Distinguishing bipolar from unipolar depression: the importance of clinical symptoms and illness features

A. K. Leonpacher[1,†], **D. Liebers**[1,†], **M. Pirooznia**[1], **D. Jancic**[2], **D. F. MacKinnon**[1], **F. M. Mondimore**[1], **B. Schweizer**[1], **J. B. Potash**[2], **P. P. Zandi**[1,3], **NIMH Genetics Initiative Bipolar Disorder Consortium, GenRED Consortium**, and **F. S. Goes**[1,*]

[1]Department of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine, Baltimore, MD, USA

[2]Department of Psychiatry, Carver College of Medicine, University of Iowa, Iowa City, IA, USA

[3]Department of Mental Health, Johns Hopkins Bloomberg School of Public Health, Baltimore, MD, USA

Abstract

Background—Distinguishing bipolar disorder (BP) from major depressive disorder (MDD) has important relevance for prognosis and treatment. Prior studies have identified clinical features that differ between these two diseases but have been limited by heterogeneity and lack of replication. We sought to identify depression-related features that distinguish BP from MDD in large samples with replication.

Method—Using a large, opportunistically ascertained collection of subjects with BP and MDD we selected 34 depression-related clinical features to test across the diagnostic categories in an initial discovery dataset consisting of 1228 subjects (386 BPI, 158 BPII and 684 MDD). Features significantly associated with BP were tested in an independent sample of 1000 BPI cases and 1000 MDD cases for classifying ability in receiver operating characteristic (ROC) analysis.

Results—Seven clinical features showed significant association with BPI compared with MDD: delusions, psychomotor retardation, incapacitation, greater number of mixed symptoms, greater number of episodes, shorter episode length, and a history of experiencing a high after depression treatment. ROC analyses of a model including these seven factors showed significant evidence for discrimination between BPI and MDD in an independent dataset (area under the curve = 0.83). Only two features (number of mixed symptoms, and feeling high after an antidepressant) showed an association with BPII *versus* MDD.

---

*Address for correspondence: F. S. Goes, M.D., Department of Psychiatry and Behavioral Sciences, Johns Hopkins School of Medicine, Meyer 4-119A, 600 N. Wolfe Street, Baltimore, MD 21287, USA. (fgoes1@jhmi.edu).
†These authors contributed equally.

Declaration of Interest
None.

Conclusions—Our study suggests that clinical features distinguishing depression in BPI *versus* MDD have important classification potential for clinical practice, and should also be incorporated as 'baseline' features in the evaluation of novel diagnostic biomarkers.

Introduction

The syndrome of depression is a functionally debilitating condition common in both community and treatment settings (Murray *et al.* 2012). Although symptoms of depression are present in many psychiatric and somatic illnesses, depressive episodes are most prominent in bipolar disorder (BP) and in major depressive disorder (MDD), the two most common mood disorders that, together, afflict up to one-fifth of the world's population (Bromet *et al.* 2011; Merikangas *et al.* 2011). The distinction between bipolar and unipolar illness, first made decades ago, was originally based on differing patterns of external validators such as family history, sex and premorbid personality (Leonhard *et al.* 1962; Angst, 1966; Perris, 1966). The discovery of psychopharmacological agents, with their relative specificity for depressive or manic syndromes, largely supported this distinction (Pacchiarotti *et al.* 2013), and highlighted the need to minimize the misdiagnosis of MDD in patients with BP, since antidepressant treatment may tend to worsen mood stability in BP (Wehr & Goodwin, 1987; Altshuler *et al.* 1995; Henry *et al.* 2001). Nevertheless, misdiagnosis between BP and MDD depression remains common in both primary care and psychiatric clinics (Ghaemi *et al.* 1999, 2000; Angst *et al.* 2011).

A major challenge in the diagnosis of BP is the relative infrequency of episodes of mania and hypomania in comparison with the longer and more frequent periods of depression (Judd *et al.* 2003; Altshuler *et al.* 2010). In addition, most individuals diagnosed with BP experience the onset of their illness with a depressive rather than manic episode (Lish *et al.* 1994). In recent decades, studies have attempted to identify features of illness that help distinguish patients with bipolar depression from those with MDD. Several of these features have been consistently found to be more prominent in bipolar depression: earlier age at onset, increased number of depressive episodes, and greater propensity for hypersomnia, psychomotor abnormalities, and psychotic symptoms (Mitchell *et al.* 2001, 2011; Serretti *et al.* 2002; Perlis *et al.* 2006; Goes *et al.* 2007; Souery *et al.* 2012; Tondo *et al.* 2014). However, differing phenotype assessments and widely varying ascertainment schemes have limited the comparison of anything but broad trends across studies.

More recently, attention has also focused on the identification of biomarkers that could assist in differentiating BP from MDD depression. Structural and functional MRI studies have been performed, though their interpretation is limited by small sample size, differing protocols, and often divergent results (Cardoso de Almeida & Phillips, 2013). Similarly, biomarker studies, mainly focused on peripheral proteins such as brain-derived neurotrophic factor, have identified potential differences across the diagnoses, but these findings remain preliminary (Fernandes *et al.* 2009; Li *et al.* 2014). Finally, while robust genetic association

findings are emerging for BP, the modest effect sizes of both individual markers and their combinations limit their diagnostic utility at present (Psychiatric GWAS Consortium Bipolar Disorder Working Group, 2011). Hence, while biomarker studies will probably provide insights into illness pathophysiology, they are currently neither sufficiently robust nor sufficiently predictive to aid in the important differentiation between depressive episodes between BP compared with MDD.

Given the limited current availability of diagnostic biomarkers and the difficulties in interpreting across previous studies of differential diagnostic features, we sought to revisit the question of which symptoms, clinical characteristics, and co-morbidities may be of use in identifying patients with depression due to BP. We took advantage of a large dataset of subjects with BP and recurrent major depression diagnosed with substantively identical semi-structured interviews and best-estimate procedures, thus allowing for appropriate cross-diagnosis and cross-study comparability. We identify depression-related features specifically associated with BP and implement a simple predictive model that performs robustly in a fully independent dataset.

Method

Subjects

We analysed diagnostic and interview data from large BP [National Institute of Mental Health (NIMH) Genetics Initiative] and MDD (Genetics of Recurrent Early-Onset Depression; GenRED) collections, initially focused on ascertainment of samples for genetic studies. Both studies began by ascertaining families with at least two affected family members and transitioned to the collection of singleton subjects. Diagnoses for both studies were based on the Diagnostic Interview for Genetic Studies (DIGS) (Nurnberger *et al.* 1994), followed by best-estimate procedures. In this analysis, we utilized samples with high confidence BPI, BPII or MDD diagnoses from both familial and singleton samples in both the BP and MDD samples. For the family samples, only one affected member was used per family. To minimize potential effects from ascertainment bias, we excluded all probands from each family and used a randomly selected affected relative.

We split the BP and MDD samples into a discovery and replication dataset. For both samples, the discovery dataset consisted of an affected relative (one per family) from the family-based collection. To increase the number of BPII cases, we preferentially selected BPII relatives if they were available. The final number of subjects in the discovery dataset was: 386 BPI, 158 BPII and 684 MDD. The replication datasets were drawn from the later singleton collections of the BP and MDD studies. We randomly selected 1000 cases with BPI and 1000 cases with MDD.

Phenotype selection

We examined three categories of clinical features: (*a*) symptoms during the most severe depression; (*b*) lifetime clinical characteristics of depression; and (*c*) co-morbidities. Diagnoses were obtained from the best-estimate consensus, while symptoms were obtained directly from the DIGS interview (most severe depression section). Clinical characteristics

were obtained from the DIGS and the best-estimate interviews. In total, we tested 34 features, including both categorical and dimensional variables (Tables 1–3)

Analysis

We compared (*a*) symptoms, (*b*) clinical characteristics and (*c*) co-morbidities, between BP and MDD, using univariate and multivariate logistic regression. Significance testing was initially performed with nominal two-sided *p* values. We subsequently selected features that remained significant after correction for the 34 clinical variables examined (Bonferroni $p < 0.0014$). These features were included in a multivariate logistic regression model, controlling for age and sex. Features that remained significant in the full model were retained and this final model was then tested for its ability to distinguish cases of BP and MDD using a receiver operating characteristic (ROC) analysis in both the discovery and independent replication samples. In secondary analyses, we compared BPII subjects with MDD subjects. All analyses were conducted in Stata 12.1 (USA).

Results

Sociodemographic characteristics

Table 1 shows the major demographic features of the selected subjects by diagnosis, combining the discovery and independent target datasets. Of the 1228 participants initially analysed in the first dataset, 386 had a BPI diagnosis, 158 had a BPII diagnosis and 684 had an MDD diagnosis. Major demographic variables were consistent across the various diagnostic groups, with certain expected differences (Table 1): for example, BPI participants were more likely to be on disability (17.1%) compared with BPII and MDD depression participants (3.2% and 4.2%, respectively).

Distinguishing BPI from MDD

We divided the associated phenotypes into three categories: (1) symptoms during the most severe depressive episode; (2) lifetime clinical characteristics; and (3) co-morbidities. In the primary analysis of subjects with BPI *versus* those with MDD (Table 2), six symptoms showed evidence of association with BPI (using Bonferroni-corrected *p* value cut-off of 0.0014): psychomotor agitation [odds ratio (OR) 1.56, $p < 0.001$], psychomotor retardation (OR 2.51, $p < 0.001$), thoughts of self-harm (OR 2.11, $p < 0.001$), delusions (OR 7.72, $p < 0.001$), hallucinations (OR 4.75, $p < 0.001$) and functional incapacitation (OR 5.00, $p < 0.001$). Similarly, six lifetime clinical characteristics were also found to be strongly associated with BPI (Table 3): high after anti-depressive treatment (OR 6.24, $p < 0.001$), hospitalization for depression (OR 3.40, $p < 0.001$), greater number of lifetime depressive episodes (OR 1.02, $p < 0.001$), shorter most severe episode (OR 0.97, $p < 0.001$), greater number of mixed symptoms (defined as the count of seven potential mixed symptoms during the most severe depression) (OR 1.43, $p < 0.001$), and greater number of lifetime suicide attempts (OR 1.20, $p < 0.001$). The differences for the co-morbid diagnoses were less pronounced: a nominal association was seen for social phobia, but only the category of alcohol abuse and dependence was found to be significantly associated with BPI (OR 1.66, $p < 0.001$) after correcting for multiple testing.

We subsequently included all 13 associated features into a multiple logistic regression, and found that seven features remained significant in that model after Bonferroni correction (Table 4). These seven features were included in a final model and its ROCs were examined for their ability to distinguish BPI from MDD. ROC analysis of the discovery dataset showed, as expected, evidence for very good differentiation [AUC 0.84, 95% confidence interval (CI) 0.81– 0.87] (see online Supplementary Fig. S1). Importantly, very similar results were found when we tested these seven factors in an independent dataset of 1000 cases with BPI and 1000 cases with MDD. As shown in Fig. 1, the ROC analysis in the independent dataset had a significant AUC of 0.83 (95% CI 0.82–0.85), with a sensitivity of 75.6% and specificity of 77.8% at the point on the curve furthest from the null. To explore which features were driving the classification performance, we split the features into symptoms during a most severe depressive episode and those lifetime clinical characteristics. As shown in Fig. 1, the four clinical characteristics (AUC = 0.78, 95% CI 0.76–0.80) and the three symptoms (AUC = 0.74, 95% CI 0.72– 0.76) contributed about equally to the classification model.

Distinguishing BPII from MDD

In contrast to the BPI comparisons, there were very few differences in symptoms of depression between BPII and MDD, and none were significant after correction for multiple testing (see online Supplementary Table S1). Further, only two clinical characteristics were significant after correcting for multiple testing: the number of mixed symptoms during the most severe depression (OR 1.48, $p < 0.001$) and feeling 'high after depression treatment' (OR 3.41, $p < 0.001$) (see online Supplementary Table S2). Comparison of clinical co-morbidities between BPII and MDD showed a similar pattern as was observed between BPI and MDD, with only alcohol abuse being significantly associated (OR 2.21, $p < 0.001$). When these three features were examined in a multiple logistic regression model, only the two clinical characteristics remained significant (see online Supplementary Table S3). As expected, ROC analyses of these features showed a much more limited ability to discriminate between BPII and MDD (AUC = 0.63, 95% CI 0.57–0.70) (see online Supplementary Fig. S2). Since the AUC value was low in the discovery dataset, we did not seek to test the final model in the independent dataset.

Discussion

This study provided a broad evaluation of depressive symptoms and illness features that differ between subjects with bipolar and unipolar depressions. Seven such distinguishing features were found to be statistically significant after correction for multiple testing, showing a significant ability to discriminate between the BPI and MDD diagnoses (AUC = 0.83) when tested in an independent sample. In contrast, we found relatively few differences between features of depression in BPII compared with MDD, highlighting the potentially intermediary role that BPII may occupy across the MDD–BP spectrum.

The main differences between BPI and MDD were driven by symptoms during the most severe episodes and by lifetime clinical characteristics. Among the most important symptom differences (with ORs >2) were increased likelihood of psychomotor retardation, suicidal

behavior, psychotic symptoms and overall incapacitation. Importantly, they generally indicate a pattern of greater severity in bipolar depression, which is consistent with much of the published literature (Souery *et al.* 2012) albeit with some exceptions potentially due to differences in ascertainment and diagnosis. Comparison of other associated clinical characteristics showed a greater risk of hospitalization, a higher number of shorter depressive episodes, and a greater likelihood of experiencing a greater number of mixed symptoms during the most severe depression. In agreement with prior studies (Mitchell *et al.* 2001; Moreno *et al.* 2012), we found that BPI subjects experienced more hypersomnia (Table 2), although this difference did not survive correction for multiple testing in our study. A number of other important differences in symptoms and clinical characteristics were consistent with prior studies, which have shown BPI subjects to be more likely to experience mixed symptoms during a depression (Angst *et al.* 2011), psychomotor abnormalities, delusions (Mitchell *et al.* 2011) and overall impairment (Das *et al.* 2005). Both the significant effect sizes and the consistency with the prior literature suggest that these may be particularly important features and worthy of further study.

The strongest association with BPI in our sample was experiencing a 'high after depression treatment' (OR = 6.24, $p < 0.001$), which has been shown in a multinational, large-scale study to be one of the most strongly associated clinical features with a BP diagnosis (Angst *et al.* 2011). However, although pharmacologically induced switching of mood may be one of the most distinctive clinical features associated with BP, it is feature limited to subjects who have been treated and will not aid the clinician in distinguishing first-onset or untreated depressions.

Perhaps surprisingly, there were relatively few differences among co-morbid diagnoses, with only an increased rate of alcoholism being significantly associated with BPI (OR = 1.66, $p < 0.001$). This supports a similar finding by Souery *et al.* (2012), but differs from other reports that suggest that anxiety disorders and drug use may also be more prevalent in epidemiologically ascertained BPI patients (Moreno *et al.* 2012).

The comparison between depressions in BPII *versus* MDD yielded very few differences. The overall pattern was one of more subtle differences, with only two features (high after depression treatment and a greater number of mixed symptoms) remaining significant in the multivariate logistic regression models; consequently, there was limited discrimination seen in the ROC analysis (AUC = 0.63, 95% CI 0.57–0.70). Notably, the number of mixed symptoms during a most severe depression was also elevated in BPII (OR = 1.48, $p < 0.001$), consistent with a number of prior studies more focused on BPII (Benazzi, 2007). An important caveat is that our BPII sample size was less than half that for BPI, and thus our power to detect significant differences was substantially less. However, the ORs were much smaller in most of the BPII *versus* MDD comparisons, suggesting that our failure to detect significant results was largely driven by a lack of true differences.

Having identified associated features in our discovery sample, we subsequently performed ROC analysis in an independent sample, also diagnosed with the DIGS, but which focused on collection of singleton subjects rather than families. ROC analysis of the independent sample showed an essentially identical ability of the selected seven features to discriminate

between BPI and MDD (AUC = 0.83, (95% CI 0.82–0.85). Under optimal conditions (the point on the ROC curve most displaced from the null), the clinical features would have a sensitivity of 76% and a specificity of 78%. Although there is no specific AUC score cut-off that renders a test or diagnostic procedure clinically 'useful', values > 0.70 are usually recommended, with AUC scores > 0.9 being particularly desirable for tests that require particularly high accuracy (Swets, 1988). Our AUC value in the independent sample of 0.83 falls within a range that has been typically described as providing 'very good discrimination' in studies screening for psychopathology in the population (Kessler *et al.* 2003). As a comparison, this AUC is similar, if not slightly superior to widely used cardiac outcome predictor models, which range in AUC from 0.7 to 0.8 (Cook, 2012).

As research increasingly turns towards blood-based or imaging biomarkers, what role should such clinical findings play in decision-making? Because of their ready availability, clinically based markers should represent a baseline for prediction of diagnosis and/or illness course upon which biomarkers are tested. Hence, if the goal of a hypothetical biomarker were to help distinguish between two diagnoses, the important outcome of the biomarker study would be to determine the extent to which the use of a biomarker aids in classification over and above what is provided by the 'baseline' of clinical features (Kendler, 2014). For example, a recent volumetric imaging study of BP compared with MDD identified a number of gray matter differences, which yielded slightly inferior classification performance (depending on classifier models, sensitivity ranged from 66% to 76% with specificity of 59% to 73%) compared with the clinical results presented in this paper (Redlich *et al.* 2014). In such studies, it would be of interest to know what proportion of the classification performance may be indexed by the clinical features alone, and how much additional ability to classify is provided by the putative biomarker(s). Of course, biomarkers have the added potential of providing new mechanistic insights, but in their predictive role, they are probably most likely to be useful in combination with clinical features (Ioannidis & Tzoulaki, 2012).

An important limitation of our study, common to most previous studies, is the cross-sectional nature of the diagnoses. A more informative longitudinal design would have allowed us to test whether the identified illness features could predict the development of mania in a patient who initially presents with depression. Further limitations arising from the cross-sectional design include reliance on recalled symptoms and, for feasibility, a focus only on a single 'worst' depressive episode. A few prior studies have benefitted from a longitudinal design, but they have been also limited by the pragmatic need to follow fewer patients (Akiskal *et al.* 1995) or to perform a more limited phenotypic assessment (Tondo *et al.* 2014). At present, the feasibility of collecting sufficient information on a sufficiently large number of affected subjects remains a challenge for the field. A second important potential limitation might be potential ascertainment bias that may emerge from collecting cases for genetic studies. For example, one potential bias is that the GenRED sample specifically recruited early-onset cases of MDD (age of onset <31 years), which precludes an inquiry into whether age at onset could be used as a distinguishing feature between bipolar and unipolar patients as has been found in previous studies (Souery *et al.* 2012). To limit the potential for ascertainment for more severe cases, we excluded probands from the family dataset and, reassuringly, found few differences across the discovery and independent

datasets. Although population-based surveys may provide more external validity, they are by necessity limited to self-report questionnaires [such as in the National Epidemiologic Survey on Alcohol and Related Conditions (NESARC) study] (Moreno *et al.* 2012) or lay interviewer-based diagnoses (as used in the Epidemiologic Catchment Area study or the National Comorbidity Survey) (Robins & Regier, 1991; Kessler *et al.* 1994). In contrast, our diagnoses were confirmed by a well-validated interview with best-estimate diagnoses made by two supervising clinicians. Additionally, we included only diagnoses that were made with a high degree of confidence; however, one potential limitation of excluding low confidence diagnoses is that our results may be less informative to cases with more nuanced presentations. Finally, in an era when diagnostic boundaries are being increasingly called into question, our diagnostic instrument was based on Statistical Manual of Mental Disorders (DSM) criteria, which limits its use for exploration of alternative and more dimensional methods of classification.

In summary, our study, encompassing one of the largest collections of subjects with BP and MDD mood disorders, identified seven important clinical features that successfully distinguished BP from MDD patients, both in an initial and an independent dataset, which suggests that our result may have sufficient accuracy to be relevant for clinical use. Despite increasing emphasis on the discovery of novel neuroimaging and peripheral biomarkers, our study suggests that clinical features continue to have important classification potential that should not be ignored. Rather, these features should be integrated with biological markers in future studies aiming to predict diagnosis and course of illness.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Psychiatric Institute, New York, NY, Myrna M. Weissman, Ph.D. and James K. Knowles, M.D., Ph.D.; University of Pittsburgh, Pittsburgh, PA, George S. Zubenko, M.D., Ph.D. and Wendy N. Zubenko, Ed.D., R.N., C.S.; Johns Hopkins University, Baltimore, J. Raymond DePaulo, M.D., Melvin G. McInnis, M.D. and Dean MacKinnon, M.D.; University of Pennsylvania, Philadelphia, PA, Douglas F. Levinson, M.D. (GenRED coordinator), Madeleine M. Gladis, Ph.D., Kathleen Murphy-Eberenz, Ph.D. and Peter Holmans, Ph.D. (University of Wales College of Medicine); University of Iowa, Iowa City, IW, Raymond R. Crowe, M.D. and William H. Coryell, M.D.; Rush University Medical Center, Chicago, IL, William A. Scheftner, M.D. (Rush-Presbyterian).

References

Akiskal HS, Maser JD, Zeller PJ, Endicott J, Coryell W, Keller M, Warshaw M, Clayton P, Goodwin F. Switching from 'unipolar' to bipolar II. An 11-year prospective study of clinical and temperamental predictors in 559 patients. Archives of General Psychiatry. 1995; 52:114–123. [PubMed: 7848047]

Altshuler LL, Kupka RW, Hellemann G, Frye MA, Sugar CA, McElroy SL, Nolen WA, Grunze H, Leverich GS, Keck PE, Zermeno M, Post RM, Suppes T. Gender and depressive symptoms in 711 patients with bipolar disorder evaluated prospectively in the Stanley Foundation bipolar treatment outcome network. American Journal of Psychiatry. 2010; 167:708–715. [PubMed: 20231325]

Altshuler LL, Post RM, Leverich GS, Mikalauskas K, Rosoff A, Ackerman L. Antidepressant-induced mania and cycle acceleration: a controversy revisited. American Journal of Psychiatry. 1995; 152:1130–1138. [PubMed: 7625459]

Angst J. On the etiology and nosology of endogenous depressive psychoses. A genetic, sociologic and clinical study [article in German]. Monographien aus dem Gesamtgebiete der Neurologie und Psychiatrie. 1966; 112:1–118. [PubMed: 5943846]

Angst J, Azorin JM, Bowden CL, Perugi G, Vieta E, Gamma A, Young AH. BRIDGE Study Group. Prevalence and characteristics of undiagnosed bipolar disorders in patients with a major depressive episode: the BRIDGE study. Archives of General Psychiatry. 2011; 68:791–798. [PubMed: 21810644]

Benazzi F. Bipolar disorder – focus on bipolar II disorder and mixed depression. Lancet. 2007; 369:935–945. [PubMed: 17368155]

Bromet E, Andrade LH, Hwang I, Sampson NA, Alonso J, de Girolamo G, de Graaf R, Demyttenaere K, Hu C, Iwata N, Karam AN, Kaur J, Kostyuchenko S, Lépine J-P, Levinson D, Matschinger H, Mora MEM, Browne MO, Posada-Villa J, Viana MC, Williams DR, Kessler RC. Cross-national epidemiology of DSM-IV major depressive episode. BMC Medicine. 2011; 9:90. [PubMed: 21791035]

Cardoso de Almeida JR, Phillips ML. Distinguishing between unipolar depression and bipolar depression: current and future clinical and neuroimaging perspectives. Biological Psychiatry. 2013; 73:111–118. [PubMed: 22784485]

Cook NR. Clinically relevant measures of fit? A note of caution. American Journal of Epidemiology. 2012; 176:488–491. [PubMed: 22875759]

Das AK, Olfson M, Gameroff MJ, Pilowsky DJ, Lantigua R, Shea S, Weissman MM. Screening for bipolar disorder in a primary care practice. Journal of the American Medical Association. 2005; 293:956–963. [PubMed: 15728166]

Fernandes BS, Gama CS, Kauer-Sant'Anna M, Lobato MI, Belmonte-de-Abreu P, Kapczinski F. Serum brain-derived neurotrophic factor in bipolar and unipolar depression: a potential adjunctive tool for differential diagnosis. Journal of Psychiatric Research. 2009; 43:1200–1204. [PubMed: 19501841]

Ghaemi SN, Boiman EE, Goodwin FK. Diagnosing bipolar disorder and the effect of antidepressants: a naturalistic study. Journal of Clinical Psychiatry. 2000; 61:804–808. quiz 809. [PubMed: 11078046]

Ghaemi SN, Sachs GS, Chiou AM, Pandurangi AK, Goodwin K. Is bipolar disorder still underdiagnosed? Are antidepressants overutilized? Journal of Affective Disorders. 1999; 52:135–144. [PubMed: 10357026]

Goes FS, Sadler B, Toolan J, Zamoiski RD, Mondimore FM, Mackinnon DF, Schweizer B, Raymond Depaulo J, Potash JB. Psychotic features in bipolar and unipolar depression. Bipolar Disorders. 2007; 9:901–906. [PubMed: 18076541]

Henry C, Sorbara F, Lacoste J. Antidepressant-induced mania in bipolar patients: identification of risk factors. Journal of Clinical Psychiatry. 2001; 62:249–255. [PubMed: 11379838]

Ioannidis JP, Tzoulaki I. Minimal and null predictive effects for the most popular blood biomarkers of cardiovascular disease. Circulation Research. 2012; 110:658–662. [PubMed: 22383708]

Judd LL, Akiskal HS, Schettler PJ, Coryell W, Maser J, Rice JA, Solomon DA, Keller MB. The comparative clinical phenotype and long term longitudinal episode course of bipolar I and II: a clinical spectrum or distinct disorders? Journal of Affective Disorders. 2003; 73:19–32. [PubMed: 12507734]

Kendler KS. DSM issues: incorporation of biological tests, avoidance of reification, and an approach to the "Box Canyon Problem". American Journal of Psychiatry. 2014; 171:1248–1250. [PubMed: 25756765]

Kessler RC, Barker PR, Colpe LJ, Epstein JF, Gfroerer JC, Hiripi E, Howes MJ, Normand SL, Manderscheid RW, Walters EE, Zaslavsky AM. Screening for serious mental illness in the general population. Archives of General Psychiatry. 2003; 60:184–189. [PubMed: 12578436]

Kessler RC, McGonagle KA, Zhao S, Nelson CB, Hughes M, Eshleman S, Wittchen HU, Kendler KS. Lifetime and 12-month prevalence of DSM-III-R psychiatric disorders in the United States. Results from the National Comorbidity Survey. Archives of General Psychiatry. 1994; 51:8–19. [PubMed: 8279933]

Leonhard K, Korff I, Schulz H. Temperament in families with monopolar and bipolar phasic psychoses. Psychiatria et Neurologia. 1962; 143:416–434. [PubMed: 14464117]

Li Z, Zhang C, Fan J, Yuan C, Huang J, Chen J, Yi Z, Wang Z, Hong W, Wang Y, Lu W, Guan Y, Wu Z, Su Y, Cao L, Hu Y, Hao Y, Liu M, Yu S, Cui D, Xu L, Song Y, Fang Y. Brain-derived neurotrophic factor levels and bipolar disorder in patients in their first depressive episode: 3-year prospective longitudinal study. British Journal of Psychiatry: The Journal of Mental Science. 2014; 205:29–35.

Lish JD, Dime-Meenan S, Whybrow PC, Price RA, Hirschfeld RM. The National Depressive and Manic-depressive Association (DMDA) survey of bipolar members. Journal of Affective Disorders. 1994; 31:281–294. [PubMed: 7989643]

Merikangas KR, Jin R, He J-P, Kessler RC, Lee S, Sampson NA, Viana MC, Andrade LH, Hu C, Karam EG, Ladea M, Medina-Mora ME, Ono Y, Posada-Villa J, Sagar R, Wells JE, Zarkov Z. Prevalence and correlates of bipolar spectrum disorder in the world mental health survey initiative. Archives of General Psychiatry. 2011; 68:241–251. [PubMed: 21383262]

Mitchell PB, Frankland A, Hadzi-Pavlovic D, Roberts G, Corry J, Wright A, Loo CK, Breakspear M. Comparison of depressive episodes in bipolar disorder and in major depressive disorder within bipolar disorder pedigrees. British Journal of Psychiatry: The Journal of Mental Science. 2011; 199:303–309.

Mitchell PB, Wilhelm K, Parker G, Austin MP, Rutgers P, Malhi GS. The clinical features of bipolar depression: a comparison with matched major depressive disorder patients. Journal of Clinical Psychiatry. 2001; 62:212–216. quiz 217. [PubMed: 11305713]

Moreno C, Hasin DS, Arango C, Oquendo MA, Vieta E, Liu S, Grant BF, Blanco C. Depression in bipolar disorder versus major depressive disorder: results from the National Epidemiologic Survey on Alcohol and Related Conditions. Bipolar Disorders. 2012; 14:271–282. [PubMed: 22548900]

Murray CJL, Vos T, Lozano R, Naghavi M, Flaxman AD, Michaud C, Ezzati M, Shibuya K, Salomon JA, Abdalla S, Aboyans V, Abraham J, Ackerman I, Aggarwal R, Ahn SY, Ali MK, Alvarado M, Anderson HR, Anderson LM, Andrews KG, Atkinson C, Baddour LM, Bahalim AN, Barker-Collo S, Barrero LH, Bartels DH, Basáñez MG, Baxter A, Bell ML, Benjamin EJ, Bennett D, Bernabé E, Bhalla K, Bhandari B, Bikbov B, Bin Abdulhak A, Birbeck G, Black JA, Blencowe H, Blore JD, Blyth F, Bolliger I, Bonaventure A, Boufous S, Bourne R, Boussinesq M, Braithwaite T, Brayne C, Bridgett L, Brooker S, Brooks P, Brugha TS, Bryan-Hancock C, Bucello C, Buchbinder R, Buckle G, Budke CM, Burch M, Burney P, Burstein R, Calabria B, Campbell B, Canter CE, Carabin H, Carapetis J, Carmona L, Cella C, Charlson F, Chen H, Cheng AT, Chou D, Chugh SS, Coffeng LE, Colan SD, Colquhoun S, Colson KE, Condon J, Connor MD, Cooper LT, Corriere M, Cortinovis M, de Vaccaro KC, Couser W, Cowie BC, Criqui MH, Cross M, Dabhadkar KC, Dahiya M, Dahodwala N, Damsere-Derry J, Danaei G, Davis A, De Leo D, Degenhardt L, Dellavalle R, Delossantos A, Denenberg J, Derrett S, Des Jarlais DC, Dharmaratne SD, Dherani M, Diaz-Torne

C, Dolk H, Dorsey ER, Driscoll T, Duber H, Ebel B, Edmond K, Elbaz A, Ali SE, Erskine H, Erwin PJ, Espindola P, Ewoigbokhan SE, Farzadfar F, Feigin V, Felson DT, Ferrari A, Ferri CP, Fèvre EM, Finucane MM, Flaxman S, Flood L, Foreman K, Forouzanfar MH, Fowkes FG, Fransen M, Freeman MK, Gabbe BJ, Gabriel SE, Gakidou E, Ganatra HA, Garcia B, Gaspari F, Gillum RF, Gmel G, Gonzalez-Medina D, Gosselin R, Grainger R, Grant B, Groeger J, Guillemin F, Gunnell D, Gupta R, Haagsma J, Hagan H, Halasa YA, Hall W, Haring D, Haro JM, Harrison JE, Havmoeller R, Hay RJ, Higashi H, Hill C, Hoen B, Hoffman H, Hotez PJ, Hoy D, Huang JJ, Ibeanusi SE, Jacobsen KH, James SL, Jarvis D, Jasrasaria R, Jayaraman S, Johns N, Jonas JB, Karthikeyan G, Kassebaum N, Kawakami N, Keren A, Khoo JP, King CH, Knowlton LM, Kobusingye O, Koranteng A, Krishnamurthi R, Laden F, Lalloo R, Laslett LL, Lathlean T, Leasher JL, Lee YY, Leigh J, Levinson D, Lim SS, Limb E, Lin JK, Lipnick M, Lipshultz SE, Liu W, Loane M, Ohno SL, Lyons R, Mabweijano J, MacIntyre MF, Malekzadeh R, Mallinger L, Manivannan S, Marcenes W, March L, Margolis DJ, Marks GB, Marks R, Matsumori A, Matzopoulos R, Mayosi BM, McAnulty JH, McDermott MM, McGill N, McGrath J, Medina-Mora ME, Meltzer M, Mensah GA, Merriman TR, Meyer AC, Miglioli V, Miller M, Miller TR, Mitchell PB, Mock C, Mocumbi AO, Moffitt TE, Mokdad AA, Monasta L, Montico M, Moradi-Lakeh M, Moran A, Morawska L, Mori R, Murdoch ME, Mwaniki MK, Naidoo K, Nair MN, Naldi L, Narayan KM, Nelson PK, Nelson RG, Nevitt MC, Newton CR, Nolte S, Norman P, Norman R, O'Donnell M, O'Hanlon S, Olives C, Omer SB, Ortblad K, Osborne R, Ozgediz D, Page A, Pahari B, Pandian JD, Rivero AP, Patten SB, Pearce N, Padilla RP, Perez-Ruiz F, Perico N, Pesudovs K, Phillips D, Phillips MR, Pierce K, Pion S, Polanczyk GV, Polinder S, Pope CA 3rd, Popova S, Porrini E, Pourmalek F, Prince M, Pullan RL, Ramaiah KD, Ranganathan D, Razavi H, Regan M, Rehm JT, Rein DB, Remuzzi G, Richardson K, Rivara FP, Roberts T, Robinson C, De Leòn FR, Ronfani L, Room R, Rosenfeld LC, Rushton L, Sacco RL, Saha S, Sampson U, Sanchez-Riera L, Sanman E, Schwebel DC, Scott JG, Segui-Gomez M, Shahraz S, Shepard DS, Shin H, Shivakoti R, Singh D, Singh GM, Singh JA, Singleton J, Sleet DA, Sliwa K, Smith E, Smith JL, Stapelberg NJ, Steer A, Steiner T, Stolk WA, Stovner LJ, Sudfeld C, Syed S, Tamburlini G, Tavakkoli M, Taylor HR, Taylor JA, Taylor WJ, Thomas B, Thomson WM, Thurston GD, Tleyjeh IM, Tonelli M, Towbin JA, Truelsen T, Tsilimbaris MK, Ubeda C, Undurraga EA, van der Werf MJ, van Os J, Vavilala MS, Venketasubramanian N, Wang M, Wang W, Watt K, Weatherall DJ, Weinstock MA, Weintraub R, Weisskopf MG, Weissman MM, White RA, Whiteford H, Wiebe N, Wiersma ST, Wilkinson JD, Williams HC, Williams SR, Witt E, Wolfe F, Woolf AD, Wulf S, Yeh PH, Zaidi AK, Zheng ZJ, Zonies D, Lopez AD, AlMazroa MA, Memish ZA. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet. 2012; 380:2197–2223. [PubMed: 23245608]

Nurnberger JI, Blehar MC, Kaufmann CA, York-Cooler C, Simpson SG, Harkavy-Friedman J, Severe JB, Malaspina D, Reich T. Diagnostic Interview for Genetic Studies. Rationale, unique features, and training. NIMH Genetics Initiative. Archives of General Psychiatry. 1994; 51:849–859. discussion 863–864. [PubMed: 7944874]

Pacchiarotti I, Bond DJ, Baldessarini RJ, Nolen WA, Grunze H, Licht RW, Post RM, Berk M, Goodwin GM, Sachs GS, Tondo L, Findling RL, Youngstrom EA, Tohen M, Undurraga J, González-Pinto A, Goldberg JF, Yildiz A, Altshuler LL, Calabrese JR, Mitchell PB, Thase ME, Koukopoulos A, Colom F, Frye MA, Malhi GS, Fountoulakis KN, Vázquez G, Perlis RH, Ketter TA, Cassidy F, Akiskal H, Azorin J-M, Valentí M, Mazzei DH, Lafer B, Kato T, Mazzarini L, Martínez-Aran A, Parker G, Souery D, Ozerdem A, McElroy SL, Girardi P, Bauer M, Yatham LN, Zarate CA, Nierenberg AA, Birmaher B, Kanba S, El-Mallakh RS, Serretti A, Rihmer Z, Young AH, Kotzalidis GD, MacQueen GM, Bowden CL, Ghaemi SN, Lopez-Jaramillo C, Rybakowski J, Ha K, Perugi G, Kasper S, Amsterdam JD, Hirschfeld RM, Kapczinski F, Vieta E. The International Society for Bipolar Disorders (ISBD) task force report on antidepressant use in bipolar disorders. American Journal of Psychiatry. 2013; 170:1249–1262. [PubMed: 24030475]

Perlis RH, Brown E, Baker RW, Nierenberg AA. Clinical features of bipolar depression *versus* major depressive disorder in large multicenter trials. American Journal of Psychiatry. 2006; 163:225–231. [PubMed: 16449475]

Perris C. A study of bipolar (manic–depressive) and unipolar recurrent depressive psychoses. Acta Psychiatrica Scandinavica. 1966; 194:9–14. [PubMed: 5229234]

Psychiatric GWAS Consortium Bipolar Disorder Working Group. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. Nature Genetics. 2011; 43:977–983. [PubMed: 21926972]

Redlich R, Almeida JJR, Grotegerd D, Opel N, Kugel H, Heindel W, Arolt V, Phillips ML, Dannlowski U. Brain morphometric biomarkers distinguishing unipolar and bipolar depression: a voxel-based morphometry-pattern classification approach. JAMA Psychiatry. 2014; 71:1222–1230. [PubMed: 25188810]

Robins, LN., Regier, DA. Introduction. In: Robins, LN., Regier, DA., editors. Psychiatric Disorders in America: the Epidemiologic Catchment Area Study. The Free Press; New York: 1991. p. 1-10.

Serretti A, Mandelli L, Lattuada E, Cusin C, Smeraldi E. Clinical and demographic features of mood disorder subtypes. Psychiatry Research. 2002; 112:195–210. [PubMed: 12450629]

Souery D, Zaninotto L, Calati R, Linotte S, Mendlewicz J, Sentissi O, Serretti A. Depression across mood disorders: review and analysis in a clinical sample. Comprehensive Psychiatry. 2012; 53:24–38. [PubMed: 21414619]

Swets JA. Measuring the accuracy of diagnostic systems. Science. 1988; 240:1285–1293. [PubMed: 3287615]

Tondo L, Visioli C, Preti A, Baldessarini RJ. Bipolar disorders following initial depression: modeling predictive clinical factors. Journal of Affective Disorders. 2014; 167:44–49. [PubMed: 25082113]

Wehr TA, Goodwin FK. Can antidepressants cause mania and worsen the course of affective illness? American Journal of Psychiatry. 1987; 144:1403–1411. [PubMed: 3314536]
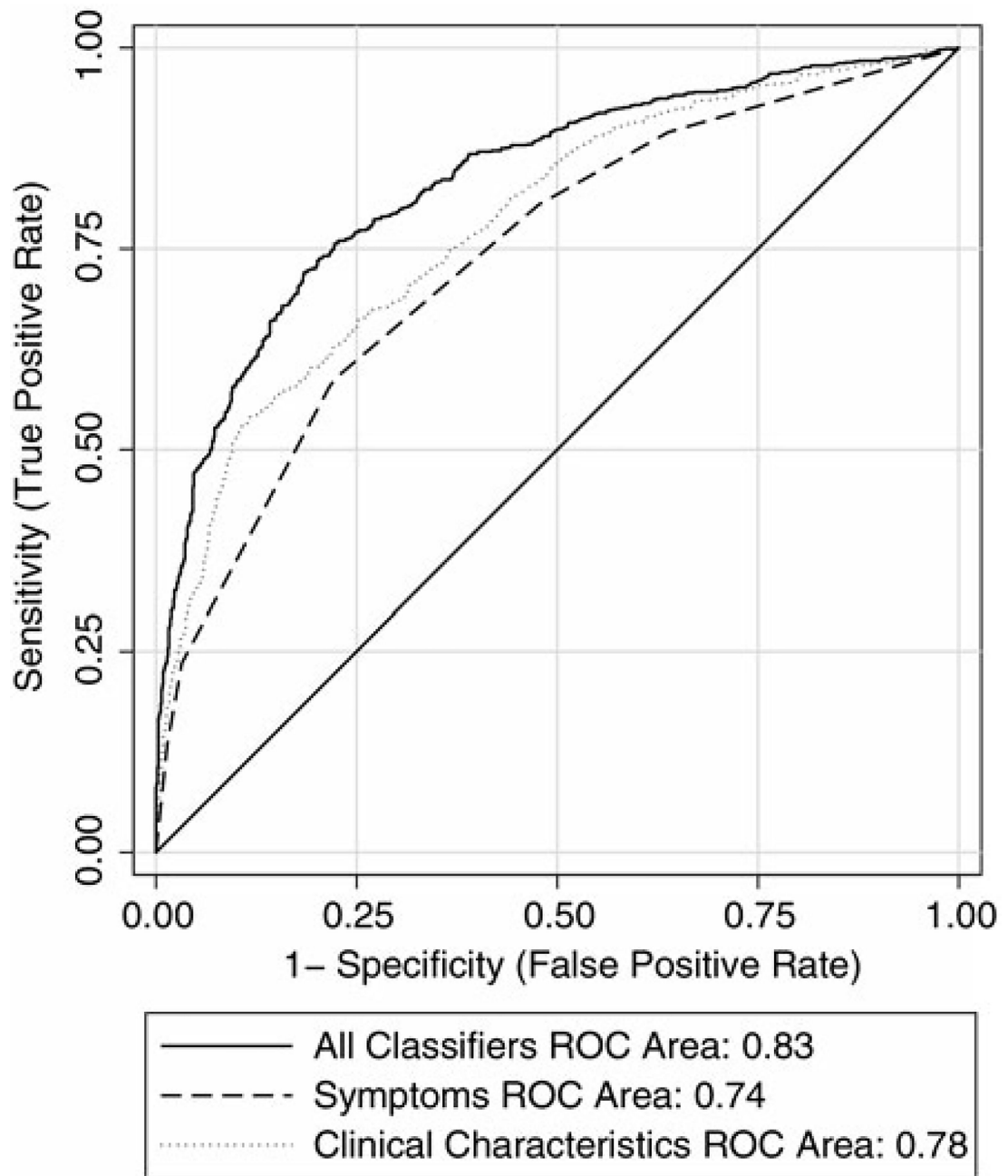
Fig. 1.
Receiver-operating characteristic (ROC) curves for seven clinical features criteria in an independent dataset.

Table 1

Sociodemographic characteristics of participants in BPI, BPII and MDD diagnosis groups encompassing both the discovery and target datasets

Characteristic	BPI (*n* = 1386)	BPII (*n* = 158)	MDD (*n* = 1684)
Mean age at interview, years (S.D.)	42.3 (12.1)	41.3 (14.4)	39.8 (12.8)
Mean age at most severe episode, years (S.D.)	30.9 (11.4)	30.8 (12.2)	29.1 (11.0)
Women, *n* (%)	886 (63.9)	104 (65.8)	1356 (80.6)
Married, *n* (%)	460 (33.2)	91 (58.0)	680 (40.4)
Disabled, *n* (%)	371 (26.8)	5 (3.2)	108 (6.4)
Mean duration of schooling, years (S.D.)	14.7 (2.6)	14.3 (3.1)	15.7 (2.8)

BP, Bipolar disorder; MDD, major depressive disorder; S.D., standard deviation.

Table 2

Prevalence of depressive symptoms during most severe episode in BPI and MDD samples in discovery dataset

Symptom	BPI, % (*n* = 386)	MDD, % (*n* = 684)	OR*a* (95% CI)	*p*
Decrease in appetite	58.0	53.2	1.35 (1.04–1.75)	0.025
Increase in appetite	18.4	22.2	0.85 (0.62–1.17)	0.321
Difficulty falling asleep	54.1	49.9	1.26 (0.97–1.65)	0.088
Early morning awakening	30.1	35.8	0.82 (0.62–1.08)	0.157
Oversleeping	56.5	49.4	1.43 (1.09–1.87)	0.009
Psychomotor agitation	46.6	36.8	1.56 (1.21–2.02)	<0.001 [*]
Psychomotor retardation	51.3	31.1	2.51 (1.93–3.26)	<0.001 [*]
Anhedonia	93.0	93.9	0.94 (0.54–1.62)	0.821
Fatigue	93.5	92.8	1.58 (0.88–2.84)	0.127
Feeling guilty	74.1	71.5	1.23 (0.92–1.65)	0.157
Feelings of worthlessness	86.0	84.2	1.31 (0.91–1.91)	0.148
Poor concentration	90.7	90.2	1.21 (0.76–1.93)	0.415
Passive death wish	71.5	64.5	1.46 (1.11–1.93)	0.007
Actually harmed self	23.3	13.3	2.11 (1.52–2.93)	<0.001 [*]
Worse in morning	27.8	27.6	1.05 (0.80–1.40)	0.706
Worse in evening	20.0	28.7	0.64 (0.47–0.87)	0.004
Delusions	22.0	3.8	7.72 (4.85–12.3)	<0.001 [*]
Hallucinations	10.4	2.5	4.75 (2.64–8.55)	<0.001 [*]
Incapacitated	72.5	36.0	5.00 (3.79–6.62)	<0.001 [*]

BPI, Bipolar disorder I; MDD, major depressive disorder; OR, odds ratio; CI, confidence interval.

*a*ORs controlled for age at interview and sex. ORs >1 indicate that a symptom is associated with greater likelihood of BPI diagnosis.

[*]Association *p* values that meet Bonferroni-corrected threshold of 0.0014.

Table 3

Clinical characteristics and co-morbidities of depressive episodes in discovery datasets

Feature	BPI (*n* = 386)	MDD (*n* = 684)	OR[a] (95% CI)	*p*
Clinical characteristic				
Sought professional treatment, %	81.0	80.9	1.01 (0.77–1.47)	0.718
Took medication, %	72.1	64.0	1.46 (1.15–1.92)	0.003
High after depression treatment, %	27.7	6.2	6.24 (4.20–9.28)	<0.001[*]
Hospitalized for depression, %	41.7	17.4	3.40 (2.56–4.52)	<0.001[*]
Mean number of lifetime depressive episodes (S.D.)	14.2 (36.5)	6.7 (11.9)	1.02 (1.01–1.03)	<0.001[*]
Mean number of suicide attempts (S.D.)	1.1 (4.3)	0.4 (0.89)	1.20 (1.09–1.32)	<0.001[*]
Mean duration of most severe episode, months (S.D.)	8.2 (13.5)	16.8 (31.2)	0.97 (0.95–0.98)	<0.001[*]
Mean number of mixed symptoms (S.D.)	1.2 (1.9)	0.5 (1.0)	1.43 (1.30–1.58)	<0.001[*]
Any mixed symptoms, %	37.1	29.6	1.45 (1.10–1.93)	0.009
Three or more mixed symptoms, %	21.0	4.4	6.05 (3.84–9.54)	<0.001[*]
Co-morbidity				
Alcoholism, %	35.2	23.8	1.66 (1.25–2.19)	<0.001[*]
Substance abuse, %	12.4	9.6	1.25 (0.84–1.87)	0.275
Panic disorder, %	26.9	24.4	1.20 (0.90–1.61)	0.207
Simple phobia, %	10.1	10.1	1.10 (0.73–1.68)	0.642
Social phobia, %	9.1	13.3	0.67 (0.44–1.01)	0.056
OCD, %	6.7	7.4	0.91 (0.56–1.49)	0.712
Anorexia/bulimia, %	6.2	5.8	1.22 (0.71–2.07)	0.468

BPI, Bipolar disorder I; MDD, major depressive disorder; OR, odds ratio; CI, confidence interval; S.D., standard deviation; OCD, obsessive–compulsive disorder.

[a] ORs controlled for age at interview and sex. ORs >1 indicate that a symptom is associated with greater likelihood of BPI diagnosis.

[*] Association *p* values that meet Bonferroni-corrected threshold of 0.0014.

Table 4

Multivariate logistic regression model for factors associated with BPI versus MDD in discovery dataset

Feature	Multivariate OR (95% CI)	P
High after depression treatment	5.92 (3.50–10.01)	<0.001 [*]
Number of suicide attempts	1.07 (0.92–1.24)	0.367
Actually harmed self	0.82 (0.44–1.50)	0.518
Hallucinations	2.24 (0.96–5.27)	0.063
Delusions	4.27 (2.16–8.45)	<0.001 [*]
Incapacitated	2.94 (1.87–4.60)	<0.001 [*]
Psychomotor agitation	1.21 (0.82–1.80)	0.332
Psychomotor retardation	1.62 (1.10–2.39)	0.015 [*]
Number of mixed symptoms	1.32 (1.15–1.52)	<0.001 [*]
Alcoholism	1.27 (0.83–1.94)	0.265
Hospitalized for depression	1.38 (0.83–2.28)	0.216
Number of lifetime depressive episodes	1.01 (1.00–1.02)	0.013 [*]
Most severe episode, months	0.95 (0.93–0.97)	<0.001 [*]

BPI, Bipolar disorder I; MDD, major depressive disorder; OR, odds ratio; CI, confidence interval.

[*] *p* Values meet Bonferroni threshold and were used as factors in the receiver-operating characteristic analysis.