predictors of multiple indicators of persistence, chronicity, and severity call out for a more diverse and integrated analysis of clusters and within-cluster subtypes among the predictors of such outcomes.

In thinking of these future developments, it is important to recognize that the recursive partitioning methods used here require a much larger sample size than is likely to exist in prospective clinical samples. This means that the most feasible way to extend the current results in prospective clinical studies would be to evaluate the significance of the synergistic symptom profiles found here rather than to attempt independent data mining exercises, although independent Lasso and cluster analyses using larger sets of predictors (possibly including measures of endophenotypes) and alternative indicators of outcomes would be quite feasible in such studies. Although it is unlikely that clinicians would be willing to collect such data for purposes of making the subtyping distinctions made here, it is conceivable that future studies will document powerful effects of other predictors that could be examined using similar methods and shown to have sufficiently important clinical implications that it would motivate clinicians to collect such information as a routine part of their initial evaluations to guide treatment planning. The technology described here holds great promise in facilitating analyses aimed at documenting such predictors.

## CONCLUSION

Despite our analysis being based on retrospective reports, our results suggest that useful symptom-based MDD subtyping distinctions can be made with data mining methods that focus on prediction rather than internal consistency and that the resulting subtypes have meaningful relationships with course of illness. The practical value of this approach, though, can only be judged by replication with prospective data, ideally expanding the analysis to use a wider range of predictors and outcomes.

**Acknowledgments.** The World Health Organization World Mental Health (WMH) Survey Initiative is supported by the National Institute of Mental Health (NIMH; R01 MH070884); John D. and Catherine T. MacArthur Foundation; Pfizer Foundation; U.S. Public Health Service (R13-MH066849, R01-MH069864, and R01 DA016558); Fogarty International Center (FIRCA R03-TW006481); Pan American Health Organization; Eli Lilly & Company Foundation; Ortho-McNeil Pharmaceutical, Inc.; GlaxoSmithKline; Sanofi Aventis; and Bristol-Myers Squibb. Peter de Jonge is supported by a VICI grant (91812607) from the Netherlands Research Foundation (NWO-ZonMW). We thank the WMH staff for assistance with instrumentation, fieldwork, and data analysis. A complete list of WMH publications can be found at <http://www.hcp.med.harvard.edu/wmh/>.

Each WMH country obtained funding for its own survey. The São Paulo Megacity Mental Health Survey is supported by the State of São Paulo Research Foundation (FAPESP) Thematic Project grant 03/00204–3. The Bulgarian Epidemiological Study of common mental disorders EPIBUL is supported by the Ministry of Health and the National Center for Public Health Protection. The Beijing, People’s Republic of China World Mental Health Survey Initiative is supported by the Pfizer Foundation. The Shenzhen, People’s Republic of China Mental Health Survey is supported by the Shenzhen Bureau of Health and the Shenzhen Bureau of Science, Technology, and Information. The Colombian National Study of Mental Health (NSMH) is supported by the Ministry of Social Protection. Implementation of the Iraq Mental Health Survey (IMHS) and data entry were carried out by the staff of the Iraqi MOH and MOP with direct support from the Iraqi IMHS team with funding from both the Japanese and European funds through United Nations Development Group Iraq Trust Fund (UNDG ITF). The Israel National Health Survey is funded by the Ministry of Health with support from the Israel National Institute for Health Policy and Health Services Research and the National Insurance Institute of Israel. The World Mental Health Japan (WMHJ) Survey is supported by the Grant for Research on Psychiatric and Neurological Diseases and Mental Health (H13-SHOGAI-023, H14-TOKUBETSU-026, H16-KOKORO-013) from the Japan Ministry of Health, Labour and Welfare. The Lebanese National Mental Health Survey (L.E.B.A.N.O.N.) is supported by the Lebanese Ministry of Public Health, the WHO (Lebanon), National Institute of Health/Fogarty International Center (R03 TW006481–01), Sheikh Hamdan Bin Rashid Al Maktoum Award for Medical Sciences, anonymous private donations to IDRAAC, Lebanon, and unrestricted grants from AstraZeneca, Eli Lilly, GlaxoSmithKline, Hikma Pharm, Pfizer, Roche, Sanofi-

Aventis, Servier and Novartis. The Mexican National Comorbidity Survey (MNCS) is supported by the National Institute of Psychiatry Ramon de la Fuente (INPRFMDIES 4280) and National Council on Science and Technology (CONACyT-G30544- H), with supplemental support from the PanAmerican Health Organization (PAHO). Te Rau Hinengaro, the New Zealand Mental Health Survey (NZMHS) is supported by the New Zealand Ministry of Health, Alcohol Advisory Council, and the Health Research Council. The Nigerian Survey of Mental Health and Wellbeing (NS-MHW) is supported by the WHO (Geneva), WHO (Nigeria), and Federal Ministry of Health, Abuja, Nigeria. The Northern Ireland Study of Mental Health was funded by the Health and Social Care Research and Development Division of the Public Health Agency. The Portuguese Mental Health Study was carried out by the Department of Mental Health, Faculty of Medical Sciences, NOVA University of Lisbon, with collaboration of the Portuguese Catholic University, and was funded by Champalimaud Foundation, Gulbenkian Foundation, Foundation for Science and Technology (FCT) and Ministry of Health. The Romania WMH study projects Policies in Mental Health Area and National Study regarding Mental Health and Services Use were carried out by National School of Public Health & Health Services Management (former National Institute for Research & Development in Health, present National School of Public Health Management & Professional Development, Bucharest), with technical support of Metro Media Transylvania, the National Institute of Statistics—National Centre for Training in Statistics, SC Cheyenne Services SRL, Statistics Netherlands, and were funded by Ministry of Public Health (former Ministry of Health) with supplemental support of Eli Lilly Romania SRL. The Ukraine Comorbid Mental Disorders during Periods of Social Disruption (CMDPSD) study is funded by the U.S. National Institute of Mental Health (R01-MH61905). The U.S. National Comorbidity Survey Replication (NCS-R) is supported by the National Institute of Mental Health (NIMH; U01-MH60220) with supplemental support from the National Institute of Drug Abuse (NIDA), Substance Abuse and Mental Health Services Administration (SAMHSA), Robert Wood Johnson Foundation (RWJF; grant 044708), and John W. Alden Trust. Additional support for preparation of this report was provided by Janssen Pharmaceuticals.

**Disclosure.** Dr. Wilcox is an employee of Janssen Pharmaceuticals. Dr. Kessler has been a consultant for AstraZeneca; Analysis Group; Bristol-Myers Squibb; Cerner-Galt Associates; Eli Lilly & Company;

GlaxoSmithKline Inc.; HealthCore Inc.; Health Dialog; Hoffman-LaRoche, Inc.; Integrated Benefits Institute; J & J Wellness & Prevention, Inc.; John Snow Inc.; Kaiser Permanente; Lake Nona Institute; Matria Inc.; Mensante; Merck & Co, Inc.; Ortho-McNeil Janssen Scientific Affairs; Pfizer Inc.; Primary Care Network; Research Triangle Institute; Sanofi-Aventis Groupe; Shire US Inc.; SRA International, Inc.; Takeda Global Research & Development; Transcept Pharmaceuticals Inc.; and Wyeth-Ayerst. Kessler has served on advisory boards for Appliance Computing II, Eli Lilly & Company, Mindsite, Ortho-McNeil Janssen Scientific Affairs, Johnson & Johnson, Plus One Health Management, and Wyeth-Ayerst. Kessler has had research support for his epidemiological studies from Analysis Group Inc., Bristol-Myers Squibb, Eli Lilly & Company, EPI-Q, GlaxoSmithKline, Johnson & Johnson Pharmaceuticals, Ortho-McNeil Janssen Scientific Affairs, Pfizer Inc., Sanofi-Aventis Groupe, Shire US, Inc., and Walgreens Co. Kessler owns 25% share in DataStat, Inc.

The views and opinions expressed in this report are those of the authors and should not be construed to represent the views or policies of any of the sponsoring organizations, agencies, or the World Health Organization.

## REFERENCES

- Cooper C, Jones L, Dunn E, et al. Clinical presentation of postnatal and non-postnatal depressive episodes. *Psychol Med* 2007;37:1273–1280.
- Cooper PJ, Murray L. Course and recurrence of postnatal depression. Evidence for the specificity of the diagnostic concept. *Br J Psychiatry* 1995;166:191–195.
- Fink M, Rush AJ, Knapp R, et al. DSM melancholic features are unreliable predictors of ECT response: a CORE publication. *J ECT* 2007;23:139–146.
- Uher R, Dernovsek MZ, Mors O, et al. Melancholic, atypical and anxious depression subtypes and outcome of treatment with escitalopram and nortriptyline. *J Affect Disord* 2011;132:112–120.
- Andreasen NC, Grove WM. The classification of depression: traditional versus mathematical approaches. *Am J Psychiatry* 1982;139:45–52.
- Romera I, Delgado-Cohen H, Perez T, et al. Factor analysis of the Zung self-rating depression scale in a large sample of patients with major depressive disorder in primary care. *BMC Psychiatry* 2008;8:4.
- Lamers F, Burstein M, He JP, et al. Structure of major depressive disorder in adolescents and adults in the US general population. *Br J Psychiatry* 2012;201:143–150.
- Baumeister H, Parker G. Meta-review of depressive subtyping models. *J Affect Disord* 2012;139:126–140.
- Carragher N, Adamson G, Bunting B, et al. Subtypes of depression in a nationally representative sample. *J Affect Disord* 2009;113:88–99.
- van Loo HM, de Jonge P, Romeijn JW, et al. Data-driven subtypes of major depressive disorder: a systematic review. *BMC Med* 2012;10:156.
- Glahn DC, Curran JE, Winkler AM, et al. High-dimensional endophenotype ranking in the search for major depression risk genes. *Biol Psychiatry* 2012;71:6–14.
- Hasler G, Northoff G. Discovering imaging endophenotypes for major depression. *Mol Psychiatry* 2011;16:604–619.
- Strobl C, Malley J, Tutz G. An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychol Methods* 2009;14:323–348.
- Zhang H, Singer BH. *Recursive Partitioning and Applications*. 2nd ed. New York, NY: Springer; 2010.
- Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. New York, NY: Springer; 2009.
- van der Laan MJ, Rose S. *Targeted Learning: Causal Inference for Observational and Experimental Data*. New York, NY: Springer; 2011.
- Chang YJ, Chen LJ, Chung KP, et al. Risk groups defined by Recursive Partitioning Analysis of patients with colorectal adenocarcinoma treated with colorectal resection. *BMC Med Res Methodol* 2012;12:2.
- Chao ST, Koefman SA, Woody N, et al. Recursive partitioning analysis index is predictive for overall survival in patients undergoing spine stereotactic body radiation therapy for spinal metastases. *Int J Radiat Oncol Biol Phys* 2012;82:1738–1743.
- Andreescu C, Mulsant BH, Houck PR, et al. Empirically derived decision trees for the treatment of late-life depression. *Am J Psychiatry* 2008;165:855–862.
- Jain FA, Hunter AM, Brooks JO, 3rd, et al. Predictive socioeconomic and clinical profiles of antidepressant response and remission. *Depress Anxiety* 2013;30:624–630.
- Nelson JC, Zhang Q, Deberdt W, et al. Predictors of remission with placebo using an integrated study database from patients with major depressive disorder. *Curr Med Res Opin* 2012;28:325–334.
- Rabinoff M, Kitchen CM, Cook IA, et al. Evaluation of quantitative EEG by classification and regression trees to characterize responders to antidepressant and placebo treatment. *Open Med Inform J* 2011;5:1–8.
- Riedel M, Moller HJ, Obermeier M, et al. Clinical predictors of response and remission in inpatients with depressive syndromes. *J Affect Disord* 2011;133:137–149.
- Ilgen MA, Downing K, Zivin K, et al. Exploratory data mining analysis identifying subgroups of patients with depression who are at high risk for suicide. *J Clin Psychiatry* 2009;70:1495–1500.
- Musil R, Zill P, Seemuller F, et al. Genetics of emergent suicidality during antidepressive treatment—data from a naturalistic study on a large sample of inpatients with a major depressive episode. *Eur Neuropsychopharmacol* 2013;23:663–674.
- Seemuller F, Riedel M, Obermeier M, et al. The controversial link between antidepressants and suicidality risks in adults: data from a naturalistic study on a large sample of in-patients with a major depressive episode. *Int J Neuropsychopharmacol* 2009;12:181–189.
- Kessler RC, Üstün TB, editors. *The WHO World Mental Health Surveys: Global Perspectives on the Epidemiology of Mental Disorders*. New York, NY: Cambridge University Press; 2008.
- Alonso J, Chatterji S, He Y. *The Burden of Mental Disorders: Global Perspectives from the WHO World Mental Health Surveys*. New York, NY: Cambridge University Press; 2013.
- Nock MK, Borges G, Ono Y. *Suicide: Global Perspectives from the WHO World Mental Health Surveys*. New York, NY: Cambridge University Press; 2012.
- von Korff MR, Scott KM, Gureje O. *Global Perspectives on Mental-Physical Comorbidity in the WHO World Mental Health Surveys*. New York, NY: Cambridge University Press; 2009.



31. World Bank. Data, Country and Lending Groups. 2009 [cited 2009 May 12]. Available at: <http://go.worldbank.org/D7SN0B8YU0>.
32. Heeringa SG, Wells EJ, Hubbard F, et al. Sample designs and sampling procedures. In: Kessler RC, Üstün TB, editors. *The WHO World Mental Health Surveys: Global Perspectives on the Epidemiology of Mental Disorders*. New York, NY: Cambridge University Press; 2008:14–32.
33. Harkness J, Pennell BE, Villar A, et al. Translation procedures and translation assessment in the world mental health survey initiative. In: Kessler RC, Üstün TB, editors. *The WHO World Mental Health Surveys: Global Perspectives on the Epidemiology of Mental Disorders*. New York, NY: Cambridge University Press; 2008:91–113.
34. Pennell B-E, Mneimneh Z, Bowers A, et al. Implementation of the world mental health surveys. In: Kessler RC, Üstün TB, editors. *The WHO World Mental Health Surveys: Global Perspectives on the Epidemiology of Mental Disorders*. New York, NY: Cambridge University Press; 2008:33–57.
35. Kessler RC, Üstün TB. The World Mental Health (WMH) survey initiative version of the World Health Organization (WHO) Composite International Diagnostic Interview (CIDI). *Int J Methods Psychiatr Res* 2004;13:93–121.
36. Haro JM, Arbabzadeh-Bouchez S, Brugha TS, et al. Concordance of the Composite International Diagnostic Interview Version 3.0 (CIDI 3.0) with standardized clinical assessments in the WHO World Mental Health surveys. *Int J Methods Psychiatr Res* 2006;15:167–180.
37. Altman D, Machin D, Bryant T, et al., editors. *Statistics with Confidence: Confidence intervals and statistical guidelines*. 2nd ed. London: BMJ Books; 2000.
38. Endicott J, Andreasen N, Spitzer RL. *Family History Research Diagnostic Criteria*. New York, NY: Biometrics Research, New York State Psychiatric Institute; 1978.
39. Thernau T, Atkinson B, Ripley B. Rpart: recursive partitioning. R Package 4.1–0. 2012.
40. Draper N, Smith H. *Applied Regression Analysis*. 2nd ed. New York, NY: Wiley; 1981.
41. Berk RA. *Regression Analysis: A Constructive Critique*. Newbury Park, CA: Sage; 2003.
42. Tibshirani R. Regression shrinkage and selection via the lasso. *J Roy Stat Soc B* 1996;58:267–288.
43. Berk RA. *Statistical Learning from a Regression Perspective*. New York, NY: Springer; 2008.
44. Friedman J, Hastie T, Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 2010;33:1–22.
45. R Core Team. R: a language and environment for statistical computing. 2013 [cited 2013 November 8]. Available at: <http://www.R-project.org/>.
46. Chambless LE, Diao G. Estimation of time-dependent area under the ROC curve for long-term risk prediction. *Stat Med* 2006;25:3474–3486.
47. Judd LL, Schettler PJ, Coryell W, et al. Overt irritability/anger in unipolar major depressive episodes: past and current characteristics and implications for long-term course. *JAMA Psychiatry* 2013;70:1171–1180.
48. Penninx BW, Beekman AT, Smit JH, et al. The Netherlands Study of Depression and Anxiety (NESDA): rationale, objectives and methods. *Int J Methods Psychiatr Res* 2008;17:121–140.
49. Bottomley C, Nazareth I, Torres-Gonzalez F, et al. Comparison of risk factors for the onset and maintenance of depression. *Br J Psychiatry* 2010;196:13–17.
50. Patten SB, Wang JL, Williams JV, et al. Predictors of the longitudinal course of major depression in a Canadian population sample. *Can J Psychiatry* 2010;55:669–676.
51. Spijker J, de Graaf R, Bijl RV, et al. Determinants of persistence of major depressive episodes in the general population. Results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). *J Affect Disord* 2004;81:231–240.
52. Coryell W, Fiedorowicz JG, Solomon D, et al. Effects of anxiety on the long-term course of depressive disorders. *Br J Psychiatry* 2012;200:210–215.
53. Gerrits MM, Vogelzangs N, van Oppen P, et al. Impact of pain on the course of depressive and anxiety disorders. *Pain* 2012;153:429–436.
54. Wardenaar KJ, Giltay EJ, van Veen T, et al. Symptom dimensions as predictors of the two-year course of depressive and anxiety disorders. *J Affect Disord* 2012;136:1198–1203.
55. Vogelzangs N, Beekman AT, Boelhouwer IG, et al. Metabolic depression: a chronic depressive subtype? Findings from the InCHIANTI study of older persons. *J Clin Psychiatry* 2011;72:598–604.
56. Lamers F, Vogelzangs N, Merikangas KR, et al. Evidence for a differential role of HPA-axis function, inflammation and metabolic syndrome in melancholic versus atypical depression. *Mol Psychiatry* 2013;18:692–699.
57. Becking K, Boschloo L, Vogelzangs N, et al. The association between immune activation and manic symptoms in patients with a depressive disorder. *Transl Psychiatry* 2013;3:e314.
58. Kapur S, Phillips AG, Insel TR. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol Psychiatry* 2012;17:1174–1179.
59. Kessler RC, Cox BJ, Green JG, et al. The effects of latent variables in the development of comorbidity among common mental disorders. *Depress Anxiety* 2011;28:29–39.
60. Kessler RC, Ormel J, Petukhova M, et al. Development of lifetime comorbidity in the World Health Organization World Mental Health Surveys. *Arch Gen Psychiatry* 2011;68:90–100.
61. Kessler RC, Petukhova M, Zaslavsky AM. The role of latent internalizing and externalizing predispositions in accounting for the development of comorbidity among common mental disorders. *Curr Opin Psychiatry* 2011;24:307–312.
62. Chan SS, Kyba M. What is a master regulator? *J Stem Cell Res Ther* 2013;3:e114.
63. Rabinowitz JD, Silhavy TJ. Systems biology: metabolite turns master regulator. *Nature* 2013;500:283–284.
64. Rayner BS, Figtree GA, Sabaretnam T, et al. Selective inhibition of the master regulator transcription factor egr-1 with catalytic oligonucleotides reduces myocardial injury and improves left ventricular systolic function in a preclinical model of myocardial infarction. *J Am Heart Assoc* 2013;2:e000023.
65. Ryu B, Kim DS, Deluca AM, et al. Comprehensive expression profiling of tumor cell lines identifies molecular signatures of melanoma progression. *PloS One* 2007;2:e594.