

Original Paper

Toward an Extended Definition of Major Depressive Disorder Symptomatology: Digital Assessment and Cross-validation Study

Nayra A Martin-Key¹, PhD; Dan-Mircea Mirea^{1,2}, MSc; Tony Olmert^{1,3}, MPhil; Jason Cooper^{1,4}, PhD; Sung Yeon Sarah Han¹, PhD; Giles Barton-Owen⁵, MA; Lynn Farrag⁵, BSc; Emily Bell⁵, BA; Pawel Eljasz¹; Daniel Cowell^{5,6}, BSc; Jakub Tomasik¹, PhD; Sabine Bahn^{1,5}, MRCPsych, MD, PhD

¹Cambridge Centre for Neuropsychiatric Research, Department of Chemical Engineering and Biotechnology, University of Cambridge, Cambridge, United Kingdom

²Princeton Neuroscience Institute, Princeton University, Princeton, NJ, United States

³UC San Diego School of Medicine, University of California, San Diego, CA, United States

⁴Owlstone Medical Ltd, Cambridge, United Kingdom

⁵Psyomics Ltd, Cambridge, United Kingdom

⁶Sentinel Oncology Ltd, Cambridge, United Kingdom

Corresponding Author:

Sabine Bahn, MRCPsych, MD, PhD

Cambridge Centre for Neuropsychiatric Research

Department of Chemical Engineering and Biotechnology

University of Cambridge

Philippa Fawcett Drive

Cambridge, CB3 0AS

United Kingdom

Phone: 44 1223 334151

Email: sb209@cam.ac.uk

Abstract

Background: Diagnosing major depressive disorder (MDD) is challenging, with diagnostic manuals failing to capture the wide range of clinical symptoms that are endorsed by individuals with this condition.

Objective: This study aims to provide evidence for an extended definition of MDD symptomatology.

Methods: Symptom data were collected via a digital assessment developed for a delta study. Random forest classification with nested cross-validation was used to distinguish between individuals with MDD and those with subthreshold symptomatology of the disorder using disorder-specific symptoms and transdiagnostic symptoms. The diagnostic performance of the Patient Health Questionnaire–9 was also examined.

Results: A depression-specific model demonstrated good predictive performance when distinguishing between individuals with MDD (n=64) and those with subthreshold depression (n=140) (area under the receiver operating characteristic curve=0.89; sensitivity=82.4%; specificity=81.3%; accuracy=81.6%). The inclusion of transdiagnostic symptoms of psychopathology, including symptoms of depression, generalized anxiety disorder, insomnia, emotional instability, and panic disorder, significantly improved the model performance (area under the receiver operating characteristic curve=0.95; sensitivity=86.5%; specificity=90.8%; accuracy=89.5%). The Patient Health Questionnaire–9 was excellent at identifying MDD but overdiagnosed the condition (sensitivity=92.2%; specificity=54.3%; accuracy=66.2%).

Conclusions: Our findings are in line with the notion that current diagnostic practices may present an overly narrow conception of mental health. Furthermore, our study provides proof-of-concept support for the clinical utility of a digital assessment to inform clinical decision-making in the evaluation of MDD.

(*JMIR Form Res* 2021;5(10):e27908) doi: [10.2196/27908](https://doi.org/10.2196/27908)

KEYWORDS

major depressive disorder; subthreshold depression, transdiagnostic symptoms; digital assessment; digital mental health; mobile phone

Introduction

Background

Major depressive disorder (MDD) is a common and heterogeneous condition representing the leading cause of disability worldwide [1]. MDD has been associated with poor global outcomes, including impaired social functioning, lower quality of life, inability to return to work, and suicide [2]. The condition is typically diagnosed in primary care settings, with most help seekers exhibiting subthreshold or subsyndromal presentations of the disorder [3,4]. Critically, recognizing diagnosable symptomatology of MDD can be particularly challenging, with any 2 individuals' meeting criteria for the condition potentially having no symptoms in common [5]. In fact, short consultation times coupled with the complexity and subjectivity of diagnosing MDD results in primary care practitioners misdiagnosing >50% of low-mood help seekers [6]. This means that many patients do not receive the most effective treatment and support.

In an attempt to improve the current diagnostic practice, the search for objective diagnostic tests and valid biomarkers for depression has received a lot of attention. However, despite substantial research expenditures and large-scale genome-wide studies, no pathognomonic biological markers of depression have been identified [7-11]. In fact, with the exception of a few neuropsychiatric disorders, not a single psychiatric diagnosis can still be validated by molecular, genetic, or imaging biomarkers [12]. Importantly, psychiatric diagnostic criteria were not conceptualized to facilitate biological differentiation [13], with extensive comorbidity across conditions being the rule rather than the exception [14] and no single condition representing a discrete entity [15-17].

Another issue pertaining to psychiatric nosology is *incomplete symptom capture* [18,19]. It has been argued that the symptom profiles described in diagnostic manuals, such as the *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (DSM-5)* [20] and the *International Statistical Classification of Diseases, Eleventh Revision (ICD-11)* [21], may be overly narrow, failing to capture the wide range of clinical symptoms that are endorsed by individuals with MDD [18]. For instance, although anxiety is not listed as a core symptom of the condition, many individuals with MDD experience co-occurring symptoms of anxiety and typically meet the criteria for at least one anxiety disorder [22-26].

Some authors suggest that assessing the presence of anxiety symptoms in patients with MDD is critical [27,28]. Others propose combining depression and anxiety disorders, which present with largely overlapping symptomatology. Combining these disorders may be a useful strategy for better clinical evaluation and management of patients with MDD [29]. This is important, given that estimates of prevalence rates for treatment-resistant depression range between 30% and 50% [30,31], with incomplete remission often leading to relapse [32], increased chronicity and severity of episodes [33], greater functional impairments [34], and higher risk of suicide [35]. One of the established risk factors that predispose patients to develop treatment-resistant depression are comorbid anxiety

symptoms or anxiety disorders, especially generalized anxiety disorder (GAD) [31].

In this regard, a transdiagnostic view of MDD encompassing symptoms of anxiety and other commonly co-occurring disorders may improve early and accurate diagnosis, reflect biological disease understanding (eg, twin studies have shown shared genetic predisposition for MDD and GAD [36]), and allow for personalized treatment strategies. An extended definition of MDD symptomatology may also reduce the misdiagnosis of bipolar disorder (BD) as MDD [37,38], which is particularly problematic, with many individuals having to wait 8-10 years before receiving a correct diagnosis [39,40].

Critically, time is premium in primary care settings, where relying on brief symptom-count checklists, such as the Patient Health Questionnaire-9 (PHQ-9) [41], is a common practice. Importantly, the PHQ-9 may overestimate depression severity in primary care patients relative to other self- and clinician-rated scales [42], resulting in a greater reliance on medication as a first-line treatment option and an increased potential for adverse drug effects [43,44]. In addition, some researchers have suggested that the PHQ-9 may be missing the presence of symptoms that are meaningful for patients and that longer assessments may be better at capturing diagnosable levels of low mood [45]. To this end, digital technologies allow for the cost- and time-effective collection of a vast range of important patient and symptom data [46]. Such an approach offers an innovative way to improve and advance mental health care provision. In turn, the use of digital technologies could help alleviate the load on the health care system by providing individuals with subthreshold or mild MDD with self-help tips and psychoeducation. This would reserve the limited and specialized services for more severe or highly comorbid and complex patients.

Objectives of This Study

This study aims to provide evidence for an extended definition of MDD symptomatology using a digital assessment that was developed for the delta study [47]. The digital assessment was designed following an extensive analysis of existing validated questionnaires for mood disorders [48-57], the *DSM-5* [20], and the *International Statistical Classification of Diseases, Tenth Revision (ICD-10)* [21], as well as input from psychiatrists and a service user group. In an attempt to move away from a symptom-count approach to psychopathology, our study uses machine learning (ML) methods (advanced statistical and probabilistic techniques that automatically learn from data) to construct a data-driven view of MDD. We examined the extent to which (1) disorder-specific symptoms (ie, symptoms of depression) and (2) transdiagnostic symptoms (ie, cross-disorder symptoms) could be used to answer the following question: "when does depression become a mental disorder?" To do this, we compared individuals with MDD with those with subthreshold levels of depressive symptoms. Although this was largely an exploratory study, it was predicted that, relative to a disorder-specific model of psychopathology, an extended model would be better at identifying individuals with diagnosable levels of MDD. In particular, we predicted that symptoms of anxiety would be highly indicative of the disorder.

Methods

The Delta Study

This study used data from the delta study that was conducted by the Cambridge Centre for Neuropsychiatric Research between April 2018 and November 2019. Olmert et al [47] provided a detailed description of the delta study design and sampling procedures. In brief, the key objectives of the study were to develop and validate a diagnostic algorithm to (1) reduce the misdiagnosis of MDD and BD and (2) achieve a more accurate and earlier diagnosis of MDD in individuals presenting with depressive symptoms. The target population for the primary objective was those who had received a recent diagnosis of MDD (within the past 5 years) by a general practitioner or psychiatrist and those who were experiencing depressive symptoms at the time of recruitment. The target population for the secondary objective included those without a previous mood disorder diagnosis and who were experiencing depressive symptoms at the time of recruitment. Individuals aged between 18 and 45 years could take part in this study. This age group was selected in consultation with a practicing psychiatrist (SB) on the basis that individuals aged between 18 and 45 years are most likely to have undiagnosed BD (primary objective of the delta study). Further inclusion criteria were being a resident in the United Kingdom, not pregnant or breastfeeding, not suicidal, and a score of at least five on the PHQ-9 [41]. Information on treatment history was collected but was not deemed an inclusion or exclusion criterion. All participants provided informed consent to participate in the study, which was approved by the University of Cambridge Human Biology Research Ethics Committee (approval number HBREC 2017.11).

Over 5000 participants were recruited on the web through email, via paid Facebook (Facebook Inc) advertisements, and updates on the Cambridge Centre for Neuropsychiatric Research laboratory website. Eligible participants were invited to take part in the main study; of these, 3232 completed the digital assessment via the delta study website. The digital assessment was designed following an extensive analysis of validated questionnaires for mood disorders [48-57], the *DSM-5* [20], and the *ICD-10* [21], as well as input from psychiatrists to ensure the inclusion of a wide range of clinically meaningful and well-validated symptoms of MDD and BD and other symptoms of interest (eg, other psychiatric conditions). The assessment was further refined following advice from a service user group on features, including tone of voice and user journey. For each participant, an individualized dashboard guided their progress through the study. The digital assessment could be completed

on a smartphone, laptop, tablet, or desktop computer and comprised 635 distinct questions. One question was presented at a time, with participants required to select the answer that best described their feelings and experiences (eg, from “No, not at all” to “Yes, very much”). The questions were grouped into six sections: (1) demographics and personal history; (2) bipolar and manic and hypomanic symptoms; (3) depressive symptoms; (4) personality traits; (5) medication, treatment, and substance use; and (6) other psychiatric disorders and symptoms, including GAD, social anxiety disorder, emotional instability, panic disorder, eating disorders, insomnia, and obsessive-compulsive disorder. Participants were required to press *next* to take them to the following question. Each section could be completed in 10-15 minutes, although response times varied because of the adaptive nature of the assessment, where only relevant questions were asked based on responses to previous questions.

A subgroup of the original study cohort (n=1740) consented to provide dried blood spot samples and complete a telephone interview for MDD and BD using the *Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition (DSM-IV)*-based Composite International Diagnostic Interview (CIDI [48]), with 924 participants completing both steps. Of these 924 participants, 241 (26.1%) self-reported having been diagnosed with MDD by a general practitioner or psychiatrist and were confirmed to have met the criteria for the condition by the CIDI. Of these 241 participants, 64 (26.6%) participants (11/64, 17% participants were male) met the criteria for *current* MDD (ie, symptoms present in the past month according to the CIDI). This formed the MDD group. The subthreshold depression group included 15.2% (140/924) participants (male: 35/140, 25%) who self-reported no diagnosis of MDD and whose symptoms of depression were confirmed to not meet the criteria for MDD according to the CIDI. None of the selected participants had ever experienced a manic or hypomanic episode.

Data Analytic Strategy

Participant Characteristics

Participant characteristics and comorbidities were collected via digital assessments and are shown in Table 1. Group differences in continuous variables were explored using Mann-Whitney U tests as the data were nonnormally distributed, with effect sizes reported as *r* (small ≥ 0.1 ; medium ≥ 0.3 ; and large ≥ 0.5) [58]. Group differences in categorical variables were evaluated using chi-square tests or Fisher exact test for low-frequency data (ie, values < 5). Effect sizes are reported as Cramer V (ϕ_c ; small ≥ 0.1 ; medium ≥ 0.3 ; and large ≥ 0.5) [58].

Table 1. Participant characteristics and comorbidities: major depressive disorder versus subthreshold depression group comparisons.

Characteristics	Subthreshold depression (n=140)	MDD ^a (n=64)	<i>U</i> ^b	<i>P</i> value	<i>r</i> ^c	Chi-square (<i>df</i>)	ϕ_c ^d
Age (years), mean (SD)	25.84 (6.66)	25.94 (5.7)	4209	.49	0.05	N/A ^e	N/A
BMI, mean (SD)	24.62 (4.73)	28.29 (6.8)	2901	<.001	0.28	N/A	N/A
Sex, n (%)							
Male	49 (35)	11 (17.2)	N/A	N/A	N/A	N/A	N/A
Female	91 (65)	53 (82.8)	N/A	.01	N/A	6.7 (1)	0.18
Higher education^f, n (%)							
Yes	87 (62.1)	37 (42.2)	N/A	N/A	N/A	N/A	N/A
No	53 (37.1)	27 (57.8)	N/A	.56	N/A	0.4 (1)	0.04
Employment, n (%)							
Employed	81 (57.9)	35 (54)	N/A	N/A	N/A	N/A	N/A
Unemployed	5 (3.6)	8 (12.5)	N/A	N/A	N/A	N/A	N/A
Student	54 (38.6)	21 (32.8)	N/A	.05	N/A	6.0 (2)	0.17
Support network: relationships, n (%)							
Secure and stable relationship	84 (59.6)	133 (55.4)	N/A	N/A	N/A	N/A	N/A
Insecure and unstable relationship	9 (6.4)	14 (5.8)	N/A	N/A	N/A	N/A	N/A
Single	48 (34)	93 (38.8)	N/A	.54	N/A	1.3 (2)	0.08
Support network: living alone, n (%)							
Yes	9 (6.4)	9 (14.1)	N/A	N/A	N/A	N/A	N/A
No	131 (93.6)	55 (85.9)	N/A	.07	N/A	3.2 (1)	0.13
Psychiatric history, n (%)							
GAD ^g	12 (8.6)	47 (73.4)	N/A	<.001	N/A	89.9	0.66
Personality disorder	0	3 (4.7)	N/A	N/A	N/A	N/A	N/A
OCD ^h	3 (2.1)	3 (4.7)	N/A	.28	N/A	0.4 (2)	0.07
Panic disorder	0	7 (10.9)	N/A	N/A	N/A	N/A	N/A
Social anxiety	1 (0.7)	9 (14.1)	N/A	<.001	N/A	16.8 (2)	0.29
Eating disorder	2 (1.4)	5 (7.8)	N/A	.03	N/A	5.4 (2)	0.16
Medical history, n (%)							
Thyroid disease	4 (2.9)	2 (3.1)	N/A	.99	N/A	0.0 (3)	0.01
Cardiovascular disease	1 (0.7)	0	N/A	N/A	N/A	N/A	N/A
Irritable bowel syndrome	6 (4.3)	2 (3.1)	N/A	.99	N/A	0.2 (3)	0.03
Chronic pain	30 (21.4)	14 (21.9)	N/A	.94	N/A	0.0 (3)	0.01
Migraines	53 (37.9)	31 (48.4)	N/A	.15	N/A	2.0 (3)	0.10
Current psychiatric treatment, n (%)							
SSRI ⁱ antidepressants	12 (8.6)	32 (50)	N/A	<.001	N/A	44.6 (3)	0.47
SNRI ^j antidepressants	0	5 (7.8)	N/A	N/A	N/A	N/A	N/A
Tricyclic antidepressants	0	3 (4.7)	N/A	N/A	N/A	N/A	N/A
Other antidepressants	1 (0.7)	8 (12.5)	N/A	<.001	N/A	14.5 (3)	0.27
Anxiety medication	4 (2.9)	10 (15.6)	N/A	.002	N/A	11.2 (3)	0.23
Antipsychotics	0	3 (4.7)	N/A	N/A	N/A	N/A	N/A
Mood stabilizers	0	2 (3.1)	N/A	N/A	N/A	N/A	N/A

Characteristics	Subthreshold depression (n=140)	MDD ^a (n=64)	U^b	P value	r^c	Chi-square (df)	ϕ_c^d
Psychotherapy	4 (2.9)	16 (25)	N/A	<.001	N/A	24.4 (3)	0.35

^aMDD: major depressive disorder.

^bMann–Whitney U test.

^cEffect size (r).

^dEffect size (Cramer V).

^eN/A: not applicable.

^fUndergraduate degree or equivalent and above was coded as *yes*, whereas A level or equivalent and below was coded as *no*.

^gGAD: generalized anxiety disorder.

^hOCD: obsessive-compulsive disorder.

ⁱSNRI: serotonin–norepinephrine reuptake inhibitor.

^jSSRI: selective serotonin reuptake inhibitor.

Model Construction and Performance

Random forest classification models were constructed in Python 3.7.4 (Python Software Foundation) using the scikit-learn library 0.21.3 to distinguish between MDD and subthreshold depression using (1) disorder-specific symptoms (ie, symptoms of depression), and (2) transdiagnostic symptoms (ie, cross-disorder symptoms). We constructed two models: a *depression model*, including 36 symptoms of depression and an *extended model*, comprising 134 symptoms (36 symptoms of depression, 12 symptoms of GAD, 19 symptoms of BD or mania, 15 symptoms of hypomania, 6 symptoms of social anxiety, 11 symptoms of emotional instability, 14 symptoms of panic disorder, nine symptoms of obsessive-compulsive disorder, two symptoms of eating disorders, and 10 symptoms of insomnia). These symptoms were manually selected based on the maximum number of available symptoms from the digital assessment. Scores per symptom ranged from 0–1, with higher scores indicating increased severity. Mean symptom severity per group can be found in Table S1, [Multimedia Appendix 1](#).

Although some symptoms overlapped across disorders (eg, tiredness, low energy, and irritability), we did not feel that it would be appropriate to combine these as the questions were framed in the context of each condition. Furthermore, although all participants were asked about symptoms of depression, BD or mania, and hypomania, the questions for the remaining conditions were adaptive in nature, such that only relevant questions were asked based on responses to previous questions. This resulted in some participants having *missing* data. These data were imputed as zeros. Furthermore, owing to the adaptive nature of the digital mental health assessment, participants answered questions on current (eg, present in the past 2 weeks) or past symptoms of all disorders.

For each of the models, nested cross-validation (NCV) was performed to obtain the highest algorithmic accuracy while ensuring the generalizability of the models. At each iteration of NCV, the data were randomly split into three folds; two-thirds of the data were used in the inner loop for model training and validation, and one-third was used for testing the model in the outer loop. The inner loop was (further) randomly split into three folds, whereby the hyperparameters (ie, number of estimators and maximum depth) were tuned, and the best cross-validated model was selected. To do this, two of the three

folds were used to tune the model parameters and train the model, which was then validated on the third fold. This procedure was repeated with the remaining combinations of training and validation folds. The final model (ie, the optimized classifier) was obtained by fitting a model with the tuned parameters to all three data folds from the inner loop and then evaluating the hold-out test data in the outer loop. This procedure was repeated 100 times with different splits of the data (into train and test sets), resulting in a total of 300 unique models for each feature set (ie, depression model vs extended model).

Model performance was evaluated by measuring the area under the receiver operating characteristic curve (AUC) for the 300 models and averaging across all models for each feature set. The AUC shows the degree of separability between two conditions (ie, MDD vs subthreshold depression) and represents the probability that a randomly selected subject with the condition is rated or ranked as more likely to have the condition than a randomly selected individual without the condition (AUC: ≥ 0.9 =excellent; ≥ 0.8 =good; ≥ 0.7 =fair; ≥ 0.6 =poor; ≥ 0.5 =fail) [59]. Mann–Whitney U tests were used to determine significant differences in AUCs across the 300 models between the depression and transdiagnostic models.

The mean sensitivity, specificity, and accuracy scores per model were also evaluated. Here, sensitivity refers to the model's ability to classify MDD cases correctly (ie, true positives), whereas specificity refers to the model's ability to classify subthreshold depression cases correctly (ie, true negatives). Accuracy corresponds to the model's ability to classify all true cases (ie, both true positives and true negatives).

Feature Importance and Occurrence

Relative feature importances (ie, Gini impurity [60]) were calculated for the 300 models and averaged across all models for each feature set, with features with higher values showing better discrimination between MDD and subthreshold depression. Feature occurrence was calculated by summing the number of times each feature contributed to each of the 300 models and computing a percentage score per feature for each feature set, with higher values representing higher feature occurrence.

Diagnostic Performance of the PHQ-9

Finally, to establish the diagnostic performance of the PHQ-9 on the basis of its intended use, we calculated the sensitivity, specificity, and accuracy in the current sample using the standard cut-off score of ≥ 10 [41]. This mimics what would ordinarily happen in the clinic (ie, those scoring ≥ 10 would be classified as MDD).

Results

Participant Characteristics

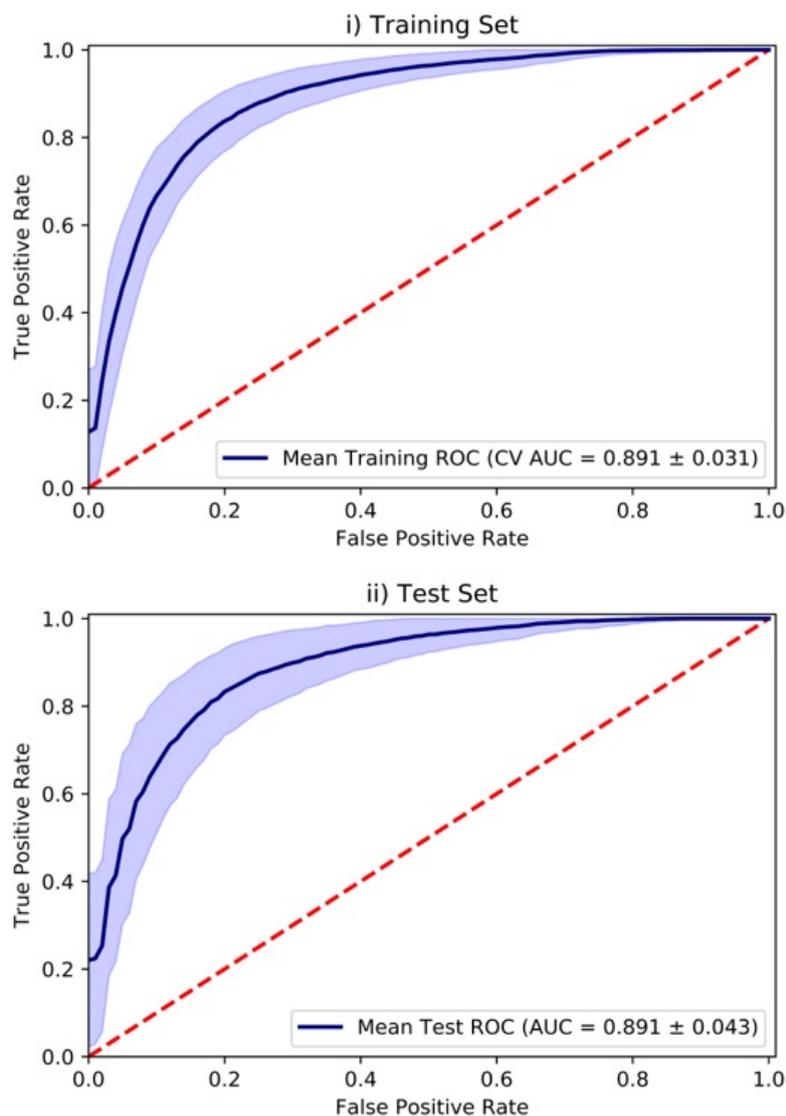
Table 1 presents information on the characteristics of each group, with statistical comparisons. The groups did not differ significantly in age, level of education, support network (ie, relationships and living conditions), or medical history. However, the MDD group had a significantly higher proportion of women and a higher mean BMI than the subthreshold depression group. The MDD group was significantly more likely

to be unemployed than the subthreshold depression group, which was more likely to be employed or in full-time education. The groups had different psychiatric histories, with the MDD group having a significantly higher proportion of individuals with comorbid GAD, social anxiety disorder, and eating disorders. Finally, relative to the subthreshold depression group, individuals with MDD were more likely to be currently taking psychiatric medication and receiving psychotherapy.

Depression Model

This model comprised 36 features, with analyses demonstrating good discriminatory performance on both the training (AUC=0.89±0.03) and test sets (AUC=0.89±0.04; Figure 1). Approximately 81% (52/64) of the MDD cases and 81.4% (114/140) of the subthreshold depression cases were correctly classified by the model, corresponding to the mean sensitivity and specificity scores, respectively. The mean accuracy of the model was 81.6%, corresponding to the proportion of individuals correctly classified by the model.

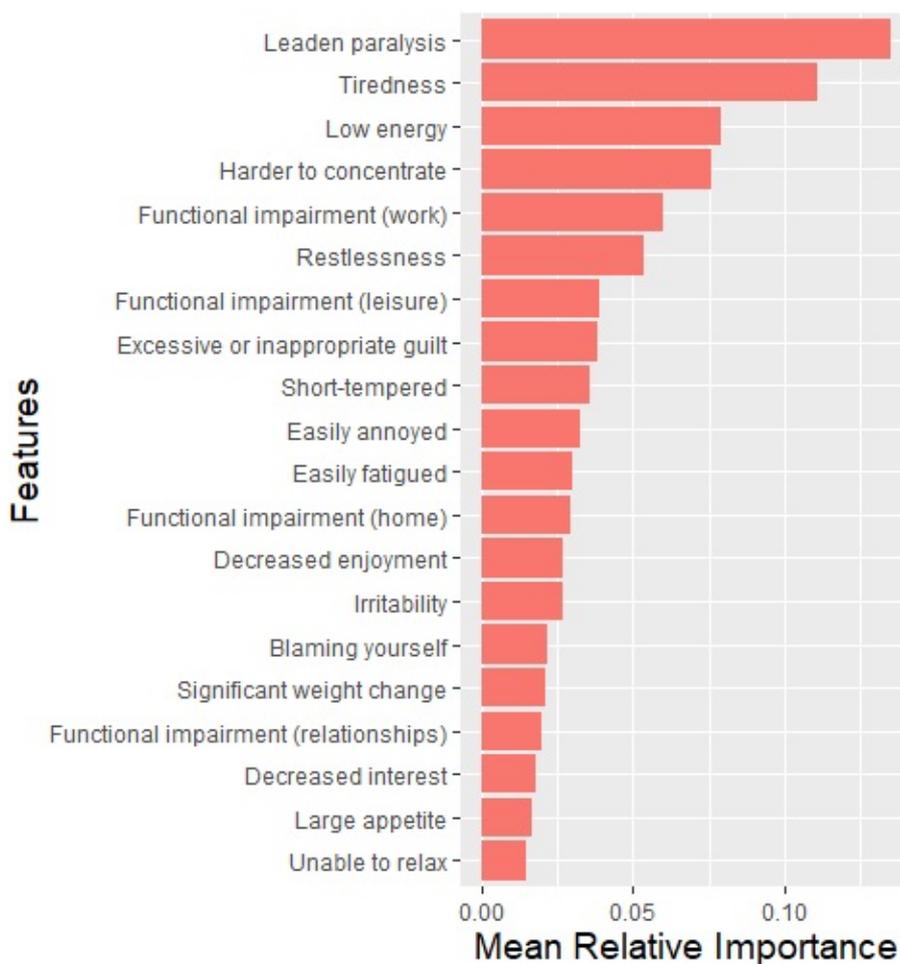
Figure 1. Area under the receiver operating characteristic curves showing mean predictive performance of the depression model. The models were applied to predict the probability of major depressive disorder in the: (1) training and (2) test sets. AUC: area under the receiver operating characteristic curve; CV AUC: cross-validated area under the receiver operating characteristic curve; MDD: major depressive disorder; ROC: receiver operating characteristic.



The top 20 features contributing to the depression model (averaged across all 300 models) were leaden paralysis, tiredness, low energy, harder to concentrate, functional impairment (work), restlessness, functional impairment (leisure), excessive or inappropriate guilt, short-tempered, easily annoyed, easily fatigued, functional impairment (home), decreased enjoyment, irritability, blaming oneself, significant weight

change, functional impairment (relationships), decreased interest, large appetite, and unable to relax. Figure 2 shows the mean relative feature importances. These features appeared across at least 81.3% (244/300) of the models. Multimedia Appendices 2 and 3 show the relative feature importances and percentage occurrences of all 36 features.

Figure 2. Top 20 mean relative importance for the depression-specific model. Features have been ordered from most to least important.



Extended Model

Next, we added 98 features to the model, resulting in an extended model comprising 134 features. Analyses demonstrated excellent discriminatory performance on both the training (AUC=0.94±0.03) and test sets (AUC=0.94±0.04; Multimedia Appendix 4). Mann–Whitney U tests confirmed a significant improvement in model performance (ie, AUC) relative to the depression-specific model (training set: $U=12922.50$, $P<.001$; test set: $U=33525.50$, $P<.001$). Here, 83% (53/64) of MDD cases (ie, sensitivity) and 90% (126/140) of subthreshold depression cases (ie, specificity) were correctly classified by the model, whereas the ability of the model to correctly classify both MDD and subthreshold depression cases was 87.7% (ie, accuracy). Feature importances can be found in Multimedia Appendix 5, with percentage occurrences found in Multimedia Appendix 6.

On the basis of these findings, we then reran the analyses using a *truncated* version of the extended model, which only included features that appeared across at least 90.3% (271/300) of the models (Multimedia Appendix 6). The truncated model comprised 12 symptoms of depression, 11 symptoms of GAD, six symptoms of insomnia, three symptoms indicative of emotional instability, and one panic disorder symptom, resulting in a total of 33 features.

The analyses revealed a significant improvement in the model's discriminatory performance on both the training (AUC=0.95±0.02) and test sets (AUC=0.95±0.03; Figure 3) relative to the full extended model (training set: $U=12856$, $P<.001$; test set: $U=322798.50$, $P<.001$). The mean sensitivity, specificity, and accuracy scores were 86.5%, 90.8%, and 89.5%, respectively. The mean relative feature importance for the 33 features can be found in Figure 4, with percentage occurrences found in Multimedia Appendix 7.

Figure 3. Area under the receiver operating characteristic curves showing mean model performance of the truncated version of the extended model. The models were applied to predict the probability of major depressive disorder in the: (1) training and (2) test sets. AUC: area under the receiver operating characteristic curve; CV AUC: cross-validated area under the receiver operating characteristic curve; MDD: major depressive disorder; ROC: receiver operating characteristic.

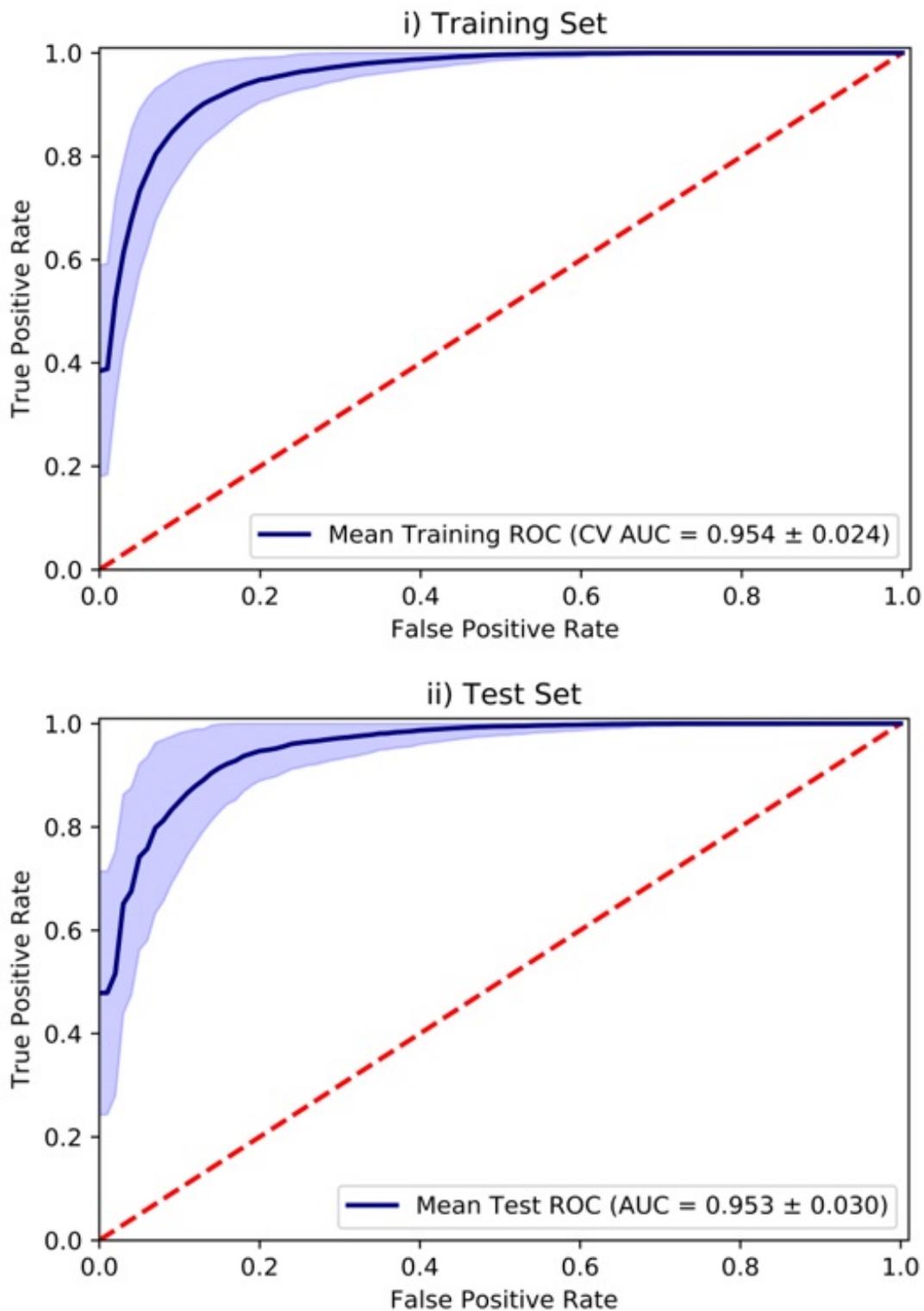
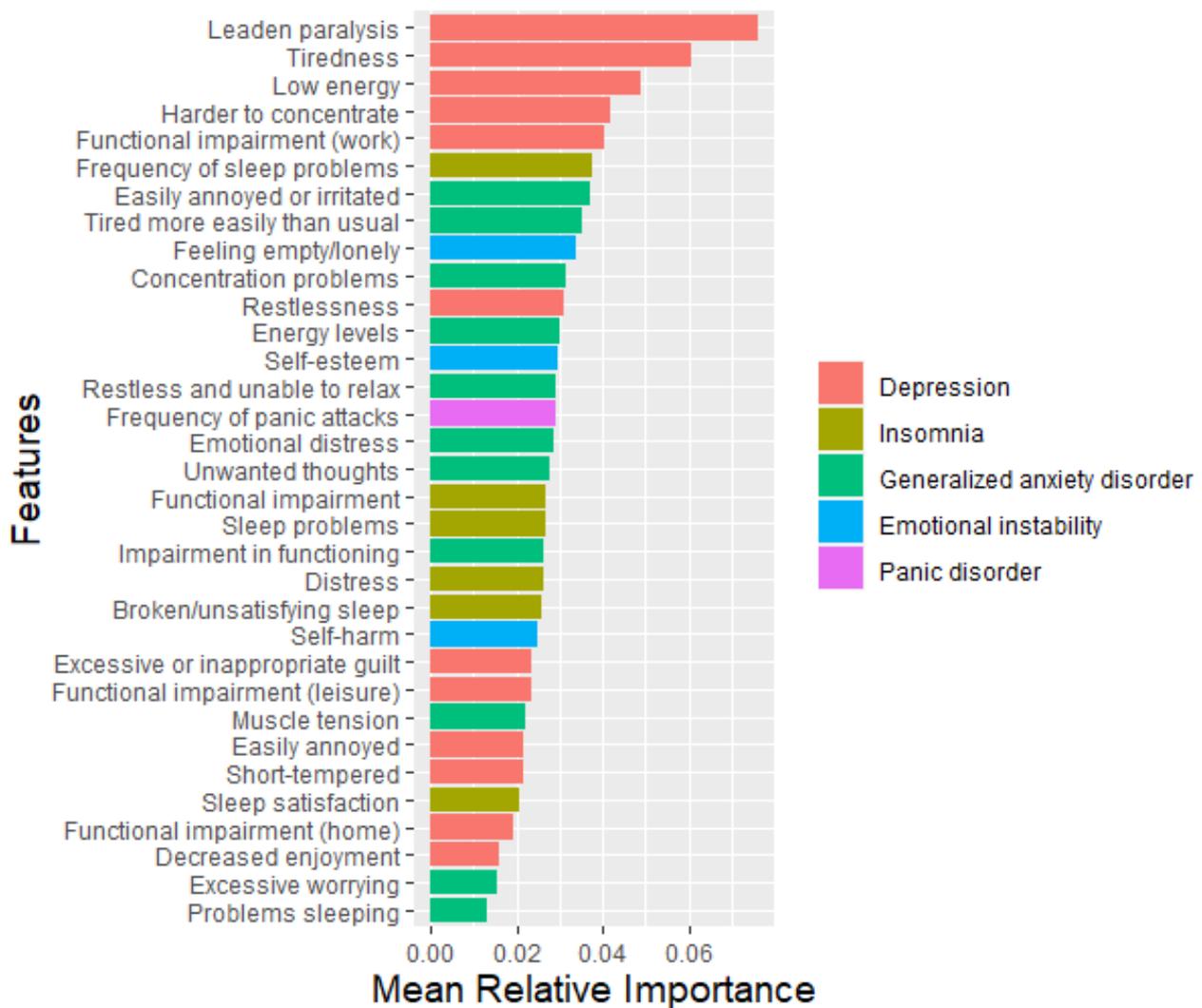


Figure 4. Mean relative importance for the 33 features in the truncated version of the extended model. Features have been ordered from most to least important and colored according to the disorder or symptom cluster they correspond to.



Diagnostic Performance of the PHQ-9

The sensitivity of the PHQ-9 for detecting MDD was 92.2%, whereas the specificity was 54.3%, and the overall diagnostic accuracy was 66.2%.

Discussion

Principal Findings

This study provides evidence for an extended definition of MDD symptomatology and supports the use of a digital assessment as an aid to clinical decision-making in the identification of MDD. Relative to a disorder-specific model of MDD psychopathology, an extended model of symptomatology was better at distinguishing between individuals with MDD and those with subthreshold levels of the disorder. In particular, a truncated version of the model, comprising symptoms of depression, GAD, insomnia, emotional instability, and panic disorder, demonstrated excellent predictive performance (AUC=0.95; sensitivity=86.5%; specificity=90.8%; and accuracy=89.5%).

Critically, although the PHQ-9 was particularly good at detecting MDD in the current sample, it tended to overdiagnose MDD in subthreshold depression and, in turn, was associated with poor overall diagnostic performance. Overdiagnosis of MDD presents a significant problem and has the potential for antidepressant overprescription and adverse drug effects in individuals who may benefit from alternative treatment options [43], such as psychotherapy or psychoeducation. Furthermore, relying on a simple cut-off score does not allow for personalized treatment plans and strategies, potentially resulting in incomplete remission rates.

Overall, the findings from our models are in line with the notion that current diagnostic practices may present a narrow conception of mental health that does not allow for the wide range of clinical signs and symptoms that are endorsed by individuals with MDD. Across our models, the most predictive symptom of MDD was leadren paralysis, which refers to an extreme form of fatigue or heavy, leaden feelings in the arms and legs. This finding is in line with a recent study by Han et al [61], whose findings revealed that leadren paralysis was a

robust and important predictor of first-onset MDD. Critically, although leaden paralysis is included in the *DSM-5* specifier for atypical depression, it is not deemed a core feature of the disorder [20]. In fact, even fatigue or loss of energy and tiredness are not considered essential symptoms according to the *DSM-5* [20]. Importantly, exhaustion, extreme tiredness, and loss of energy are typically seen in primary care settings and are often the predominant presenting complaint [62]. In fact, in a large European study comprising approximately 2000 depressed primary care patients across 6 countries, almost two-thirds of patients reported feeling tired [63]. These findings suggest that leaden paralysis and its lesser extreme variants (ie, tiredness and low energy) may be particularly important for the recognition of MDD in the primary care setting.

As expected, symptoms of GAD were among the most predictive features of MDD, with *overlapping* symptoms between GAD and depression (eg, tiredness, low energy, and irritability) being particularly indicative of the disorder. For instance, regarding the latter, irritability has been seen to occur in one-third to one-half of patients with MDD [64-66] and is associated with greater severity and chronicity, a history of suicide attempts, and reduced quality of life [64]. Other GAD-specific symptoms that were highly predictive of MDD included unwanted thoughts, excessive worrying, and emotional distress. Importantly, it could be argued that our model simply reflects the significantly higher rates of comorbidity in the MDD group relative to the subthreshold depression group, particularly with regard to GAD. Similarly, our model may capture higher levels of severity or a higher p factor in the MDD group [67] rather than important components of the condition.

Although our findings should be interpreted with caution, our view is that this should not detract from the importance of assessing for transdiagnostic symptoms of MDD, especially as these are likely to share common underlying pathophysiology and genetics [36,68]. Indeed, the evaluation of anxiety symptoms in the context of MDD is critical, particularly given the chronicity of the conditions, with research indicating that anxiety disorders may be a precursor to MDD [69]. Notably, combining clinical information with biological biomarkers, such as serum analytes, can be used to predict the development of future depressive episodes in individuals presenting with social anxiety [70] and panic disorder [71]. Identifying those who may be at a heightened risk for comorbid anxiety and depression is of clinical importance, particularly given that these individuals are likely to exhibit more pervasive and recurrent forms of illness, reduced remission rates, and increased suicidality [72-74].

Indeed, suicidality has particularly been associated with the co-occurrence of depression and panic disorder [75,76], with our findings indicating that the frequency of panic attacks was an important predictor of MDD. Importantly, comorbid panic disorder and MDD have been related to increased depression severity, an earlier age of onset, increased functional impairment, and a poorer clinical prognosis [77], suggesting that assessing for panic disorder may be important when diagnosing and treating MDD. Similarly, our findings indicate that a more in-depth evaluation of the symptoms associated with sleep problems or insomnia, including emotional distress caused by disordered sleep, warrants inclusion when diagnosing

MDD. This is particularly important as sleep problems have been shown to reduce the efficacy of depression treatment [78].

Finally, symptoms of emotional instability or personality disorder, including feeling empty, low self-esteem, and self-harm, were also seen to be important when distinguishing between MDD and subthreshold depression. Feeling empty or chronic emptiness has been closely related to depression and suicidal ideation [79]. Interestingly, a recent qualitative study in adolescents with depression revealed that a partial or complete blunting of any emotion (negative or positive), with feelings of flatness, emptiness, and lack of emotions, was an important component of anhedonia [80]. It is interesting that self-esteem but not self-worth (a diagnostic criterion for MDD) was a predictor of the disorder. This finding suggests that wording may also have an important impact on individuals' subjective evaluations of their symptoms, with the concept of self-esteem perhaps being easier to grasp than that of self-worth. Finally, although diagnostic descriptions of MDD symptomatology include suicidality as a criterion for the disorder, expanding this to include self-harm may facilitate its identification.

Taken together, our findings indicate that the current diagnostic criteria for MDD may fail to evaluate relevant clinical information that is important for the diagnosis and treatment of individuals with the disorder. Although time is a luxury in the primary care setting, our study supports the use of digital technologies as a means for obtaining a more comprehensive depiction of MDD symptomatology in a time-efficient manner. Notably, related research using the same digital mental health assessment has highlighted the utility of the tool in distinguishing individuals with MDD from those with BD [81]. Indeed, digital technologies have the potential to aid in the recognition of a wide range of psychiatric conditions, allowing for more time to be spent managing and treating symptoms. In turn, digital technologies can reduce the number of in-person appointments, alleviate health care professionals' workload, and reduce the risk of burnout. The use of digital technologies also has the potential to reduce some of the barriers associated with disclosing mental health difficulties, such as discomfort as well as issues related to stigma and discrimination. Furthermore, research has demonstrated that patients are more likely to report severe symptoms on technology platforms than to a health care professional [82] and value the independence and empowerment that can be obtained from the use of a digital platform [83].

Strengths and Limitations

To our knowledge, this is the first study to provide evidence for an extended definition of MDD symptomatology using a digital assessment. Furthermore, the digital assessment was designed following an extensive analysis of existing validated questionnaires for psychiatric disorders and diagnostic manuals, as well as input from psychiatrists and a service user group. In addition, as opposed to the use of healthy controls as a reference population against which patients are compared, our subthreshold depression group represented a clinically relevant reference group. Finally, the use of ML methods meant that patterns in data could be more readily and accurately identified, whereas our NCV approach allowed us to obtain high

algorithmic accuracy while ensuring the generalizability of the models.

This study also had several limitations. First, as with any supervised ML approach to psychopathology, our analyses were limited by the *truthfulness* of the diagnostic labels (MDD vs subthreshold depression). Second, given the adaptive (nonlinear) nature of the question flow, missing data were imputed with zeros, which may have resulted in an overly artificial data set. This should be borne in mind when interpreting these findings. In addition, the MDD sample was small in size and primarily comprised women, reflecting the higher prevalence of MDD in women than in men [84] and the difficulties in recruiting males with MDD. Furthermore, given that suicidality was an exclusion criterion in this study, measures of suicidal thoughts, ideations, plans, or impulses do not appear in our list of important features of MDD. This is a key limitation of the study, as suicidality denotes an important component of the condition. Similarly, symptoms from other disorders that frequently co-occur with MDD, such as posttraumatic stress disorder, were not available

for inclusion in our analyses, which may have allowed for a more comprehensive depiction of MDD symptomatology.

Conclusions

In an attempt to answer the question “when does depression become a mental disorder?”, our study demonstrated that a data-driven view of MDD may improve our understanding of the condition. A more comprehensive conceptualization of the psychopathology of MDD, including symptoms of depression, GAD, insomnia, panic disorder, and emotional instability, may not only facilitate patient stratification but also allow for personalized treatment plans and strategies. Although further studies with larger sample sizes are required to replicate our findings, our study shines a positive light on the use of digital technologies as an innovative way to help develop and facilitate mental health care provision. In particular, digital technologies have the capacity to collect a vast range of key clinical information that may be important for the diagnosis and treatment of individuals with MDD.

Acknowledgments

This study was funded by the Stanley Medical Research Institute (grant number: 07R-1888). The authors thank the participants for taking part in the delta study, as well as the CIDI interviewers, without whom this work would not have been possible.

Conflicts of Interest

SB is a director of Psynova Neurotech Ltd and Psyomics Ltd. SB, DC, GBO, LF, and EB have financial interests in Psyomics, Ltd, which provided funding and support for the delta study. SB, PE, and TO could benefit financially from any product that arises from work performed in the delta study.

Multimedia Appendix 1

Mean symptom severity per group was separated by disorder or symptom clusters.

[\[DOCX File , 30 KB-Multimedia Appendix 1\]](#)

Multimedia Appendix 2

Depression model: mean relative feature importance.

[\[DOCX File , 15 KB-Multimedia Appendix 2\]](#)

Multimedia Appendix 3

Depression model: percentage feature occurrences.

[\[DOCX File , 14 KB-Multimedia Appendix 3\]](#)

Multimedia Appendix 4

Receiver operating characteristic curves showing the mean predictive performance of the extended model. The models were applied to predict the probability of major depressive disorder in the training and test sets.

[\[DOCX File , 141 KB-Multimedia Appendix 4\]](#)

Multimedia Appendix 5

Extended model: mean relative feature importance colored by disorder or symptom clusters.

[\[DOCX File , 33 KB-Multimedia Appendix 5\]](#)

Multimedia Appendix 6

Extended model: percentage feature occurrences colored by disorder or symptom clusters.

[\[DOCX File , 32 KB-Multimedia Appendix 6\]](#)

Multimedia Appendix 7

Truncated model: percentage of feature occurrences colored by disorder or symptom clusters.

[\[DOCX File , 20 KB-Multimedia Appendix 7\]](#)

References

1. Ferrari AJ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray CJ, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS Med* 2013 Nov;10(11):e1001547 [[FREE Full text](#)] [doi: [10.1371/journal.pmed.1001547](https://doi.org/10.1371/journal.pmed.1001547)] [Medline: [24223526](https://pubmed.ncbi.nlm.nih.gov/24223526/)]
2. Holden C. Mental health. Global survey examines impact of depression. *Science* 2000 Apr 07;288(5463):39-40. [doi: [10.1126/science.288.5463.39](https://doi.org/10.1126/science.288.5463.39)] [Medline: [10766633](https://pubmed.ncbi.nlm.nih.gov/10766633/)]
3. Wang PS, Aguilar-Gaxiola S, Alonso J, Angermeyer MC, Borges G, Bromet EJ, et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *Lancet* 2007 Sep 8;370(9590):841-850 [[FREE Full text](#)] [doi: [10.1016/S0140-6736\(07\)61414-7](https://doi.org/10.1016/S0140-6736(07)61414-7)] [Medline: [17826169](https://pubmed.ncbi.nlm.nih.gov/17826169/)]
4. Goldney RD, Fisher LJ, Dal GE, Taylor AW. Subsyndromal depression: prevalence, use of health services and quality of life in an Australian population. *Soc Psychiatry Psychiatr Epidemiol* 2004 Apr;39(4):293-298. [doi: [10.1007/s00127-004-0745-5](https://doi.org/10.1007/s00127-004-0745-5)] [Medline: [15085331](https://pubmed.ncbi.nlm.nih.gov/15085331/)]
5. Fried EI, Nesse RM, Zivin K, Guille C, Sen S. Depression is more than the sum score of its parts: individual DSM symptoms have different risk factors. *Psychol Med* 2014 Jul;44(10):2067-2076 [[FREE Full text](#)] [doi: [10.1017/S0033291713002900](https://doi.org/10.1017/S0033291713002900)] [Medline: [24289852](https://pubmed.ncbi.nlm.nih.gov/24289852/)]
6. Mitchell AJ, Vaze A, Rao S. Clinical diagnosis of depression in primary care: a meta-analysis. *Lancet* 2009 Aug 22;374(9690):609-619. [doi: [10.1016/S0140-6736\(09\)60879-5](https://doi.org/10.1016/S0140-6736(09)60879-5)] [Medline: [19640579](https://pubmed.ncbi.nlm.nih.gov/19640579/)]
7. Hek K, Demirkan A, Lahti J, Terracciano A, Teumer A, Cornelis MC, et al. A genome-wide association study of depressive symptoms. *Biol Psychiatry* 2013 Apr 01;73(7):667-678 [[FREE Full text](#)] [doi: [10.1016/j.biopsych.2012.09.033](https://doi.org/10.1016/j.biopsych.2012.09.033)] [Medline: [23290196](https://pubmed.ncbi.nlm.nih.gov/23290196/)]
8. Lewis CM, Ng MY, Butler AW, Cohen-Woods S, Uher R, Pirilo K, et al. Genome-wide association study of major recurrent depression in the U.K. population. *Am J Psychiatry* 2010 Aug;167(8):949-957. [doi: [10.1176/appi.ajp.2010.09091380](https://doi.org/10.1176/appi.ajp.2010.09091380)] [Medline: [20516156](https://pubmed.ncbi.nlm.nih.gov/20516156/)]
9. Shi J, Potash JB, Knowles JA, Weissman MM, Coryell W, Scheftner WA, et al. Genome-wide association study of recurrent early-onset major depressive disorder. *Mol Psychiatry* 2011 Feb;16(2):193-201 [[FREE Full text](#)] [doi: [10.1038/mp.2009.124](https://doi.org/10.1038/mp.2009.124)] [Medline: [20125088](https://pubmed.ncbi.nlm.nih.gov/20125088/)]
10. Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium, Ripke S, Wray NR, Lewis CM, Hamilton SP, Weissman MM, et al. A mega-analysis of genome-wide association studies for major depressive disorder. *Mol Psychiatry* 2013 Apr;18(4):497-511 [[FREE Full text](#)] [doi: [10.1038/mp.2012.21](https://doi.org/10.1038/mp.2012.21)] [Medline: [22472876](https://pubmed.ncbi.nlm.nih.gov/22472876/)]
11. Wray NR, Pergadia ML, Blackwood DH, Penninx BW, Gordon SD, Nyholt DR, et al. Genome-wide association study of major depressive disorder: new results, meta-analysis, and lessons learned. *Mol Psychiatry* 2012 Jan;17(1):36-48 [[FREE Full text](#)] [doi: [10.1038/mp.2010.109](https://doi.org/10.1038/mp.2010.109)] [Medline: [21042317](https://pubmed.ncbi.nlm.nih.gov/21042317/)]
12. 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 Dec;17(12):1174-1179. [doi: [10.1038/mp.2012.105](https://doi.org/10.1038/mp.2012.105)] [Medline: [22869033](https://pubmed.ncbi.nlm.nih.gov/22869033/)]
13. Insel T, Cuthbert B, Garvey M, Heinssen R, Pine DS, Quinn K, et al. Research domain criteria (RDoC): toward a new classification framework for research on mental disorders. *Am J Psychiatry* 2010 Jul;167(7):748-751. [doi: [10.1176/appi.ajp.2010.09091379](https://doi.org/10.1176/appi.ajp.2010.09091379)] [Medline: [20595427](https://pubmed.ncbi.nlm.nih.gov/20595427/)]
14. Wittchen H, Beesdo K, Bittner A, Goodwin RD. Depressive episodes—evidence for a causal role of primary anxiety disorders? *Eur Psychiatry* 2003 Dec;18(8):384-393. [doi: [10.1016/j.eurpsy.2003.10.001](https://doi.org/10.1016/j.eurpsy.2003.10.001)] [Medline: [14680714](https://pubmed.ncbi.nlm.nih.gov/14680714/)]
15. Hankin BL, Fraley RC, Lahey BB, Waldman ID. Is depression best viewed as a continuum or discrete category? A taxometric analysis of childhood and adolescent depression in a population-based sample. *J Abnorm Psychol* 2005 Feb;114(1):96-110. [doi: [10.1037/0021-843X.114.1.96](https://doi.org/10.1037/0021-843X.114.1.96)] [Medline: [15709816](https://pubmed.ncbi.nlm.nih.gov/15709816/)]
16. Haslam N, Holland E, Kuppens P. Categories versus dimensions in personality and psychopathology: a quantitative review of taxometric research. *Psychol Med* 2012 May;42(5):903-920. [doi: [10.1017/S0033291711001966](https://doi.org/10.1017/S0033291711001966)] [Medline: [21939592](https://pubmed.ncbi.nlm.nih.gov/21939592/)]
17. Wright AG, Krueger RF, Hobbs MJ, Markon KE, Eaton NR, Slade T. The structure of psychopathology: toward an expanded quantitative empirical model. *J Abnorm Psychol* 2013 Feb;122(1):281-294 [[FREE Full text](#)] [doi: [10.1037/a0030133](https://doi.org/10.1037/a0030133)] [Medline: [23067258](https://pubmed.ncbi.nlm.nih.gov/23067258/)]
18. Dalgleish T, Black M, Johnston D, Bevan A. Transdiagnostic approaches to mental health problems: current status and future directions. *J Consult Clin Psychol* 2020 Mar;88(3):179-195 [[FREE Full text](#)] [doi: [10.1037/ccp0000482](https://doi.org/10.1037/ccp0000482)] [Medline: [32068421](https://pubmed.ncbi.nlm.nih.gov/32068421/)]
19. Kendler KS. The nature of psychiatric disorders. *World Psychiatry* 2016 Feb;15(1):5-12 [[FREE Full text](#)] [doi: [10.1002/wps.20292](https://doi.org/10.1002/wps.20292)] [Medline: [26833596](https://pubmed.ncbi.nlm.nih.gov/26833596/)]
20. American Psychiatric Association. *Diagnostic and Statistical Manual of Mental Disorders (5th Edition)*. Washington, DC: American Psychiatric Publication; 2013.

21. The ICD-11 classification of mental and behavioural disorders: diagnostic criteria for research. World Health Organization. 2018. URL: <https://www.who.int/classifications/icd/en/GRNBOOK.pdf> [accessed 2021-09-14]
22. Braam AW, Copeland JR, Delespaul PA, Beekman AT, Como A, Dewey M, et al. Depression, subthreshold depression and comorbid anxiety symptoms in older europeans: results from the EURODEP concerted action. *J Affect Disord* 2014 Feb;155:266-272. [doi: [10.1016/j.jad.2013.11.011](https://doi.org/10.1016/j.jad.2013.11.011)] [Medline: [24355647](https://pubmed.ncbi.nlm.nih.gov/24355647/)]
23. Brown TA, Campbell LA, Lehman CL, Grisham JR, Mancill RB. Current and lifetime comorbidity of the DSM-IV anxiety and mood disorders in a large clinical sample. *J Abnorm Psychol* 2001 Nov;110(4):585-599. [doi: [10.1037//0021-843x.110.4.585](https://doi.org/10.1037//0021-843x.110.4.585)] [Medline: [11727948](https://pubmed.ncbi.nlm.nih.gov/11727948/)]
24. Kessler RC, Berglund P, Demler O, Jin R, Merikangas KR, Walters EE. Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the national comorbidity survey replication. *Arch Gen Psychiatry* 2005 Jun;62(6):593-602. [doi: [10.1001/archpsyc.62.6.593](https://doi.org/10.1001/archpsyc.62.6.593)] [Medline: [15939837](https://pubmed.ncbi.nlm.nih.gov/15939837/)]
25. Kessler RC, Merikangas KR, Wang PS. Prevalence, comorbidity, and service utilization for mood disorders in the united states at the beginning of the twenty-first century. *Annu Rev Clin Psychol* 2007;3:137-158. [doi: [10.1146/annurev.clinpsy.3.022806.091444](https://doi.org/10.1146/annurev.clinpsy.3.022806.091444)] [Medline: [17716051](https://pubmed.ncbi.nlm.nih.gov/17716051/)]
26. Zimmerman M, McDermt W, Mattia JI. Frequency of anxiety disorders in psychiatric outpatients with major depressive disorder. *Am J Psychiatry* 2000 Aug;157(8):1337-1340. [doi: [10.1176/appi.ajp.157.8.1337](https://doi.org/10.1176/appi.ajp.157.8.1337)] [Medline: [10910803](https://pubmed.ncbi.nlm.nih.gov/10910803/)]
27. Pedrelli P, Blais MA, Alpert JE, Shelton RC, Walker RS, Fava M. Reliability and validity of the symptoms of depression questionnaire (SDQ). *CNS Spectr* 2014 Dec;19(6):535-546 [FREE Full text] [doi: [10.1017/S1092852914000406](https://doi.org/10.1017/S1092852914000406)] [Medline: [25275853](https://pubmed.ncbi.nlm.nih.gov/25275853/)]
28. Chelminski I, Zimmerman M. Pathological worry in depressed and anxious patients. *J Anxiety Disord* 2003;17(5):533-546. [doi: [10.1016/s0887-6185\(02\)00246-3](https://doi.org/10.1016/s0887-6185(02)00246-3)] [Medline: [12941364](https://pubmed.ncbi.nlm.nih.gov/12941364/)]
29. Gask L, Klinkman M, Fortes S, Dowrick C. Capturing complexity: the case for a new classification system for mental disorders in primary care. *Eur Psychiatry* 2008 Oct;23(7):469-476. [doi: [10.1016/j.eurpsy.2008.06.006](https://doi.org/10.1016/j.eurpsy.2008.06.006)] [Medline: [18774269](https://pubmed.ncbi.nlm.nih.gov/18774269/)]
30. Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am J Psychiatry* 2006 Nov;163(11):1905-1917. [doi: [10.1176/appi.ajp.163.11.1905](https://doi.org/10.1176/appi.ajp.163.11.1905)] [Medline: [17074942](https://pubmed.ncbi.nlm.nih.gov/17074942/)]
31. Souery D, Oswald P, Massat I, Bailer U, Bollen J, Demyttenaere K, Group for the Study of Resistant Depression. Clinical factors associated with treatment resistance in major depressive disorder: results from a European Multicenter Study. *J Clin Psychiatry* 2007 Jul;68(7):1062-1070. [doi: [10.4088/jcp.v68n0713](https://doi.org/10.4088/jcp.v68n0713)] [Medline: [17685743](https://pubmed.ncbi.nlm.nih.gov/17685743/)]
32. Paykel ES. Partial remission, residual symptoms, and relapse in depression. *Dialogues Clin Neurosci* 2008;10(4):431-437 [FREE Full text] [doi: [10.31887/DCNS.2008.10.4/espaykel](https://doi.org/10.31887/DCNS.2008.10.4/espaykel)] [Medline: [19170400](https://pubmed.ncbi.nlm.nih.gov/19170400/)]
33. Saragoussi D, Touya M, Haro JM, Jönsson B, Knapp M, Botrel B, et al. Factors associated with failure to achieve remission and with relapse after remission in patients with major depressive disorder in the PERFORM study. *Neuropsychiatr Dis Treat* 2017;13:2151-2165 [FREE Full text] [doi: [10.2147/NDT.S136343](https://doi.org/10.2147/NDT.S136343)] [Medline: [28860772](https://pubmed.ncbi.nlm.nih.gov/28860772/)]
34. Jaffe DH, Rive B, Denee TR. The humanistic and economic burden of treatment-resistant depression in europe: a cross-sectional study. *BMC Psychiatry* 2019 Aug 07;19(1):247 [FREE Full text] [doi: [10.1186/s12888-019-2222-4](https://doi.org/10.1186/s12888-019-2222-4)] [Medline: [31391065](https://pubmed.ncbi.nlm.nih.gov/31391065/)]
35. Bergfeld IO, Mantione M, Figeo M, Schuurman PR, Lok A, Denys D. Treatment-resistant depression and suicidality. *J Affect Disord* 2018 Aug 01;235:362-367. [doi: [10.1016/j.jad.2018.04.016](https://doi.org/10.1016/j.jad.2018.04.016)] [Medline: [29665520](https://pubmed.ncbi.nlm.nih.gov/29665520/)]
36. Kendler KS. Major depression and generalised anxiety disorder. same genes, (partly)different environments--revisited. *Br J Psychiatry Suppl* 1996 Jun(30):68-75. [doi: [10.1192/S0007125000298437](https://doi.org/10.1192/S0007125000298437)] [Medline: [8864151](https://pubmed.ncbi.nlm.nih.gov/8864151/)]
37. Frankland A, Cerrillo E, Hadzi-Pavlovic D, Roberts G, Wright A, Loo CK, et al. Comparing the phenomenology of depressive episodes in bipolar I and II disorder and major depressive disorder within bipolar disorder pedigrees. *J Clin Psychiatry* 2015 Jan;76(1):32-38. [doi: [10.4088/JCP.14m09293](https://doi.org/10.4088/JCP.14m09293)] [Medline: [25650671](https://pubmed.ncbi.nlm.nih.gov/25650671/)]
38. Hirschfeld RM. Differential diagnosis of bipolar disorder and major depressive disorder. *J Affect Disord* 2014 Dec;169 Suppl 1:12-16 [FREE Full text] [doi: [10.1016/S0165-0327\(14\)70004-7](https://doi.org/10.1016/S0165-0327(14)70004-7)] [Medline: [25533909](https://pubmed.ncbi.nlm.nih.gov/25533909/)]
39. Bauer M, Andreassen OA, Geddes JR, Vedel Kessing L, Lewitzka U, Schulze TG, et al. Areas of uncertainties and unmet needs in bipolar disorders: clinical and research perspectives. *Lancet Psychiatry* 2018 Nov;5(11):930-939. [doi: [10.1016/S2215-0366\(18\)30253-0](https://doi.org/10.1016/S2215-0366(18)30253-0)] [Medline: [30146246](https://pubmed.ncbi.nlm.nih.gov/30146246/)]
40. Patel R, Shetty H, Jackson R, Broadbent M, Stewart R, Boydell J, et al. Delays before diagnosis and initiation of treatment in patients presenting to mental health services with bipolar disorder. *PLoS One* 2015;10(5):e0126530 [FREE Full text] [doi: [10.1371/journal.pone.0126530](https://doi.org/10.1371/journal.pone.0126530)] [Medline: [25992560](https://pubmed.ncbi.nlm.nih.gov/25992560/)]
41. Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001 Sep;16(9):606-613 [FREE Full text] [doi: [10.1046/j.1525-1497.2001.016009606.x](https://doi.org/10.1046/j.1525-1497.2001.016009606.x)] [Medline: [11556941](https://pubmed.ncbi.nlm.nih.gov/11556941/)]
42. Zimmerman M, Martinez JH, Friedman M, Boerescu DA, Attiullah N, Toba C. How can we use depression severity to guide treatment selection when measures of depression categorize patients differently? *J Clin Psychiatry* 2012 Oct;73(10):1287-1291. [doi: [10.4088/JCP.12m07775](https://doi.org/10.4088/JCP.12m07775)] [Medline: [23058983](https://pubmed.ncbi.nlm.nih.gov/23058983/)]

43. Jerant A, Kravitz RL, Garcia EF, Feldman MD, Cipri C, Nishio D, et al. Potential antidepressant overtreatment associated with office use of brief depression symptom measures. *J Am Board Fam Med* 2014;27(5):611-620 [FREE Full text] [doi: [10.3122/jabfm.2014.05.140038](https://doi.org/10.3122/jabfm.2014.05.140038)] [Medline: [25201931](https://pubmed.ncbi.nlm.nih.gov/25201931/)]
44. Mitchell AJ, Rao S, Vaze A. International comparison of clinicians' ability to identify depression in primary care: meta-analysis and meta-regression of predictors. *Br J Gen Pract* 2011 Feb;61(583):72-80 [FREE Full text] [doi: [10.3399/bjgp11X556227](https://doi.org/10.3399/bjgp11X556227)] [Medline: [21276327](https://pubmed.ncbi.nlm.nih.gov/21276327/)]
45. Malpass A, Dowrick C, Gilbody S, Robinson J, Wiles N, Duffy L, et al. Usefulness of PHQ-9 in primary care to determine meaningful symptoms of low mood: a qualitative study. *Br J Gen Pract* 2016 Feb;66(643):78-84 [FREE Full text] [doi: [10.3399/bjgp16X683473](https://doi.org/10.3399/bjgp16X683473)] [Medline: [26823268](https://pubmed.ncbi.nlm.nih.gov/26823268/)]
46. Chen M, Mao S, Liu Y. Big data: a survey. *Mobile Netw Appl* 2014 Jan 22;19(2):171-209. [doi: [10.1007/s11036-013-0489-0](https://doi.org/10.1007/s11036-013-0489-0)]
47. Olmert T, Cooper JD, Han SY, Barton-Owen G, Farrag L, Bell E, et al. A combined digital and biomarker diagnostic aid for mood disorders (the delta trial): protocol for an observational study. *JMIR Res Protoc* 2020 Aug 10;9(8):e18453 [FREE Full text] [doi: [10.2196/18453](https://doi.org/10.2196/18453)] [Medline: [32773373](https://pubmed.ncbi.nlm.nih.gov/32773373/)]
48. Kessler RC, Ustün 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(2):93-121. [doi: [10.1002/mpr.168](https://doi.org/10.1002/mpr.168)] [Medline: [15297906](https://pubmed.ncbi.nlm.nih.gov/15297906/)]
49. Hirschfeld RM. The mood disorder questionnaire: a simple, patient-rated screening instrument for bipolar disorder. *Prim Care Companion J Clin Psychiatry* 2002 Feb;4(1):9-11 [FREE Full text] [doi: [10.4088/pcc.v04n0104](https://doi.org/10.4088/pcc.v04n0104)] [Medline: [15014728](https://pubmed.ncbi.nlm.nih.gov/15014728/)]
50. Eckblad M, Chapman LJ. Development and validation of a scale for hypomanic personality. *J Abnorm Psychol* 1986 Aug;95(3):214-222. [doi: [10.1037//0021-843x.95.3.214](https://doi.org/10.1037//0021-843x.95.3.214)] [Medline: [3745642](https://pubmed.ncbi.nlm.nih.gov/3745642/)]
51. Ghaemi SN, Miller CJ, Berv DA, Klugman J, Rosenquist KJ, Pies RW. Sensitivity and specificity of a new bipolar spectrum diagnostic scale. *J Affect Disord* 2005 Feb;84(2-3):273-277. [doi: [10.1016/S0165-0327\(03\)00196-4](https://doi.org/10.1016/S0165-0327(03)00196-4)] [Medline: [15708426](https://pubmed.ncbi.nlm.nih.gov/15708426/)]
52. Angst J, Adolfsson R, Benazzi F, Gamma A, Hantouche E, Meyer TD, et al. The HCL-32: towards a self-assessment tool for hypomanic symptoms in outpatients. *J Affect Disord* 2005 Oct;88(2):217-233. [doi: [10.1016/j.jad.2005.05.011](https://doi.org/10.1016/j.jad.2005.05.011)] [Medline: [16125784](https://pubmed.ncbi.nlm.nih.gov/16125784/)]
53. Leung CM, Yim CL, Yan CT, Chan CC, Xiang Y, Mak AD, et al. The bipolar II depression questionnaire: a self-report tool for detecting bipolar II depression. *PLoS One* 2016;11(3):e0149752 [FREE Full text] [doi: [10.1371/journal.pone.0149752](https://doi.org/10.1371/journal.pone.0149752)] [Medline: [26963908](https://pubmed.ncbi.nlm.nih.gov/26963908/)]
54. Depue RA, Slater JF, Wolfstetter-Kausch H, Klein D, Goplerud E, Farr D. A behavioral paradigm for identifying persons at risk for bipolar depressive disorder: a conceptual framework and five validation studies. *J Abnorm Psychol* 1981 Oct;90(5):381-437. [doi: [10.1037//0021-843x.90.5.381](https://doi.org/10.1037//0021-843x.90.5.381)] [Medline: [7298991](https://pubmed.ncbi.nlm.nih.gov/7298991/)]
55. Altman EG, Hedeker D, Peterson JL, Davis JM. The altman self-rating mania scale. *Biol Psychiatry* 1997 Nov 15;42(10):948-955. [doi: [10.1016/S0006-3223\(96\)00548-3](https://doi.org/10.1016/S0006-3223(96)00548-3)] [Medline: [9359982](https://pubmed.ncbi.nlm.nih.gov/9359982/)]
56. First M, Williams J, Spitzer RL, Gibbon M. Structured clinical interview for DSM-IV axis I disorders (SCID). ePROVIDE - Mapi Research Trust. URL: <https://eprovide.mapi-trust.org/instruments/structured-clinical-interview-for-dsm-iv-axis-i-disorders> [accessed 2021-09-14]
57. Sheehan DV, Lecrubier Y, Sheehan KH, Amorim P, Janavs J, Weiller E, et al. The Mini-International Neuropsychiatric Interview (M.I.N.I.): the development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *J Clin Psychiatry* 1998;59 Suppl 20:22-33. [Medline: [9881538](https://pubmed.ncbi.nlm.nih.gov/9881538/)]
58. Cohen J. *Statistical Power Analysis for the Behavioral Sciences* (2nd Edition). New York: Academic Press; 1988:1-978.
59. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology* 1983 Sep;148(3):839-843. [doi: [10.1148/radiology.148.3.6878708](https://doi.org/10.1148/radiology.148.3.6878708)] [Medline: [6878708](https://pubmed.ncbi.nlm.nih.gov/6878708/)]
60. Pedregosa F, Gramfort A, Michel V, Thirion B, Varoquaux G, Grisel O, et al. Scikit-learn: machine learning in python. *J Mach Learn Res* 2011;12:2825-2830 [FREE Full text]
61. Han SY, Cooper JD, Ozcan S, Rustogi N, Penninx BW, Bahn S. Integrating proteomic, sociodemographic and clinical data to predict future depression diagnosis in subthreshold symptomatic individuals. *Transl Psychiatry* 2019 Nov 07;9(1):277 [FREE Full text] [doi: [10.1038/s41398-019-0623-2](https://doi.org/10.1038/s41398-019-0623-2)] [Medline: [31699963](https://pubmed.ncbi.nlm.nih.gov/31699963/)]
62. Hamilton W, Watson J, Round A. Investigating fatigue in primary care. *Br Med J* 2010 Aug 24;341:c4259. [doi: [10.1136/bmj.c4259](https://doi.org/10.1136/bmj.c4259)] [Medline: [20736254](https://pubmed.ncbi.nlm.nih.gov/20736254/)]
63. Tylee A, Gastpar M, Lépine JP, Mendlewicz J. DEPRES II (depression research in european society II): a patient survey of the symptoms, disability and current management of depression in the community. DEPRES Steering Committee. *Int Clin Psychopharmacol* 1999 May;14(3):139-151. [doi: [10.1097/00004850-199905002-00001](https://doi.org/10.1097/00004850-199905002-00001)] [Medline: [10435767](https://pubmed.ncbi.nlm.nih.gov/10435767/)]
64. Fava M, Hwang I, Rush AJ, Sampson N, Walters EE, Kessler RC. The importance of irritability as a symptom of major depressive disorder: results from the national comorbidity survey replication. *Mol Psychiatry* 2010 Aug;15(8):856-867 [FREE Full text] [doi: [10.1038/mp.2009.20](https://doi.org/10.1038/mp.2009.20)] [Medline: [19274052](https://pubmed.ncbi.nlm.nih.gov/19274052/)]
65. Judd LL, Schettler PJ, Coryell W, Akiskal HS, Fiedorowicz JG. Overt irritability/anger in unipolar major depressive episodes: past and current characteristics and implications for long-term course. *JAMA Psychiatry* 2013 Nov;70(11):1171-1180. [doi: [10.1001/jamapsychiatry.2013.1957](https://doi.org/10.1001/jamapsychiatry.2013.1957)] [Medline: [24026579](https://pubmed.ncbi.nlm.nih.gov/24026579/)]

66. Posternak MA, Zimmerman M. Anger and aggression in psychiatric outpatients. *J Clin Psychiatry* 2002 Aug;63(8):665-672. [doi: [10.4088/jcp.v63n0803](https://doi.org/10.4088/jcp.v63n0803)] [Medline: [12197446](https://pubmed.ncbi.nlm.nih.gov/12197446/)]
67. Caspi A, Houts RM, Belsky DW, Goldman-Mellor SJ, Harrington H, Israel S, et al. The p factor: one general psychopathology factor in the structure of psychiatric disorders? *Clin Psychol Sci* 2014 Mar;2(2):119-137 [FREE Full text] [doi: [10.1177/2167702613497473](https://doi.org/10.1177/2167702613497473)] [Medline: [25360393](https://pubmed.ncbi.nlm.nih.gov/25360393/)]
68. Hettema JM. What is the genetic relationship between anxiety and depression? *Am J Med Genet C Semin Med Genet* 2008 May 15;148C(2):140-146. [doi: [10.1002/ajmg.c.30171](https://doi.org/10.1002/ajmg.c.30171)] [Medline: [18412101](https://pubmed.ncbi.nlm.nih.gov/18412101/)]
69. Bittner A, Goodwin RD, Wittchen H, Beesdo K, Höfler M, Lieb R. What characteristics of primary anxiety disorders predict subsequent major depressive disorder? *J Clin Psychiatry* 2004 May;65(5):618-626. [doi: [10.4088/jcp.v65n0505](https://doi.org/10.4088/jcp.v65n0505)] [Medline: [15163247](https://pubmed.ncbi.nlm.nih.gov/15163247/)]
70. Gottschalk MG, Cooper JD, Chan MK, Bot M, Penninx BW, Bahn S. Discovery of serum biomarkers predicting development of a subsequent depressive episode in social anxiety disorder. *Brain Behav Immun* 2015 Aug;48:123-131. [doi: [10.1016/j.bbi.2015.04.011](https://doi.org/10.1016/j.bbi.2015.04.011)] [Medline: [25929723](https://pubmed.ncbi.nlm.nih.gov/25929723/)]
71. Gottschalk MG, Cooper JD, Chan MK, Bot M, Penninx BW, Bahn S. Serum biomarkers predictive of depressive episodes in panic disorder. *J Psychiatr Res* 2016 Feb;73:53-62. [doi: [10.1016/j.jpsychires.2015.11.012](https://doi.org/10.1016/j.jpsychires.2015.11.012)] [Medline: [26687614](https://pubmed.ncbi.nlm.nih.gov/26687614/)]
72. Fava M, Rush AJ, Alpert JE, Balasubramani GK, Wisniewski SR, Carmin CN, et al. Difference in treatment outcome in outpatients with anxious versus nonanxious depression: a STAR*D report. *Am J Psychiatry* 2008 Mar;165(3):342-351. [doi: [10.1176/appi.ajp.2007.06111868](https://doi.org/10.1176/appi.ajp.2007.06111868)] [Medline: [18172020](https://pubmed.ncbi.nlm.nih.gov/18172020/)]
73. Penninx BW, Nolen WA, Lamers F, Zitman FG, Smit JH, Spinhoven P, et al. Two-year course of depressive and anxiety disorders: results from the netherlands study of depression and anxiety (NESDA). *J Affect Disord* 2011 Sep;133(1-2):76-85 [FREE Full text] [doi: [10.1016/j.jad.2011.03.027](https://doi.org/10.1016/j.jad.2011.03.027)] [Medline: [21496929](https://pubmed.ncbi.nlm.nih.gov/21496929/)]
74. Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *Am J Psychiatry* 2006 Jan;163(1):28-40. [doi: [10.1176/appi.ajp.163.1.28](https://doi.org/10.1176/appi.ajp.163.1.28)] [Medline: [16390886](https://pubmed.ncbi.nlm.nih.gov/16390886/)]
75. Nam Y, Kim C, Roh D. Comorbid panic disorder as an independent risk factor for suicide attempts in depressed outpatients. *Compr Psychiatry* 2016 May;67:13-18. [doi: [10.1016/j.comppsy.2016.02.011](https://doi.org/10.1016/j.comppsy.2016.02.011)] [Medline: [27095329](https://pubmed.ncbi.nlm.nih.gov/27095329/)]
76. Katz C, Yaseen ZS, Mojtabai R, Cohen LJ, Galyunker II. Panic as an independent risk factor for suicide attempt in depressive illness: findings from the national epidemiological survey on alcohol and related conditions (NESARC). *J Clin Psychiatry* 2011 Dec;72(12):1628-1635. [doi: [10.4088/JCP.10m06186blu](https://doi.org/10.4088/JCP.10m06186blu)] [Medline: [21457675](https://pubmed.ncbi.nlm.nih.gov/21457675/)]
77. Roy-Byrne PP, Stang P, Wittchen HU, Ustun B, Walters EE, Kessler RC. Lifetime panic-depression comorbidity in the national comorbidity survey. Association with symptoms, impairment, course and help-seeking. *Br J Psychiatry* 2000 Mar;176:229-235. [doi: [10.1192/bjp.176.3.229](https://doi.org/10.1192/bjp.176.3.229)] [Medline: [10755069](https://pubmed.ncbi.nlm.nih.gov/10755069/)]
78. Hinkelmann K, Moritz S, Botzenhardt J, Muhtz C, Wiedemann K, Kellner M, et al. Changes in cortisol secretion during antidepressive treatment and cognitive improvement in patients with major depression: a longitudinal study. *Psychoneuroendocrinology* 2012 May;37(5):685-692. [doi: [10.1016/j.psyneuen.2011.08.012](https://doi.org/10.1016/j.psyneuen.2011.08.012)] [Medline: [21944955](https://pubmed.ncbi.nlm.nih.gov/21944955/)]
79. Johansen M, Karterud S, Pedersen G, Gude T, Falkum E. An investigation of the prototype validity of the borderline DSM-IV construct. *Acta Psychiatr Scand* 2004 Apr;109(4):289-298. [doi: [10.1046/j.1600-0447.2003.00268.x](https://doi.org/10.1046/j.1600-0447.2003.00268.x)] [Medline: [15008803](https://pubmed.ncbi.nlm.nih.gov/15008803/)]
80. Watson R, Harvey K, McCabe C, Reynolds S. Understanding anhedonia: a qualitative study exploring loss of interest and pleasure in adolescent depression. *Eur Child Adolesc Psychiatry* 2020 Apr;29(4):489-499 [FREE Full text] [doi: [10.1007/s00787-019-01364-y](https://doi.org/10.1007/s00787-019-01364-y)] [Medline: [31270605](https://pubmed.ncbi.nlm.nih.gov/31270605/)]
81. Tomasik J, Han SY, Barton-Owen G, Mirea D, Martin-Key NA, Rustogi N, et al. A machine learning algorithm to differentiate bipolar disorder from major depressive disorder using an online mental health questionnaire and blood biomarker data. *Transl Psychiatry* 2021 Jan 12;11(1):41 [FREE Full text] [doi: [10.1038/s41398-020-01181-x](https://doi.org/10.1038/s41398-020-01181-x)] [Medline: [33436544](https://pubmed.ncbi.nlm.nih.gov/33436544/)]
82. Torous J, Staples P, Shanahan M, Lin C, Peck P, Keshavan M, et al. Utilizing a personal smartphone custom app to assess the patient health questionnaire-9 (PHQ-9) depressive symptoms in patients with major depressive disorder. *JMIR Ment Health* 2015;2(1):e8 [FREE Full text] [doi: [10.2196/mental.3889](https://doi.org/10.2196/mental.3889)] [Medline: [26543914](https://pubmed.ncbi.nlm.nih.gov/26543914/)]
83. Knowles SE, Toms G, Sanders C, Bee P, Lovell K, Rennick-Egglestone S, et al. Qualitative meta-synthesis of user experience of computerised therapy for depression and anxiety. *PLoS One* 2014;9(1):e84323 [FREE Full text] [doi: [10.1371/journal.pone.0084323](https://doi.org/10.1371/journal.pone.0084323)] [Medline: [24465404](https://pubmed.ncbi.nlm.nih.gov/24465404/)]
84. Albert PR. Why is depression more prevalent in women? *J Psychiatry Neurosci* 2015 Jul;40(4):219-221 [FREE Full text] [doi: [10.1503/jpn.150205](https://doi.org/10.1503/jpn.150205)] [Medline: [26107348](https://pubmed.ncbi.nlm.nih.gov/26107348/)]

Abbreviations

AUC: area under the receiver operating characteristic curve

BD: bipolar disorder

CIDI: Composite International Diagnostic Interview

DSM-IV: Diagnostic and Statistical Manual of Mental Disorders, Fourth Edition

DSM-5: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition

GAD: generalized anxiety disorder

ICD-10: International Statistical Classification of Diseases, Tenth Revision

ICD-11: International Statistical Classification of Diseases, Eleventh Revision

MDD: major depressive disorder

ML: machine learning

NCV: nested cross-validation

PHQ-9: Patient Health Questionnaire-9

Edited by G Eysenbach; submitted 15.02.21; peer-reviewed by M Gottschalk, R Pine; comments to author 07.03.21; revised version received 17.03.21; accepted 01.08.21; published 28.10.21

Please cite as:

Martin-Key NA, Mirea DM, Olmert T, Cooper J, Han SYS, Barton-Owen G, Farrag L, Bell E, Eljasz P, Cowell D, Tomasik J, Bahn S

Toward an Extended Definition of Major Depressive Disorder Symptomatology: Digital Assessment and Cross-validation Study

JMIR Form Res 2021;5(10):e27908

URL: <https://formative.jmir.org/2021/10/e27908>

doi: [10.2196/27908](https://doi.org/10.2196/27908)

PMID:

©Nayra A Martin-Key, Dan-Mircea Mirea, Tony Olmert, Jason Cooper, Sung Yeon Sarah Han, Giles Barton-Owen, Lynn Farrag, Emily Bell, Pawel Eljasz, Daniel Cowell, Jakub Tomasik, Sabine Bahn. Originally published in JMIR Formative Research (<https://formative.jmir.org>), 28.10.2021. This is an open-access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work, first published in JMIR Formative Research, is properly cited. The complete bibliographic information, a link to the original publication on <https://formative.jmir.org>, as well as this copyright and license information must be included.