Published in final edited form as:

J Affect Disord. 2016 March 1; 192: 219–225. doi:10.1016/j.jad.2015.12.053.

Individualized identification of euthymic bipolar disorder using the Cambridge Neuropsychological Test Automated Battery (CANTAB) and machine learning

Mon-Ju Wu^{1,*}, Ives Cavalcante Passos^{1,2,*}, Isabelle E. Bauer¹, Luca Lavagnino¹, Bo Cao¹, Giovana B. Zunta-Soares¹, Flávio Kapczinski^{1,2}, Benson Mwangi¹, and Jair C. Soares¹

¹UT Center of Excellence on Mood Disorder, Department of Psychiatry and Behavioral Sciences, The University of Texas Science Center at Houston, Houston, Texas, USA

²Bipolar Disorder Program and Laboratory of Molecular Psychiatry, Federal University of Rio Grande do Sul, Porto Alegre, RS, Brazil

Abstract

Background—Previous studies have reported that patients with bipolar disorder (BD) present with cognitive impairments during mood episodes as well as euthymic phase. However, it is still unknown whether reported neurocognitive abnormalities can objectively identify individual BD patients from healthy controls (HC).

Methods—A total of 21 euthymic BD patients and 21 demographically matched HC were included in the current study. Participants performed the computerized Cambridge Neurocognitive Test Automated Battery (CANTAB) to assess cognitive performance. The least absolute shrinkage selection operator (LASSO) machine learning algorithm was implemented to identify neurocognitive signatures to distinguish individual BD patients from HC.

Results—The LASSO machine learning algorithm identified individual BD patients from HC with an accuracy of 71%, area under receiver operating characteristic curve of 0.7143 and significant at $p = 0.0053$. The LASSO algorithm assigned individual subjects with a probability score (0 – healthy, 1 – patient). Patients with rapid cycling (RC) were assigned increased probability scores as compared to patients without RC. A multivariate pattern of neurocognitive abnormalities comprising of affective Go/No-go and the Cambridge gambling task was relevant in distinguishing individual patients from HC.

Limitations—Our study sample was small as we only considered euthymic BD patients and demographically matched HC.

Corresponding author: Benson Mwangi, PhD, Department of Psychiatry & Behavioral Sciences, The University of Texas Health Science Center, 1941 East Road, Houston, TX, 77054, USA, Phone: +17134862686, Fax: +17134862553, benson.mwangi@uth.tmc.edu.

*These authors contributed equally to this work.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Conclusion—Neurocognitive abnormalities can distinguish individual euthymic BD patients from HC with relatively high accuracy. In addition, patients with RC had more cognitive impairments compared to patients without RC. The predictive neurocognitive signature identified in the current study can potentially be used to provide individualized clinical inferences on BD patients.

Keywords

bipolar disorder; neurocognition; rapid cycling; CANTAB; machine learning

Introduction

Bipolar disorder (BD) has a lifetime prevalence of 2.1% worldwide, with subthreshold forms affecting another 2.4% (Merikangas et al., 2007). According to the World Health Organization (WHO), BD is among the top 10 leading causes of disability-adjusted life years (DALY) in young adults (Mathers et al., 2006). Furthermore, the rates of completed suicide among patients with BD are 7.8% in men and 4.9% in women (Nordentoft et al., 2011). It is known that the diagnosis of BD has an average delay of ten years between the first symptoms and the formal diagnosis (Lish et al., 1994). Moreover, only 20% of the patients with BD who are experiencing a depressive episode are diagnosed with BD within the first year of seeking treatment (Goldberg et al., 2001). This suggests that the current diagnostic system, which is based on subjective clinical judgments, lacks the required sensitivity and specificity to identify individual patients with BD.

It has been reported that BD is associated with cognitive impairments during mood episodes and euthymic phase of the disorder. In particular, the primary deficits observed during euthymia pertain to executive control, sustained attention, verbal learning, and working memory (Arts et al., 2008; Bora et al., 2009; Bourne et al., 2013; Mann-Wrobel et al., 2011; Robinson and Ferrier, 2006; Torres et al., 2010, 2007). A positive association between neurocognitive dysfunction and functional impairment has also been shown in both cross-sectional (Dickerson et al., 2004; Laes and Sponheim, 2006; Martínez-Arán et al., 2004; Zubieta et al., 2001) and longitudinal studies (Bonnín et al., 2010; Jaeger et al., 2007; Martino et al., 2009; Tabarés-Seisdedos et al., 2008). Furthermore, findings from a recent systematic review indicated an association between previous number of mood episodes and abnormalities in neurocognitive functioning - suggesting that an increasing number of mood episodes might be related to progressive neurocognitive decline (Robinson and Ferrier, 2006).

While previous studies have undoubtedly advanced our understanding of neurocognitive abnormalities in BD – they have largely established average group-level neurocognitive differences between BD patients and healthy controls (HC). However, it is still unknown whether reported neurocognitive abnormalities can objectively identify individual BD patients from demographically matched HC with a high prediction accuracy (specificity and sensitivity). Consequently, in the current study we set out to investigate whether an advanced machine learning algorithm ‘trained’ using the Cambridge Neuropsychological Test Automated Battery (CANTAB) can objectively identify individual euthymic BD

patients from HC. The CANTAB is a validated and standardized tool for assessment of cognitive functioning (Levaux et al., 2007; Sweeney et al., 2000).

Machine learning algorithms have recently gained traction in psychiatric research in establishing predictive biomarkers. In particular, several machine learning studies have reported above chance level (>50%) prediction accuracy in distinguishing BD patients from healthy individuals. For example, two studies (Schnack et al., 2014) and (Redlich et al., 2014) used structural gray matter density measurements coupled with a support vector machine (SVM) algorithm to distinguish individual BD patients from HCs with 67% and 79.3% accuracy respectively. In two subsequent studies, Mwangi and colleagues (Mwangi et al., 2014; Mwangi et al., 2015) used neuroanatomical signatures of the amygdala and white matter connectivity pathways to distinguish individual children and adolescents with BD as compared to healthy youths with 78% and 78% accuracy respectively. Studies using functional imaging methods coupled with machine learning to distinguish BD patients from HC have also been reported with prediction accuracy ranging from 79 – 89% (Arribas et al., 2010; Costafreda et al., 2011; Redlich et al., 2014). Gene expression data coupled with demographic and clinical data analyzed using a machine learning algorithm reported excellent prediction accuracy (92%) in distinguishing individual BD patients from HC. (Struyf et al., 2008). Recently, Koutsouleris and colleagues (Koutsouleris et al., 2012) used a comprehensive neuropsychological test battery to categorize at-risk mental state individuals from HC with 94% accuracy. However, to the best of our knowledge, this is the first study to utilize the CANTAB neurocognitive battery to distinguish BD patients from HC at an individual subject level. Consequently, in the current study, we used the CANTAB neurocognitive battery coupled with the least absolute shrinkage selection operator (LASSO) machine learning algorithm (Robert Tibshirani, 2011) to distinguish individual BD patients from HC. We postulated that a machine learning algorithm is ideal in identifying a predictive neurocognitive signature of BD for the following three reasons. First, machine learning algorithms are multivariate – meaning they take into account interactions between multiple variables as compared to univariate methods. Second, machine learning algorithms utilize a ‘cross-validation’ approach which facilitates separation of observations into ‘training’ and ‘testing’ samples and subsequently allow a researcher to establish generalizability of observed findings. Third, these algorithms allow predictions at an individual subject level, which may facilitate personalized clinical inferences.

Consequently, the three main objectives of this study were to 1) Examine the utility of neurocognitive measurements from the CANTAB battery together with a machine learning algorithm in distinguishing individual euthymic BD patients from HC; 2) Elucidate neurocognitive characteristics and patterns ‘most relevant’ in identifying individual euthymic patients from HC. 3) Investigate associations between individual subjects' machine learning predicted probability scores and rapid cycling (RC). Of note, RC is the presence of at least four mood episodes in the previous 12 months according to the diagnostic and statistical manual (DSM) (American Psychiatric Association, 2000). It has been reported that patients with RC have poor clinical outcomes (Kupka et al., 2005) and potentially more cognitive impairments (Post and Kalivas, 2013; Post et al., 2012). In particular, based on previous evidence of the progressive decline in cognitive functioning in BD we

hypothesized those individual BD patients assigned higher probabilities by the machine learning algorithm would present with rapid cycling.

Methods

The study was approved by the University of North Carolina at Chapel Hill Institutional Review Board. All subjects signed informed consent before any study-related procedures were performed. This process involved a complete description of the study with ample time for questions. Participants were recruited between January 2006 and June 2010.

Participants

A total of 21 euthymic BD patients and 21 demographically matched controls were recruited from the community and psychiatric clinics through flyers, radio, and newspaper advertisements. Inclusion criteria were euthymic subjects with BD types I or II according to DSM-IV, and aged between 18 and 65 years. Exclusion criteria were head trauma with residual effects, neurological disorder, and uncontrolled major medical conditions. Clinical and demographic details are summarized in Table 1.

Assessments

Subjects were evaluated through a socio-demographic history form to assess age, gender, years of education, and occupational status. Axis-I diagnoses and clinical characteristics were assessed using the Structured Clinical Interview for DSM-IV axis-I Disorders (SCID-I) (American Psychiatric Association, 2000), which was administered by fully trained staff. In addition, we used SCID to assess psychiatric history, current mood status, number of mood episodes (depressive and manic/hypomanic). Patients who did not meet criteria for manic, hypomanic, or depressed through a SCID interview were regarded as euthymic. Current dimensional mood symptoms were assessed with the Hamilton Depression Rating Scale (HAMD) (Hamilton, 1960), Montgomery-Åsberg depression rating scale (MADRS) (Montgomery and Asberg, 1979), and the Young Mania Rating Scale (YMRS) (Young et al., 1978).

Cognitive functioning

Participants performed the computerized Cambridge Neurocognitive Test Automated Battery (CANTAB - <http://www.cantab.com>). This cognitive battery was chosen based on the established sensitivity to cognitive impairments in psychiatric disorders (Sweeney et al., 2000) and evidence of the predictive power of the CANTAB variables for psychosocial functioning in individuals with psychotic symptoms (Levaux et al., 2007). Within CANTAB assessments, participants completed tests evaluating visuomotor and cognitive processing speed (Choice Reaction Time – CRT, Motor Screening – MOT, Big/Little Circle – BLC), attention (Match to Sample Visual Search – MTS, Rapid Visual Processing – RVP), working memory (Spatial Span task - SSP), decision-making, cognitive and affective control (Cambridge Gambling Task – CGT, Affective Go/No-Go-AGN), and spatial memory (Spatial Recognition Memory – SRM). Variables of interest across tasks included reaction times and accuracy (number of correct responses, commission and omission errors). For the SSP task we calculated a score accounting for both speed and accuracy (SSP quality score).

This score was calculated by dividing the mean time to the last response given by the participant divided by the total number of attempts. These neurocognitive tasks are briefly described in Table 2 and a detailed description is given elsewhere (Robbins et al., 1998).

Machine learning data analysis

As detailed in Table S2 in the supplementary material, a total of 37 CANTAB cognitive scores were included in the current study as predictor variables. These scores were selected based on previous literature in BD showing important deficits in the domains of learning and memory, processing speed, decision making, set-shifting and inhibition (Arts et al., 2008; Bora et al., 2009; Goodwin et al., 2008; Robinson et al., 2006). The AGN scores were included based on previous evidence of difficulties with affect processing and facial emotion recognition in individuals with BD (David et al., 2014; Sepede et al., 2015). These variables were normalized through a Z-transform by subtracting the mean and dividing with the standard deviation. Z-scores where higher numbers represent poor cognitive performance (e.g. reaction time) were multiplied by -1 to allow interpretability of results as also explored in details by Gur and colleagues (Gur et al., 2014). Z-transformed predictor variables together with corresponding target labels (1 – euthymic bipolar disorder, 2 – healthy control) input into the LASSO machine learning algorithm, which was ‘trained’ to distinguish individual patients from HC. Briefly, predictor variables were represented as x_{ij} and target labels as y_i , where $i = 1, 2, 3, \dots N$ and $j = 1, 2, 3, \dots M$ represent subjects and predictor variables or CANTAB measurements, respectively. As a result, the LASSO algorithm (Tibshirani 2011, Tibshirani 1996) was ‘trained’ to compute a set of coefficients $\hat{\beta}$ by minimizing this objective function

$$\sum_{i=1}^N \left(y_i - \sum_{j=1}^M x_{ij} \beta_j \right)^2 + \lambda \sum_{j=1}^M |\beta_j|$$

Briefly, the parameter λ was used to encourage the resulting predictive model to be ‘sparse’ – translating into few model coefficients ($\hat{\beta}$) with non-zero values, which would lead to a higher prediction accuracy and better model interpretability (Lavagnino et al., 2015; Robert Tibshirani, 2011). In this study, λ was selected using ‘training’ data only through a 10-fold cross-validation approach and the λ parameter maximizing prediction accuracy (high specificity and sensitivity) within the 10-fold was selected. Importantly, the λ parameter was selected using ‘training’ data only to avoid ‘circularity’ or ‘double-dipping’ as discussed elsewhere (Kriegeskorte et al., 2009; Lavagnino et al., 2015; Mwangi et al., 2013, 2012). The LASSO objective function was minimized through the coordinate descent algorithm using a MATLAB (The Mathworks, Inc) routine freely provided by Friedman and colleagues (Friedman et al., 2010). A generalized linear model was used to estimate a probability score of an individual subject belonging to either BD or HC group given subjects' CANTAB measurements and LASSO coefficients ($\hat{\beta}$). A probability score of zero represented a perfect prediction of a subject as a HC while on the contrary a probability score of unity represented a perfect BD prediction. A leave-one-out cross-validation (LOOCV) approach was used to evaluate the performance of the algorithm, which entailed

training the algorithm with all subjects but one and testing the algorithm with the ‘left-out’ subject as shown in Figure 1 and described elsewhere (Klöppel et al., 2012; Lavagnino et al., 2015; Mwangi et al., 2015; Wu et al., 2015). The algorithm's ability to identify previously ‘unseen’ subjects was assessed using standard statistical metrics such as accuracy, sensitivity, specificity, positive predictive value, negative predictive value and area under receive operating characteristic curve (AUC) as described in details elsewhere (Savitz et al., 2013). Importantly, to determine significance of prediction accuracy a chi-square test of subjects' actual diagnostic labels and predicted labels was also performed. Neurocognitive variables that were most relevant in distinguishing euthymic BD patients from HC and assigned non-zero coefficients were also examined and reported.

Relationship between a history of RC and predicted probability scores

All subjects were assigned a probability (0 – healthy, 1 – BD). A post hoc statistical test was performed to examine the null hypothesis of no difference in predicted patients' probability scores between BD patients with a history of RC (BD with RC) and BD patients without a history of RC (BD without RC). We hypothesized that patients with a history of RC (BD with RC) will be assigned significantly higher probability scores as compared to non-RC counterparts –as an indication of more cognitive impairments in patients with RC.

Results

Table 1 summarizes study clinical and demographic characteristics. BD patients and HC did not differ significantly based on gender, age and education as shown in Table 1. The LASSO algorithm identified individual patients from HC with 71% accuracy, 76% sensitivity, 67% specificity, 70% positive predictive value, 74% negative predictive value, and an area under receiver operating characteristic curve (ROC) of 0.7143. These predictions were significantly different from random (50 %) with a χ^2 p-value = 0.0053. A ‘confusion matrix’ a ROC curve are shown in Figure 2. A full list of predicted probability scores is shown in table S1 of supplementary materials.

Predictor variables most relevant in identifying patients from HC and assigned top non-zero coefficients by LASSO included the number of commission and omission errors to negative stimuli and the number of omission errors to positive stimuli on the AGN task, the mean time to the last response to total number of attempts ratio on the SSP task, and the quality of decision making index for the CGT task. These predictor variables and their corresponding coefficients are shown in Figure 3. A full list of coefficients of predictor variables is given in Table S2 of supplementary materials.

As shown in Figure 4, predicted probability scores between BD patients with RC and HC differed significantly (Kruskal-Wallis test $p = 0.01$). In contrast, BD patients without RC were less distinguishable from HC.

Discussion

Although previous studies have extensively reported neurocognitive differences between patients with BD and HC – these studies have largely established average group-level

differences (Robinson and Ferrier, 2006). Therefore, synthesizing these cognitive assessments into a predictive neurocognitive signature able to assist in BD diagnosis and treatment management has been an elusive goal up to now. This is the first study to examine the feasibility of individualized identification of BD patients from HC using multiple neurocognitive assessments coupled with an advanced machine learning technique. The LASSO algorithm identified individual patients from HC with 71% accuracy and an AUC of 0.7143. Most notably, the prediction accuracy was established using a cross-validation approach, which allowed testing of the algorithm's ability to predict previously 'unseen' subjects' diagnostic labels. The main outcome of the machine learning algorithm was a probability score able to quantify the likelihood of a subject being a patient with BD.

From a cognitive viewpoint, using CANTAB variables considered in this study, the strongest predictor variables were the number of commission and omission errors to negative stimuli and the number of omission errors to positive stimuli on the AGN task, the SSP quality score, and the quality of decision making index for the CGT task. The most compelling finding was that, compared to HC, individuals with BD committed a greater number of commission errors when exposed to negative stimuli. This finding is in line with literature showing the impairing effects of affective, more in particular negative stimuli, on cognitive processing in both patients with BD and their relatives (Bauer, I.E., et al., 2015; Passarotti et al., 2012, 2011; Pavuluri and Passarotti, 2008). All these findings suggested that cognitive deficits and affective processing biases might be interrelated and possibly constitute markers of vulnerability to BD. Further, the elevated number of commission errors to positive stimuli, and not to negative stimuli, supports previous hypotheses that individuals with BD may have difficulty disengaging their attention from negative material (Gotlib and Joormann, 2010). Additionally, patients with BD displayed a more impulsive profile compared to HC, characterized by reduced thinking time and an excessive number of attempts on the SSP task. This finding is in line with previous findings showing that euthymic BD patients display pronounced executive deficits and impulsivity traits (Christodoulou et al., 2006; Najt et al., 2007; Olley et al., 2005; Robinson et al., 2006; Swann et al., 2009a; Thompson et al., 2005).

Further, HC had a higher quality of decision making index on the CGT task compared with individuals with BD (Quraishi and Frangou, 2002). This index measures the total number of times that the subject chooses to bet on a stimulus leading to a likely winning outcome and is traditionally considered as an estimate of impulsivity and risk taking. Our finding is consistent with previous evidence that patients with BD have a high reward-seeking response and are unable to delay gratification (Najt et al., 2007; Swann et al., 2009b). Given that our sample was medicated and included participants with BD I and II it is noteworthy to mention that impulsive tendencies are observed across bipolar phases and appear to be independent from the patients' medication status (Swann et al., 2003).

BD with RC was associated with increased predicted probability scores, which indicated that they presented more cognitive impairment compared to patients without RC. Importantly, BD patients with RC and BD patients without RC did not differ significantly on age, gender, and education. Subjects with this specifier have previously been associated with poor clinical outcomes, such as treatment refractoriness, comorbid substance use disorder, and

suicide attempts (Berk et al., 2011; Carvalho et al., 2014; da Costa et al., 2015). For instance, RC has been associated with less response to lithium and carbamazepine (Post et al., 2012). However, to the best of our knowledge, this is the first study to show that patients with BD and RC have more cognitive impairment compared to those without.

The current study has some potential limitations. First, our study sample was small as we only considered euthymic BD patients and demographically matched HC. Second, although the validity of the current predictive algorithm was established using a leave-one-out cross-validation approach (Klöppel et al., 2012; Mwangi et al., 2015; Wu et al., 2015), this did not entail using a fully independent testing cohort. Work is ongoing in our group to acquire an independent sample and validate the current predictive model. However, previous machine learning studies in psychiatry have shown that a cross-validation approach is a valid mechanism of establishing algorithm generalization to ‘novel’ data as presented in the current study. In addition, a majority of subjects were taking medication, thus making it challenging to disentangle current findings from medications. Future studies should therefore replicate these findings using medication naïve patients. However, Swann and colleagues (Swann et al., 2003) observed impulsive behaviors across BD phases regardless of medication status which further highlights that our observations are potentially not due to medications. Notwithstanding these limitations, the current study serves as a proof-of-concept and future individualized machine learning studies will focus on other clinically relevant questions such as identification of individual subjects vulnerable to BD and prediction of response to therapeutic interventions. However, being able to distinguish individual BD patients from HC with significant prediction accuracy is an important first step.

In summary, we report an accurate prediction algorithm able to distinguish individual patients with BD from HC using the CANTAB neurocognitive battery. These results suggest that neurocognitive abnormalities can distinguish individual euthymic BD patients from HC and could potentially be used clinically to provide clinical inferences on individual patients. In conclusion, we hypothesize that integration of multi-modal markers (e.g. neurocognition, genetics, blood markers, neuroimaging) through a machine learning algorithm could potentially lead to improved prediction accuracy as presented in the current study (Kapczinski and Passos, 2015).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported in part by NIMH Grant R01 MH 085667, the Dunn Foundation, and the Pat Rutherford, Jr. Endowed Chair in Psychiatry to JCS.

References

American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th., text rev. 2000.

- Arribas JJ, Calhoun VD, Adali T. Automatic Bayesian classification of healthy controls, bipolar disorder, and schizophrenia using intrinsic connectivity maps from fMRI data. *IEEE Trans Biomed Eng.* 2010; 57:2850–60.10.1109/TBME.2010.2080679 [PubMed: 20876002]
- Arts B, Jabben N, Krabbendam L, van Os J. Meta-analyses of cognitive functioning in euthymic bipolar patients and their first-degree relatives. *Psychol Med.* 2008; 38:771–85.10.1017/S0033291707001675 [PubMed: 17922938]
- Bauer IE, Frazier TW, Meyer TD, Youngstrom E, Zunta-Soares G, Soares JC. Affective Processing in Pediatric Bipolar Disorder and Offspring of Bipolar Parents. *J Child Adolesc Psychopharmacol.* 2015
- Berk M, Kapczinski F, Andreazza AC, Dean OM, Giorlando F, Maes M, Yücel M, Gama CS, Dodd S, Dean B, Magalhães PVS, Amminger P, McGorry P, Malhi GS. Pathways underlying neuroprogression in bipolar disorder: focus on inflammation, oxidative stress and neurotrophic factors. *Neurosci Biobehav Rev.* 2011; 35:804–17.10.1016/j.neubiorev.2010.10.001 [PubMed: 20934453]
- Bonnín CM, Martínez-Arán A, Torrent C, Pacchiarotti I, Rosa AR, Franco C, Murru A, Sanchez-Moreno J, Vieta E. Clinical and neurocognitive predictors of functional outcome in bipolar euthymic patients: a long-term, follow-up study. *J Affect Disord.* 2010; 121:156–60.10.1016/j.jad.2009.05.014 [PubMed: 19505727]
- Bora E, Yucel M, Pantelis C. Cognitive endophenotypes of bipolar disorder: A meta-analysis of neuropsychological deficits in euthymic patients and their first-degree relatives. *J Affect Disord.* 2009; 113:1–20.10.1016/j.jad.2008.06.009 [PubMed: 18684514]
- Bourne C, Aydemir Ö, Balanzá-Martínez V, Bora E, Brissos S, Cavanagh JTO, Clark L, Cubukcuoglu Z, Dias VV, Dittmann S, Ferrier IN, Fleck DE, Frangou S, Gallagher P, Jones L, Kieseppä T, Martínez-Arán A, Melle I, Moore PB, Mur M, Pfennig A, Raust A, Senturk V, Simonsen C, Smith DJ, Bio DS, Soeiro-de-Souza MG, Stoddart SDR, Sundet K, Szöke A, Thompson JM, Torrent C, Zalla T, Craddock N, Andreassen OA, Leboyer M, Vieta E, Bauer M, Worhunsky PD, Tzagarakis C, Rogers RD, Geddes JR, Goodwin GM. Neuropsychological testing of cognitive impairment in euthymic bipolar disorder: an individual patient data meta-analysis. *Acta Psychiatr Scand.* 2013; 128:149–62.10.1111/acps.12133 [PubMed: 23617548]
- Carvalho AF, Dimellis D, Gonda X, Vieta E, McIntyre RS, Fountoulakis KN. Rapid cycling in bipolar disorder: a systematic review. *J Clin Psychiatry.* 2014; 75:e578–86.10.4088/JCP.13r08905 [PubMed: 25004199]
- Christodoulou T, Lewis M, Ploubidis GB, Frangou S. The relationship of impulsivity to response inhibition and decision-making in remitted patients with bipolar disorder. *Eur Psychiatry.* 2006; 21:270–3.10.1016/j.eurpsy.2006.04.006 [PubMed: 16762532]
- Costafreda SG, Fu CHY, Picchioni M, Touloupoulou T, McDonald C, Kravariti E, Walshe M, Prata D, Murray RM, McGuire PK. Pattern of neural responses to verbal fluency shows diagnostic specificity for schizophrenia and bipolar disorder. *BMC Psychiatry.* 2011; 11:18.10.1186/1471-244X-11-18 [PubMed: 21276242]
- Da Costa SC, Passos IC, Lowri C, Soares JC, Kapczinski F. Refractory bipolar disorder and neuroprogression. *Prog Neuropsychopharmacol Biol Psychiatry.* 2015;10.1016/j.pnpbp.2015.09.005
- David DP, Soeiro-De-Souza MG, Moreno RA, Bio DS. Facial emotion recognition and its correlation with executive functions in bipolar i patients and healthy controls. *J Affect Disord.* 2014;152–154. 288–294.10.1016/j.jad.2013.09.027 [PubMed: 25618002]
- Dickerson F, Boronow JJ, Stallings C, Origoni AE, Cole SK, Yolken RH. Cognitive functioning in schizophrenia and bipolar disorder: comparison of performance on the Repeatable Battery for the Assessment of Neuropsychological Status. *Psychiatry Res.* 2004; 129:45–53.10.1016/j.psychres.2004.07.002 [PubMed: 15572184]
- Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw.* 2010; 33:1–22. [PubMed: 20808728]
- Goldberg JF, Harrow M, Whiteside JE. Risk for bipolar illness in patients initially hospitalized for unipolar depression. *Am J Psychiatry.* 2001; 158:1265–70. [PubMed: 11481161]
- Goodwin GM, Anderson I, Arango C, Bowden CL, Henry C, Mitchell PB, Nolen WA, Vieta E, Wittchen HU. 2008. ECNP consensus meeting. Bipolar depression. Nice. Eur

- Neuropsychopharmacol. Mar.2007 18:535–49.10.1016/j.euroneuro.2008.03.003 [PubMed: 18501566]
- Gotlib IH, Joormann J. Cognition and depression: current status and future directions. *Annu Rev Clin Psychol.* 2010; 6:285–312.10.1146/annurev.clinpsy.121208.131305 [PubMed: 20192795]
- Gur RC, Calkins ME, Satterthwaite TD, Ruparel K, Bilker WB, Moore TM, Savitt AP, Hakonarson H, Gur RE. Neurocognitive growth charting in psychosis spectrum youths. *JAMA psychiatry.* 2014; 71:366–74.10.1001/jamapsychiatry.2013.4190 [PubMed: 24499990]
- Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry.* 1960; 23:56–62. [PubMed: 14399272]
- Jaeger J, Berns S, Loftus S, Gonzalez C, Czobor P. Neurocognitive test performance predicts functional recovery from acute exacerbation leading to hospitalization in bipolar disorder. *Bipolar Disord.* 2007; 9:93–102.10.1111/j.1399-5618.2007.00427.x [PubMed: 17391353]
- Kapczinski F, Passos IC. Wet minds, dry minds, and the future of psychiatry as a science. *Acta Psychiatr Scand.* 2015; 131:7–8.10.1111/acps.12357 [PubMed: 25495018]
- Klöppel S, Abdulkadir A, Jack CR, Koutsouleris N, Mourão-Miranda J, Vemuri P. Diagnostic neuroimaging across diseases. *Neuroimage.* 2012; 61:457–63.10.1016/j.neuroimage.2011.11.002 [PubMed: 22094642]
- Koutsouleris N, Davatzikos C, Bottlender R, Patschurek-Kliche K, Scheuerecker J, Decker P, Gaser C, Möller HJ, Meisenzahl EM. Early recognition and disease prediction in the at-risk mental states for psychosis using neurocognitive pattern classification. *Schizophr Bull.* 2012; 38:1200–15.10.1093/schbul/sbr037 [PubMed: 21576280]
- Kriegeskorte N, Simmons WK, Bellgowan PSF, Baker CI. Circular analysis in systems neuroscience: the dangers of double dipping. *Nat Neurosci.* 2009; 12:535–40.10.1038/nn.2303 [PubMed: 19396166]
- Kupka RW, Luckenbaugh DA, Post RM, Suppes T, Altshuler LL, Keck PE, Frye MA, Denicoff KD, Grunze H, Leverich GS, McElroy SL, Walden J, Nolen WA. Comparison of rapid-cycling and non-rapid-cycling bipolar disorder based on prospective mood ratings in 539 outpatients. *Am J Psychiatry.* 2005; 162:1273–80.10.1176/appi.ajp.162.7.1273 [PubMed: 15994709]
- Laes JR, Sponheim SR. Does cognition predict community function only in schizophrenia?: A study of schizophrenia patients, bipolar affective disorder patients, and community control subjects. *Schizophr Res.* 2006; 84:121–131.10.1016/j.schres.2005.11.023 [PubMed: 16443348]
- Lavagnino L, Amianto F, Mwangi B, D'Agata F, Spalatro A, Zunta-Soares GB, Abbate Daga G, Mortara P, Fassino S, Soares JC. Identifying neuroanatomical signatures of anorexia nervosa: a multivariate machine learning approach. *Psychol Med.* 2015;1–8.10.1017/S0033291715000768
- Levaux MN, Potvin S, Sepehry AA, Sablier J, Mendrek A, Stip E. Computerized assessment of cognition in schizophrenia: promises and pitfalls of CANTAB. *Eur Psychiatry.* 2007; 22:104–15.10.1016/j.eurpsy.2006.11.004 [PubMed: 17227707]
- Lish JD, Dime-Meenan S, Whybrow PC, Price RA, Hirschfeld RM. The National Depressive and Manic-depressive Association (DMDA) survey of bipolar members. *J Affect Disord.* 1994; 31:281–94. [PubMed: 7989643]
- Mann-Wrobel MC, Carreno JT, Dickinson D. Meta-analysis of neuropsychological functioning in euthymic bipolar disorder: an update and investigation of moderator variables. *Bipolar Disord.* 2011; 13:334–42.10.1111/j.1399-5618.2011.00935.x [PubMed: 21843273]
- Martínez-Arán A, Vieta E, Reinares M, Colom F, Torrent C, Sánchez-Moreno J, Benabarre A, Goikolea JM, Comes M, Salamero M. Cognitive function across manic or hypomanic, depressed, and euthymic states in bipolar disorder. *Am J Psychiatry.* 2004; 161:262–70. [PubMed: 14754775]
- Martino DJ, Marengo E, Igoa A, Scápola M, Ais ED, Perinot L, Strejilevich SA. Neurocognitive and symptomatic predictors of functional outcome in bipolar disorders: a prospective 1 year follow-up study. *J Affect Disord.* 2009; 116:37–42.10.1016/j.jad.2008.10.023 [PubMed: 19033081]
- Mathers CD, Iburg KM, Begg S. Adjusting for dependent comorbidity in the calculation of healthy life expectancy. *Popul Health Metr.* 2006; 4:4.10.1186/1478-7954-4-4 [PubMed: 16620383]
- Merikangas KR, Akiskal HS, Angst J, Greenberg PE, Hirschfeld RMA, Petukhova M, Kessler RC. Lifetime and 12-month prevalence of bipolar spectrum disorder in the National Comorbidity

- Survey replication. *Arch Gen Psychiatry*. 2007; 64:543–52.10.1001/archpsyc.64.5.543 [PubMed: 17485606]
- Montgomery SA, Asberg M. A new depression scale designed to be sensitive to change. *Br J Psychiatry*. 1979; 134:382–9. [PubMed: 444788]
- Mwangi B, Ebmeier KP, Matthews K, Steele JD. Multi-centre diagnostic classification of individual structural neuroimaging scans from patients with major depressive disorder. *Brain*. 2012; 135:1508–21.10.1093/brain/aws084 [PubMed: 22544901]
- Mwangi B, Spiker D, Zunta-Soares GB, Soares JC. Prediction of pediatric bipolar disorder using neuroanatomical signatures of the amygdala. *Bipolar Disord*. 2014; 16:713–21.10.1111/bdi.12222 [PubMed: 24917530]
- Mwangi B, Tian TS, Soares JC. A Review of Feature Reduction Techniques in Neuroimaging. *Neuroinformatics*. 2013; 12:229–244.10.1007/s12021-013-9204-3 [PubMed: 24013948]
- Mwangi B, Wu MJ, Bauer IE, Modi H, Zeni CP, Zunta-Soares GB, Hasan KM, Soares JC. Predictive classification of pediatric bipolar disorder using atlas-based diffusion weighted imaging and support vector machines. *Psychiatry Res*. 2015; 234:265–71.10.1016/j.psychres.2015.10.002 [PubMed: 26459075]
- Najt P, Perez J, Sanches M, Peluso MAM, Glahn D, Soares JC. Impulsivity and bipolar disorder. *Eur Neuropsychopharmacol*. 2007; 17:313–20.10.1016/j.euroneuro.2006.10.002 [PubMed: 17140772]
- Nordentoft M, Mortensen PB, Pedersen CB. Absolute risk of suicide after first hospital contact in mental disorder. *Arch Gen Psychiatry*. 2011; 68:1058–64.10.1001/archgenpsychiatry.2011.113 [PubMed: 21969462]
- Olley AL, Malhi GS, Bachelor J, Cahill CM, Mitchell PB, Berk M. Executive functioning and theory of mind in euthymic bipolar disorder. *Bipolar Disord*. 2005; 7(Suppl 5):43–52.10.1111/j.1399-5618.2005.00254.x [PubMed: 16225560]
- Passarotti AM, Ellis J, Wegbreit E, Stevens MC, Pavuluri MN. Reduced functional connectivity of prefrontal regions and amygdala within affect and working memory networks in pediatric bipolar disorder. *Brain Connect*. 2012; 2:320–34.10.1089/brain.2012.0089 [PubMed: 23035965]
- Passarotti AM, Sweeney JA, Pavuluri MN. Fronto-limbic dysfunction in mania pre-treatment and persistent amygdala over-activity post-treatment in pediatric bipolar disorder. *Psychopharmacology (Berl)*. 2011; 216:485–99.10.1007/s00213-011-2243-2 [PubMed: 21390505]
- Pavuluri MN, Passarotti A. Neural bases of emotional processing in pediatric bipolar disorder. *Expert Rev Neurother*. 2008; 8:1381–7.10.1586/14737175.8.9.1381 [PubMed: 18759550]
- Post RM, Fleming J, Kapczinski F. Neurobiological correlates of illness progression in the recurrent affective disorders. *J Psychiatr Res*. 2012; 46:561–73.10.1016/j.jpsychires.2012.02.004 [PubMed: 22444599]
- Post RM, Kalivas P. Bipolar disorder and substance misuse: pathological and therapeutic implications of their comorbidity and cross-sensitisation. *Br J Psychiatry*. 2013; 202:172–6.10.1192/bjp.bp.112.116855 [PubMed: 23457180]
- Quraishi S, Frangou S. Neuropsychology of bipolar disorder: a review. *J Affect Disord*. 2002; 72:209–26. [PubMed: 12450638]
- 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; 71:1222–30.10.1001/jamapsychiatry.2014.1100 [PubMed: 25188810]
- Robbins TW, James M, Owen AM, Sahakian BJ, Lawrence AD, McInnes L, Rabbitt PM. A study of performance on tests from the CANTAB battery sensitive to frontal lobe dysfunction in a large sample of normal volunteers: implications for theories of executive functioning and cognitive aging Cambridge Neuropsychological Test Automated Batte. *J Int Neuropsychol Soc*. 1998; 4:474–90. [PubMed: 9745237]
- Tibshirani, Robert. Regression shrinkage and selection via the lasso: a retrospective. *J R Stat Soc*. 2011; 73:273–282.
- Robinson LJ, Ferrier IN. Evolution of cognitive impairment in bipolar disorder: a systematic review of cross-sectional evidence. *Bipolar Disord*. 2006; 8:103–16.10.1111/j.1399-5618.2006.00277.x [PubMed: 16542180]

- Robinson LJ, Thompson JM, Gallagher P, Goswami U, Young AH, Ferrier IN, Moore PB. A meta-analysis of cognitive deficits in euthymic patients with bipolar disorder. *J Affect Disord.* 2006; 93:105–15.10.1016/j.jad.2006.02.016 [PubMed: 16677713]
- Savitz JB, Rauch SL, Drevets WC. Reproduced from Habenula volume in bipolar disorder and major depressive disorder: a high-resolution magnetic resonance imaging study. *Mol Psychiatry.* 2013; 18:523.10.1038/mp.2013.43 [PubMed: 23608961]
- Schnack HG, Nieuwenhuis M, van Haren NEM, Abramovic L, Scheewe TW, Brouwer RM, Hulshoff Pol HE, Kahn RS. Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects. *Neuroimage.* 2014; 84:299–306.10.1016/j.neuroimage.2013.08.053 [PubMed: 24004694]
- Sepede G, De Berardis D, Campanella D, Perrucci MG, Ferretti A, Salerno RM, Di Giannantonio M, Romani GL, Gambi F. Neural correlates of negative emotion processing in bipolar disorder. *Prog Neuropsychopharmacol Biol Psychiatry.* 2015; 60:1–10.10.1016/j.pnpbp.2015.01.016 [PubMed: 25661850]
- Struyf J, Dobrin S, Page D. Combining gene expression, demographic and clinical data in modeling disease: a case study of bipolar disorder and schizophrenia. *BMC Genomics.* 2008; 9:531.10.1186/1471-2164-9-531 [PubMed: 18992130]
- Swann AC, Lijffijt M, Lane SD, Steinberg JL, Moeller FG. Increased traitlike impulsivity and course of illness in bipolar disorder. *Bipolar Disord.* 2009a; 11:280–8.10.1111/j.1399-5618.2009.00678.x [PubMed: 19419385]
- Swann AC, Lijffijt M, Lane SD, Steinberg JL, Moeller FG. Severity of bipolar disorder is associated with impairment of response inhibition. *J Affect Disord.* 2009b; 116:30–6.10.1016/j.jad.2008.10.022 [PubMed: 19038460]
- Swann AC, Pazzaglia P, Nicholls A, Dougherty DM, Moeller FG. Impulsivity and phase of illness in bipolar disorder. *J Affect Disord.* 2003; 73:105–11. [PubMed: 12507743]
- Sweeney JA, Kmiec JA, Kupfer DJ. Neuropsychologic impairments in bipolar and unipolar mood disorders on the CANTAB neurocognitive battery. *Biol Psychiatry.* 2000; 48:674–84. [PubMed: 11032979]
- Tabarés-Seisdedos R, Balanzá-Martínez V, Sánchez-Moreno J, Martínez-Aran A, Salazar-Fraile J, Selva-Vera G, Rubio C, Mata I, Gómez-Beneyto M, Vieta E. Neurocognitive and clinical predictors of functional outcome in patients with schizophrenia and bipolar I disorder at one-year follow-up. *J Affect Disord.* 2008; 109:286–99.10.1016/j.jad.2007.12.234 [PubMed: 18289698]
- Thompson JM, Gallagher P, Hughes JH, Watson S, Gray JM, Ferrier IN, Young AH. Neurocognitive impairment in euthymic patients with bipolar affective disorder. *Br J Psychiatry.* 2005; 186:32–40.10.1192/bjp.186.1.32 [PubMed: 15630121]
- Torres JJ, Boudreau VG, Yatham LN. Neuropsychological functioning in euthymic bipolar disorder: a meta-analysis. *Acta Psychiatr Scand Suppl.* 2007;17–26.10.1111/j.1600-0447.2007.01055.x [PubMed: 17688459]
- Torres JJ, DeFreitas CM, DeFreitas VG, Bond DJ, Kunz M, Honer WG, Lam RW, Yatham LN. Relationship between cognitive functioning and 6-month clinical and functional outcome in patients with first manic episode bipolar I disorder. *Psychol Med.* 2010; 41:971–982.10.1017/S0033291710001613 [PubMed: 20810001]
- Wu MJ, Wu HE, Mwangi B, Sanches M, Selvaraj S, Zunta-Soares GB, Soares JC. Prediction of pediatric unipolar depression using multiple neuromorphometric measurements: a pattern classification approach. *J Psychiatr Res.* 2015; 62:84–91.10.1016/j.jpsychires.2015.01.015 [PubMed: 25687738]
- Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. *Br J Psychiatry.* 1978; 133:429–35. [PubMed: 728692]
- Zubieta JK, Huguelet P, O'Neil RL, Giordani BJ. Cognitive function in euthymic bipolar I disorder. *Psychiatry Res.* 2001; 102:9–20. [PubMed: 11368835]

Highlights

- CANTAB coupled with LASSO distinguish euthymic bipolar disorder (BD) patients from healthy controls (HC).
- Prediction of BD patients from HC is performed at an individual subject level.
- Most relevant CANTAB features distinguishing individual BD patients from HC were identified.
- Patients with rapid cycling (RC) were found to be more cognitively impaired than those without.

Figure 1.

A flow diagram showing the LASSO algorithm training and testing process using a leave-one-out cross-validation (LOOCV) approach.

Figure 2.

A 'Confusion matrix' and a receiver operating characteristic (ROC) curve used to calculate LASSO algorithm's prediction accuracy, sensitivity, specificity and area under receive operating characteristic curve.

Figure 3.

A) A bar graph showing model coefficients assigned to the ‘most relevant’ CANTAB neurocognitive variables by the LASSO algorithm. These neurocognitive variables were the most relevant in distinguishing individual euthymic BD patients from HC. Positive model coefficients reflect higher scores in BD patients as compared to HC, whilst on the contrary negative coefficients represent lower scores BD patients. B) The number of commission errors in response to negative stimuli was among the ‘most relevant’ variable predicting membership to BD group. Compared to HC patients with BD committed a greater number of commission errors when exposed to negative stimuli which may indicate the presence of the affective cognitive bias commonly observed in BD. Predictor variable coefficients should be interpreted with caution as they are selected to maximize prediction accuracy as opposed to represent individual coefficient accuracy. Error bars represent the standard error of the mean - estimated by the standard deviation divided by the square root of the sample size.

Figure 4.

A) A bar graph comparing predicted probability scores between HC, BD without RC and BD with RC. Individual probability scores were ‘normalized’ by subtracting (0.5) for visualization purposes only and therefore HC had negative probability scores while BD patients had positive probabilities. Analysis of non-parametric Kruskal-Wallis statistical test comparing predicted probability scores between healthy controls, BD patients without RC and BD patients with RC was performed. Predicted probability scores in BD patients with RC and healthy control differed significantly (Kruskal-Wallis test $p = 0.015$) whilst BD without RC were less distinguishable from HC (Kruskal-Wallis test $p = 0.073$). B) A bootstrapping technique was used to estimate the distribution of the mean predicted probability for HC, BD patients without RC and BD patients with RC. For each group, the predicted probabilities scores were resampled with replication and the corresponding mean was calculated 100000, which led to a smooth estimated density function of the predicted probabilities.

Table 1
Participants' demographics and clinical characteristics

	Healthy controls mean (SD)	BD mean (SD)	P-value
Age (years)	35.20 (14.48)	36.08 (12.66)	0.8359 ^a
Female (total)	13 (21)	16 (21)	0.3167 ^b
Education (years)	16 (4.10)	15.15 (3.22)	0.4743 ^a
Bipolar type	-	BD-I 12/21 BD-II 6/21 BD-NOS 3/21	-
Current mood	-	Euthymic - 21/21	-
HAMD	0.35 (0.81)	7.78 (7.64)	P=0.0001 ^a
YMRS	0.45 (0.76)	4.70 (5.50)	P=0.0015 ^a
MADRS	0.15 (0.49)	9.78 (10.77)	P=0.0003 ^a
HSES	5.25 (2.90)	4.84 (3.04)	P=0.6706 ^a
Currently or previously taken any psychotropic medication	-	19	-

^a student t-test,

^b chi-square test, BD-bipolar disorder, HAMD- Hamilton depression scale, HSES- Hollingshead socioeconomic status (occupational score), MADRS- Montgomery-Åsberg depression rating scale, SD - standard deviation, YMRS- Young mania rating scale.

Table 2
Cognitive tasks and CANTAB measurements

No.	CANTAB Task	Evaluation	Measurements
1	Affective Go/No-Go	Affective and Cognitive control	Reaction time, accuracy
2	Big/Little Circle	Motor speed	Reaction time, accuracy
3	Cambridge Gambling Task	Decision-making	Reaction time, accuracy, proportion bets across trials with more/equally/less likely outcome
4	Choice Reaction Time	Cognitive processing speed	Reaction time, accuracy
5	Motor Screening	Motor processing speed	Reaction time
6	Match to Sample Visual Search	Visuo-motor speed	Reaction time, accuracy
7	Rapid Visual Processing	Sustained attention	Reaction time, accuracy
8	Spatial Recognition Memory	Spatial memory	Reaction time, accuracy
9	Spatial Span task	Spatial working memory	Span length, number of attempts, reaction time

* Reaction time is in milliseconds.