Psychological Medicine

cambridge.org/psm

Original Article

*Du Lei and Kun Qin contributed equally to this study.

Cite this article: Lei D *et al* (2023). Brain morphometric features predict medication response in youth with bipolar disorder: a prospective randomized clinical trial. *Psychological Medicine* **53**, 4083–4093. https://doi.org/10.1017/S0033291722000757

Received: 14 September 2021
Revised: 17 January 2022
Accepted: 27 February 2022
First published online: 8 April 2022

Key words:
Deep learning; bipolar disorder; MRI; cortical thickness; treatment; Young Mania Rating Scale

Author for correspondence:
Du Lei,
E-mail: alien18@163.com

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

CAMBRIDGE UNIVERSITY PRESS

Brain morphometric features predict medication response in youth with bipolar disorder: a prospective randomized clinical trial

Du Lei[1,*], Kun Qin[1,2,*], Wenbin Li[1,2], Walter H. L. Pinaya[3], Maxwell J. Tallman[1], L. Rodrigo Patino[1], Jeffrey R. Strawn[1], David Fleck[1], Christina C. Klein[1], Su Lui[2], Qiyong Gong[2], Caleb M. Adler[1], Andrea Mechelli[4], John A. Sweeney[1,2] and Melissa P. DelBello[1]

[1]Department of Psychiatry and Behavioral Neuroscience, University of Cincinnati College of Medicine, Cincinnati 45219, OH, USA; [2]Huaxi MR Research Center (HMRRC), Department of Radiology, West China Hospital of Sichuan University, Chengdu 610041, PR China; [3]Department of Biomedical Engineering, School of Biomedical Engineering & Imaging Sciences, King's College London, Westminster Bridge Road, London, UK and [4]Department of Psychosis Studies, Institute of Psychiatry, Psychology & Neuroscience, King's College London, De Crespigny Park, London, UK

Abstract

Background. Identification of treatment-specific predictors of drug therapies for bipolar disorder (BD) is important because only about half of individuals respond to any specific medication. However, medication response in pediatric BD is variable and not well predicted by clinical characteristics.

Methods. A total of 121 youth with early course BD (acute manic/mixed episode) were prospectively recruited and randomized to 6 weeks of double-blind treatment with quetiapine ($n = 71$) or lithium ($n = 50$). Participants completed structural magnetic resonance imaging (MRI) at baseline before treatment and 1 week after treatment initiation, and brain morphometric features were extracted for each individual based on MRI scans. Positive antimanic treatment response at week 6 was defined as an over 50% reduction of Young Mania Rating Scale scores from baseline. Two-stage deep learning prediction model was established to distinguish responders and non-responders based on different feature sets.

Results. Pre-treatment morphometry and morphometric changes occurring during the first week can both independently predict treatment outcome of quetiapine and lithium with balanced accuracy over 75% (all $p < 0.05$). Combining brain morphometry at baseline and week 1 allows prediction with the highest balanced accuracy (quetiapine: 83.2% and lithium: 83.5%). Predictions in the quetiapine and lithium group were found to be driven by different morphometric patterns.

Conclusions. These findings demonstrate that pre-treatment morphometric measures and acute brain morphometric changes can serve as medication response predictors in pediatric BD. Brain morphometric features may provide promising biomarkers for developing biologically-informed treatment outcome prediction and patient stratification tools for BD treatment development.

Introduction

Bipolar disorder (BD) frequently emerges during adolescence (Lewinsohn, Klein, & Seeley, 2000; Lewinsohn, Seeley, Buckley, & Klein, 2002; Perlis et al., 2004). Identification of treatment-specific predictors of drug therapies is important because only about half of individuals with BD respond to any specific medication (Hirschfeld et al., 2004; Keck et al., 2003; Keck, Welge, Strakowski, Arnold, & McElroy, 2000). Reliable predictive models could guide personalized therapeutics and implementation of supplemental drug and psychotherapeutic interventions to decrease the high levels of morbidity and mortality associated with the disorder.

Interest in identifying neuroimaging biomarkers of treatment outcome in BD has received increased attention over the past two decades (Ketter & Wang, 2002; Lim et al., 2013). For example, Passarotti, Sweeney, and Pavuluri (2011) used task-based functional magnetic resonance imaging (MRI) to predict the clinical effects of antipsychotic treatment in pediatric BD. Moore et al. (2009) found total brain gray matter volume was increased in treatment-responsive adults with BD after 4 weeks of lithium treatment. Fleck et al. (2017) used magnetic resonance spectroscopy to predict medication response in patients with BD, and Zhang et al. (2018) showed that pre-treatment cortical thickness measures predicted response to an antipsychotic drug in youth with BD. Among these measures, brain morphometric measures

have been most widely used to predict and track the effects of drug therapies for psychiatric disorders (Hazlett et al., 2017; Vieira et al., 2020).

Quetiapine and lithium are effective and widely used treatments for BD. Different neuroprotection effects between quetiapine and lithium have been reported in previous studies. Specifically, lithium was found to be more effective than quetiapine in slowing the progression of white matter abnormalities (Berk et al., 2017). Dandash et al. found that lithium exhibited higher efficiency in reversing the hyperconnectivity of striatal areas than quetiapine (Dandash et al., 2018). For brain gray matter morphometry, our previous work suggested that discrete patterns of baseline cortical thickness in temporal and parietal regions differentially predicted treatment response of quetiapine but not lithium (Zhang et al., 2018). Given the different neural pathways impacted by quetiapine and lithium, comparing the discriminative regions in these two predictive models may provide insights into the neuropharmacological mechanisms of clinical efficacy of the two drugs. While there has been a study using clinical features as predictors to build prediction models for these two drugs separately (Kim et al., 2019), models based on MRI features have not yet been established.

Most studies of neuroimaging predictors of treatment outcome have examined individuals with long-term BD who had received extensive treatment in the past. Therefore, some findings in prior work may have been impacted by the effects of chronicity and previous or current medications. Studying recent onset cases with limited treatment history may better illuminate medication effects on illness processes and predictive brain features, as well as provide more clinically useful information. Further, studying early course individuals with BD may provide more clinically useful information because an extensive clinical history is not available to guide treatment decisions (Brooks & Vizueta, 2014). Third, focusing on patients experiencing an acute mixed/manic episode can facilitate assessment of MRI features in relation to treatment outcome. Fourth, preclinical studies have revealed changes in brain systems induced by short-term administration of lithium and quetiapine. For example, 7 days of lithium therapy resulted in changes in forebrain membrane properties in a rat model of mania (Vošahlíková, Roubalová, Brejchová, Alda, & Svoboda, 2021), and 1-week treatment of quetiapine was found to enhance the level of hippocampal neurogenesis (Luo, Xu, & Li, 2005). In addition, evidence from clinical trials based on neuroimaging techniques also support brain changes following short-term quetiapine or lithium therapy. Brain D2 dopamine receptor occupancy can be observed following acute quetiapine treatment (Nord et al., 2011), and significant decreases in myoinositol levels were identified in the right frontal cortex after 5–7 days of lithium administration (Moore et al., 1999). Moreover, our previous study using brain structural MRI found youth with BD exhibited changes in brain structural network following 1-week quetiapine or lithium treatment (Lei et al., 2021). Although these findings suggest that brain changes can occur after acute drug treatment, few studies investigated whether such acute brain changes can serve as predictors of treatment outcome.

Previous studies have suggested that deep learning, a type of machine learning capable of capturing high orders of complexity and abstraction (Kim, Calhoun, Shim, & Lee, 2016; LeCun, Bengio, & Hinton, 2015; Lei et al., 2020), may yield higher classifier accuracy than the current widely adopted traditional machine learning models (Pinaya et al., 2016; Vieira, Pinaya, & Mechelli, 2017). It has been reported that brain structure is undergoing non-linear trajectory of brain structural maturational changes in typically developing youth (Giedd et al., 1999; Tamnes et al., 2017). Previous studies have also found non-linear characteristics in brain alterations associated with BD and increased brain complexity in mania (Bahrami, Seyedsadjadi, Babadi, & Noroozian, 2005; Fernández, Al-Timemy, Ferre, Rubio, & Escudero, 2018). Considering additional complicated neuroprogression and medication effects, non-linear models may be better positioned to address the brain complexity involved in a randomized clinical trial of pediatric BD compared with linear methods (e.g. principal component analysis, sparse learning) (Hazlett et al., 2017). Deep neural network can precisely recognize the most differentiable features related to medication response from the complex longitudinal structural patterns in a non-linear way, and facilitate the extraction of optimal low-dimensional representations for clinicians unequipped with expert feature engineering knowledge. We therefore utilized a two-stage prediction pipeline that includes a deep neural network component for non-linear dimensionality followed by an additional support vector machine (SVM) classifier (Hazlett et al., 2017).

With these considerations in mind, we recruited a cohort of young participants with BD who were early in their illness course and randomized to receive quetiapine or lithium trial. Our aims were to determine the utility of pre-treatment morphometric data and change in brain structure after 1 week of treatment for the prediction of treatment response at the end of the 6 weeks trial. We hypothesized that (i) baseline image data could significantly predict medication response in bipolar youth at the individual level; and (ii) change in morphometric features from baseline to 1 week after treatment initiation would predict treatment response, and the longitudinal joint model combining baseline data and week 1 structural changes data would predict medication response with highest accuracy. Finally, considering that the drug mechanisms of lithium and quetiapine are different (Ketter, Miller, Dell'Osso, & Wang, 2016), we hypothesized that (iii) pre-treatment morphometric features and their change may be differentially related to outcome prediction in the two treatment groups.

Methods

Participants

This study was approved by the University of Cincinnati and the Cincinnati Children's Hospital Medical Center Institutional Review Boards. Youth with bipolar I disorder were recruited from the Cincinnati Children's Hospital Medical Center, the University of Cincinnati, and the local community. Diagnosis of bipolar I disorder was confirmed using the Washington University in St. Louis Kiddie Schedule of Affective Disorders and Schizophrenia administered by raters with demonstrated inter-rater reliability ($\kappa > 0.9$) (Geller et al., 2001). Young Mania Rating Scale (YMRS) was used to assess mania symptoms (Young, Biggs, Ziegler, & Meyer, 1978). Written informed assent and consent about the study procedures and purpose were provided by all participants and their legal guardians (registered in https://clinicaltrials.gov/; registration number: NCT00893581).

To be included, participants were required to be: (1) within the age range of 10–18 years old; (2) experiencing a manic or mixed episode; (3) having a baseline YMRS score $\geqslant$20; (4) within 2 years from onset of BD; (5) having no prior psychiatric hospitalizations for mania; (6) having no treatment history with therapeutic doses of antipsychotics or mood stabilizers for over 3 months, and no

https://doi.org/10.1017/S0033291722000757 Published online by Cambridge University Press

psychotropic medication during the week prior to baseline scans. Participants were excluded if they had (1) a contraindication to MRI scanning; (2) an IQ <70; (3) a positive pregnancy test; (4) a history of a major systemic or neurological illness, or an episode of loss of consciousness >10 min; (5) any lifetime DSM-IV-TR substance use disorder (nicotine dependence was permitted); or (6) a lifetime DSM-IV-TR diagnosis of a pervasive developmental disorder or post-traumatic stress disorder.

Treatment procedures

Following initial clinical evaluation and MRI scanning, participants were randomized by an investigational pharmacist (C.C.K.) to double-blind treatment with lithium or quetiapine and evaluated clinically weekly for 6 weeks. The randomization schedule was stratified by presence/absence of attention deficit and hyperactivity disorder, presence/absence of psychosis, and the mood state (mixed *v.* manic episode). Quetiapine was initiated at 100 mg qhs and lithium carbonate was initiated at 30 mg/kg (maximum starting dose of 600 mg twice daily). Patients were also given placebo capsules for the medication to which they were not assigned. Quetiapine was titrated to a target dose of 400–600 mg/day based on tolerability and response. Lithium was titrated to a serum level of 1.0–1.2 mEq/L. Treatment was administered in a double-dummy, double-blind manner, with an unblinded study psychiatrist monitoring trough lithium levels and making dose adjustments independent from treating psychiatrist and clinical raters. Blinded clinical tolerability dose adjustment recommendations took precedent over un-blinded double dummy dose adjustment recommendations. There were no significant changes of the treatment methods or outcomes after trial commencement. We used the YMRS changes to assess antimanic treatment response at week 6, and treatment responder was defined as exhibiting ⩾50% reduction in YMRS scores from baseline (Patino et al., 2021; Wegbreit et al., 2011).

MRI acquisition

MRI scanning was performed on a 4-T Varian Unity INOVA scanner using a 12-channel head coil at baseline. All participants were instructed to be scanned at baseline prior to treatment and 1 week after treatment initiation. Earplugs and headphones were provided to reduce background noise, and foam padding around the head minimized head motion. Following a three-plane gradient echo scan for alignment and localization, a shim procedure was performed to generate a homogeneous magnetic field. High-resolution T1-weighted three-dimensional images were acquired with a Modified Driven Equilibrium Fourier Transform (MDEFT) protocol, optimized for the 4T Varian scanner [Tau (magnetization preparation time) = 1.1 s, TR = 13 ms, TE = 5.3 ms, field of view = 192 × 256 × 256 mm, matrix = 192 × 256 × 256, flip angle = 20 degrees, slice thickness = 1 mm]. Acquired T1-weighted images were inspected by two experienced neuroradiologists who made decisions about excessive motion artifact for scan exclusion. No observable scanning artifacts or gross brain abnormalities were identified in any participant included in the following analyses.

Image processing and brain morphometric feature sets

All structural MRI scans were processed on the same workstation using the FreeSurfer image analysis suite v6.0.0 (http://surfer.nmr.mgh.harvard.edu/) to obtain unbiased estimates of morphometric measures, including surface area, cortical thickness, and subcortical volumes (for detailed calculation and extraction of these morphometric measures, see online Supplementary Materials). Each participant was thus represented by a 150-dimensional feature vector consisting of morphometric features across the whole brain. Specifically, we included the surface area and cortical thickness of 68 cortical regions labeled in the Desikan/Killiany Atlas (Desikan et al., 2006), as well as the volume of bilateral hippocampus and 12 subcortical regions (i.e. bilateral thalamus, amygdala, caudate, putamen, pallidum, and accumbens). To compare the prediction value of baseline brain structure and acute structural alterations induced by medication, the model performance was investigated for baseline morphometric measures and morphometric changes from baseline to week 1 independently. To thoroughly take advantage of longitudinal brain morphometry related to antimanic treatment for better model performance, we further combined baseline and week 1 morphometric features into a concatenated 300-dimensional vector as our longitudinally joint model. Therefore, under the unified prediction framework, there are finally three classification models based on different feature sets (i.e. baseline model, 1-week change model, and longitudinally joint model) to be tested in two medication groups.

Two-stage prediction model based on structural MRI

We implemented a two-stage prediction pipeline to differentiate medication responders from non-responders as described in previous prospective studies (Hazlett et al., 2017; Zhu et al., 2021; Yang et al., 2021). A feedforward multi-layer neural network was adopted as the initial stage for dimensionality reduction (Hinton & Salakhutdinov, 2006), and SVM was included as the second stage to individually discriminate responders from non-responders (Cortes & Vapnik, 1995). The training procedures of neural network for dimensionality reduction mainly contain binary processing of raw features, unsupervised pre-training of stacked autoencoders, and supervised training of fine-tuned neural network. Once the training was finished, the representations stored in the layer before output layer were extracted as the optimal features fed into SVM. Detailed information about the rationale, architecture, and training process of the two-stage prediction model is shown in online Supplementary Materials.

The whole two-stage prediction pipeline was trained and evaluated using 10-fold stratified cross-validation. For each iteration of the cross-validation, we used one part of the dataset to evaluate our model (i.e. testing set) and the remaining nine parts for training (i.e. training set). To avoid information leakage, we performed the non-linear dimensionality reduction (i.e. training of neural network) only based on the training set, and the testing set was only used to evaluate performance and never used for tuning or training our models. The model performance was determined by the balanced accuracy, sensitivity, specificity, and area under receiver operating characteristic curve (AUC). We independently reported the model performance in quetiapine and lithium medication group. All the machine learning analyses were programmed using Python language, where the neural network was implemented in the Pytorch library (Paszke et al., 2019), and the SVM was implemented in the Scikit-Learn library (Pedregosa et al., 2011). All the codes are available at https://github.com/QKmeans0902/Two_stage_prediction_pipeline.

Model transferability between different medication groups

In this analysis, we examined the model transferability between two medication groups. Specifically, we used dataset from one

https://doi.org/10.1017/S0033291722000757 Published online by Cambridge University Press

whole treatment group (quetiapine or lithium) for drug-specific model training and optimization, and the other group to evaluate the predictive utility of the model for treatment outcome to the other drug therapy. To maintain consistency with preceding analyses, model transferability test was also implemented in the identical two-stage prediction pipeline.

Features with greatest contributions to predicting medication outcomes

Given morphometric features with the highest predictive power may differ in different medication groups which can provide useful neurobiological implications, we sought to identify these potentially discriminating patterns. Detailed calculation of feature contribution is presented in online Supplementary Materials. Top 10 features with greatest contributions are reported for different models and medication groups.

Statistical analysis

Group comparisons of demographic and clinical characteristics were performed using SPSS software [version 23 (IBM SPSS Statistics)]. The significance level of prediction model was evaluated using permutation test (Golland & Fischl, 2003). Specifically, we randomly permuted the labels of training set prior to training. The model training and evaluation under cross-validation were then performed to get the estimated accuracy based on the permuted dummy training set. The permutation was repeated 1000 times, and a distribution of the accuracy for randomly permuted data was obtained. The significance level was thus defined as the probability of observing an accuracy from the distribution based on permuted data no less than the real accuracy. Differences in age and IQ were assessed by two-sample t test, while differences in sex and number of responders were assessed using χ^2 test. The p value of differences in parental socioeconomic status was calculated by Mann–Whitney U test given the non-normality, and the statistical test for differences in longitudinal clinical scale scores was two-way repeated measures ANOVA.

Results

Demographic and clinical characteristics

This prospective randomized clinical trial recruited 149 youth with BD. Data from nine participants were excluded due to failure to finish baseline structural MRI scan or excessive head motion. Eleven participants in the lithium treatment group ($n = 61$) and eight participants in the quetiapine treatment group ($n = 79$) were lost to follow-up. Ultimately, 71 participants treated with quetiapine and 50 participants treated with lithium who had completed all follow-up structural MRI examinations and clinical assessments were finally included in the analysis.

The demographic and clinical characteristics are listed in Table 1 and online Supplementary Table S1. No evidence of significant differences in age, sex, IQ, parental socioeconomic status, or number of responders was found between the quetiapine and lithium groups (all $p > 0.05$). No significant time × treatment group interaction effects in YMRS scores were observed. For the group effect, we found no significant between-group differences in YMRS scores at any follow-up timepoint ($p > 0.05$). For the time effect, the YMRS scores were significantly decreased from baseline to week 6 in both quetiapine and lithium groups ($p < 0.001$).

Prediction performance of medication response

The prediction pipeline in the current study was shown in Fig. 1. Baseline structural MRI predicted treatment response at week 6 with balanced accuracy of 79.6% (AUC: 0.89, $p < 0.001$) for the quetiapine treatment group. For the lithium treatment group, baseline morphometric measures achieved balanced accuracy of 75.8% (AUC: 0.85, $p = 0.003$). When using 1-week brain morphometric changes as predictors, balanced accuracies in the quetiapine and lithium treatment group were 76.5% (AUC: 0.81, $p < 0.001$) and 78.9% (AUC: 0.85, $p < 0.001$), respectively. Combining brain morphometric features at baseline and week 1 allowed for the highest prediction performance. For the quetiapine treatment group, the balanced accuracy was 83.2% (AUC: 0.93, $p < 0.001$). For the lithium treatment group, the balanced accuracy was 83.5% (AUC: 0.89, $p < 0.001$). All the evaluation metrics for model performance were listed in Table 2.

Model transferability analyses between drug treatments

All models failed to reach comparable performance in the transferability test when prediction models developed for one drug were applied to the other drug treated group (all $p > 0.05$; Table 2). Baseline morphometric measures showed slightly above chance level performance (i.e. 50%) using the quetiapine model to predict lithium treatment outcome (balanced accuracy, 54.4%) and *vice versa* (balanced accuracy, 55.7%). Using week 1 structural change data, both models showed poor transferability below chance level (balanced accuracy of 43.8% for quetiapine model predicting lithium outcome and balanced accuracy of 47.7% for lithium predicting quetiapine outcome). When considering both baseline and week 1 morphometric data together, both models transferred around chance level (quetiapine model predicting lithium outcome: balanced accuracy, 51.5%; lithium model predicting quetiapine outcome: balanced accuracy, 48.9%).

Features with greatest contributions to treatment response prediction

Within the models considering the predictive utility of pre-treatment morphometric features, top predictive features in the lithium group mostly included cortical thickness measures, while in the quetiapine group, cortical thickness, cortical surface area, and subcortical volume measures were leading predictors. When using acute structural changes from baseline to week 1 for prediction, top features in the lithium group included cortical surface area and hippocampal volume, while in the quetiapine group, the leading predictors were cortical thickness measures (Table 3 and Fig. 2).

Discussion

In the present study, we tested the ability of brain morphometric features to predict medication treatment response in youth with BD. We know of only two previous studies that applied pre-treatment MRI data to predict treatment outcome in BD: Wade et al. (2016) trained an SVM with subcortical volume and cortical thickness features to predict response to electroconvulsive therapy in depressed adults, eight of whom had BD. They reported an accuracy of 89%. Fleck et al. (2017) trained a machine learning model with pre-treatment fMRI and magnetic resonance spectroscopy data to predict response to lithium at week 8 in 20 adults

https://doi.org/10.1017/S0033291722000757 Published online by Cambridge University Press

Table 1. Demographic and clinical characteristics of youth with bipolar disorder

Characteristic	Quetiapine ($n = 71$)	Lithium ($n = 50$)	p value
Age (years)[1]	14.84 ± 1.62 (10.2–17.7)	15.44 ± 1.72 (10.5–17.8)	0.053
Sex (female/male)	46/25 (65.0%/35.0%)	35/25 (70.0%/30.0%)	0.548
IQ	102.57 ± 12.14	103.97 ± 11.37	0.599
PSES	2.98 ± 1.11	2.97 ± 1.00	0.969[2]
BMI	23.85 ± 5.53	25.42 ± 7.19	0.198
Duration of current episode (weeks)[3]	19.39 ± 27.95	15.49 ± 23.56	0.430
No. of lifetime episodes[2]	1.31 ± 1.90	0.74 ± 1.03	0.066
Treatment response, *N* (%)			0.531
Responder	52 (73.2%)	34 (68.0%)	
Non-responder	19 (26.8%)	16 (32.0%)	
Mood state, *N* (%)			0.660
Mania	44 (62.0%)	29 (58.0%)	
Mixed	27 (38.0%)	21 (42.0%)	
YMRS			0.580[4]
Baseline	26.42 ± 5.06	26.34 ± 5.13	
Week 1	15.44 ± 6.36	16.56 ± 8.18	
Week 6	9.64 ± 6.84	10.83 ± 6.86	
CDRS-R			0.981[4]
Baseline	37.73 ± 9.05	38.50 ± 8.98	
Week 1	31.33 ± 6.97	31.79 ± 9.19	
Week 6	27.06 ± 8.33	27.88 ± 6.95	
CGI-S			0.983[4]
Baseline	4.84 ± 0.59	4.94 ± 0.55	
Week 1	3.82 ± 0.75	3.94 ± 0.92	
Week 6	2.86 ± 1.26	3.00 ± 1.23	

IQ, intelligence quotient; PSES, parental socioeconomic status; BMI, body mass index; YMRS, Young Mania Rating Scale; CDRS-R, Children's Depression Rating Scale - Revised; CGI-S, Clinical Global Impressions-Severity.
Note: Data are presented as mean ± standard unless otherwise indicated. The p values were calculated by two-sample t test or χ^2 test unless otherwise indicated.
[1]Age are presented as mean ± standard with age range in parenthesis.
[2]Data were available in 111 of 121 participants.
[3]Data were available in 116 of 121 participants.
[4]The p values refer to significance level of time × treatment group interaction effects using two-way repeated measures ANOVA.

with BD, classifying responders *v.* non-responders with 80% accuracy. Our study is considerably larger, which focused on youth with limited clinical and treatment history and considered acute neuroanatomic changes after 1 week of treatment as well as baseline predictors.

Consistent with our first hypothesis, we found that baseline morphometric measures predicted medication response at the individual level with significant accuracy, 79.6% for the quetiapine group and 75.8% for the lithium group. Our findings indicate that pre-treatment morphometric features may provide predictors of treatment response approaching the level needed for clinical application, consistent with findings of the two prior smaller studies (Fleck et al., 2017; Wade et al., 2016). We also assessed the ability of acute changes in morphometric features one week following treatment initiation to predict clinical outcome at week 6. The rationale for this aspect of our design is that early drug-induced changes in brain might provide an indication of clinically-relevant effects of drugs before clinical changes themselves can predict outcome of a drug trial. Changes in morphometric features from baseline achieved comparable prediction accuracy for the two study drugs (76.5% for quetiapine group and 78.9% for lithium group), though models were different for the two drugs. Previous studies using somewhat longer re-test periods (2–4 weeks) after drug treatment initiation have shown neuroanatomic changes with short-term acute treatment with lithium (Anand et al., 2020) and antipsychotic drugs (Keshavan et al., 1994). Our findings suggest that relevant changes can be detected even earlier after treatment initiation. A study of lithium treatment in BD reported effects in the hippocampus/amygdala complex after several years, regions known to be important in emotion processing (Germana et al., 2010). Our study extends these earlier findings by showing that effects occurring as early as one week can predict later treatment response. Thus, the monitoring of early drug effects on brain structure may provide a way of predicting treatment response weeks before treatment outcome can be determined clinically, potentially providing important information to

https://doi.org/10.1017/S0033291722000757 Published online by Cambridge University Press

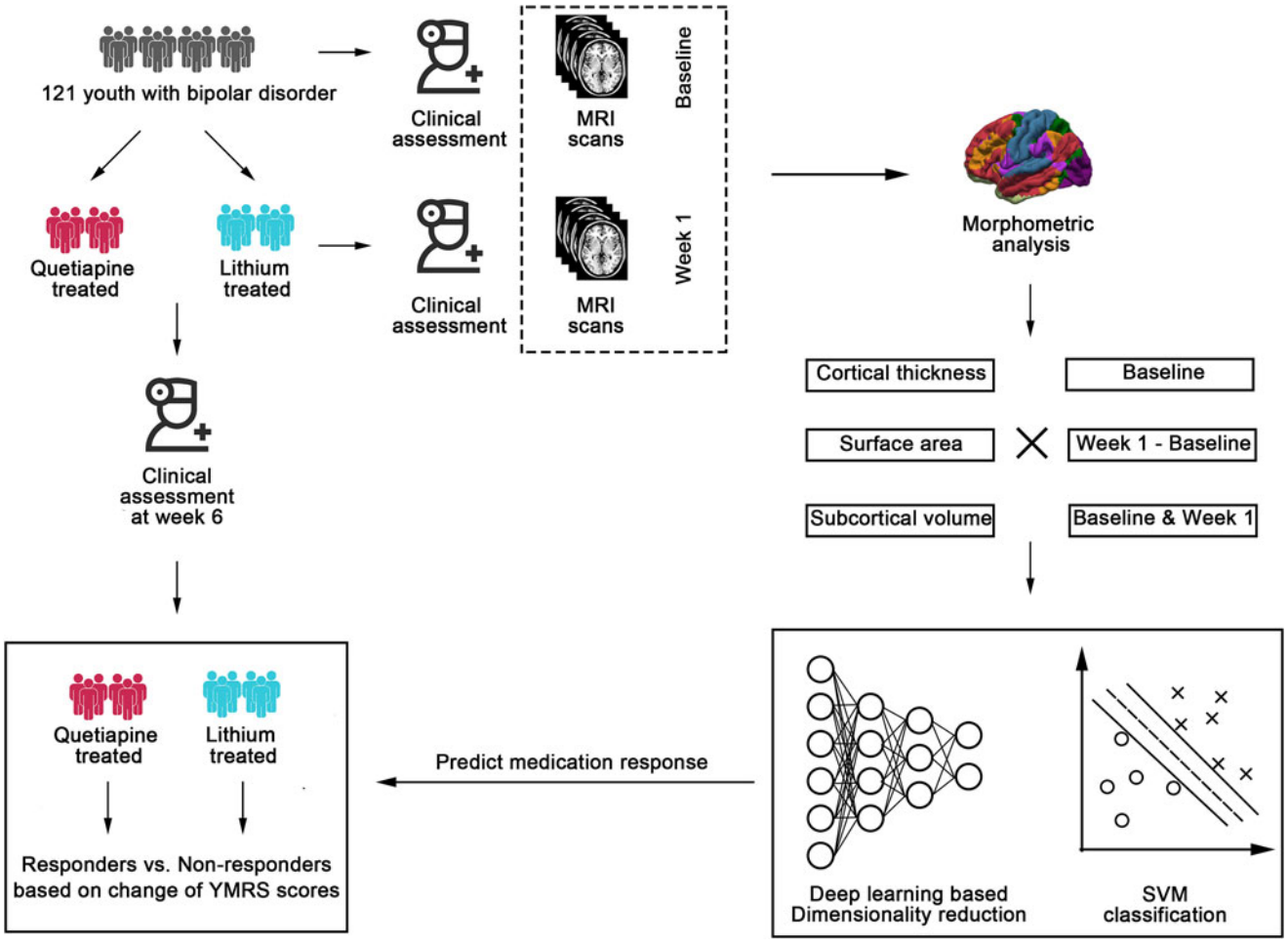


Fig. 1. The pipeline of treatment response prediction. A total of 121 youth with BD were included and randomly assigned to quetiapine and lithium treatment group. Structural MRI examination was performed prior to and at week 1 of the treatment. Clinical assessments were implemented at baseline, week 1, and week 6, respectively. To develop a medication response prediction model using structural MRI data, we extracted the morphometric measures including cortical thickness, surface area, and subcortical volume. Responders were determined as a reduction of YMRS scores >50% at week 6. Baseline, change during the first week (baseline – week 1), and longitudinally combined morphometric features (baseline + week 1) were separately investigated for both medication groups. The two-stage prediction model including non-linear dimensionality reduction and support vector machine classifier was applied consistently. SVM, support vector machine; YMRS, Young Manic Rating Scale.

clinicians considering adjustments to treatment plans in individuals who are not early treatment responders.

Combining baseline brain features and acute treatment-induced effects (1-week) in models predicting 6-week treatment response modestly improved the accuracy (83.2% for the quetiapine group and 83.5% for the lithium group). Therefore, combining baseline and follow-up data within a single model may provide a promising direction for maximizing early prediction of treatment outcome in individuals with BD. While these findings require replication in an independent sample to confirm predictive utility of the models developed, they offer a promising preliminary step toward the development of clinically useful MRI-based biomarkers for guiding optimal and flexible treatment planning for individuals in the early course of BD.

Given these differences in prediction performance, we sought to identify which morphometric features provide the greatest contribution to treatment outcome. At baseline, cortical thickness features were better predictors of lithium response, while cortical surface area and subcortical volumes were more prominent as features predicting response to quetiapine. In contrast, when considering morphometric changes at week 1, the opposite pattern was observed with changes in cortical surface area and subcortical volume measures being more prominent in lithium response prediction, and changes in cortical thickness measures being more predictive of quetiapine response. According to the radial unit hypothesis and the supra-granular layer expansion hypothesis, cortical thickness is determined by the number and size of cells within a cortical column and surrounding neuropil, while surface area is less dynamic being primarily driven by the number of cortical columns established during brain maturation (Rakic, 2009). This suggests that dynamic illness-related pathophysiological changes may be more related to lithium outcomes, while relatively enduring features of brain development at the baseline may be better predictors of quetiapine response and dynamic changes in cells and neuropil at one week may be more related to quetiapine treatment. Changes in surface area at one week following lithium treatment are less readily interpretable, but may indicate a more global shift in the neocortical mantle following administration of lithium salts (Hozer et al., 2020).

https://doi.org/10.1017/S0033291722000757 Published online by Cambridge University Press

Table 2. Model classification and transferability performance between quetiapine and lithium treatment groups

Model performance	BAC, %	SEN, %	SPE, %	AUC
Baseline model				
Quetiapine	79.6	75.0	84.2	0.89
Lithium	75.8	76.5	75.0	0.85
Quetiapine model predicts lithium outcome	54.4	58.8	50.0	0.51
Lithium model predicts quetiapine outcome	55.7	69.2	42.1	0.49
One-week change model				
Quetiapine	76.5	63.5	89.5	0.81
Lithium	78.9	73.5	81.3	0.85
Quetiapine model predicts lithium outcome	43.8	50.0	37.5	0.47
Lithium model predicts quetiapine outcome	47.7	48.1	47.4	0.50
Longitudinally joint model				
Quetiapine	83.2	76.9	89.5	0.93
Lithium	83.5	79.4	87.5	0.89
Quetiapine model predicts lithium outcome	51.5	52.9	50.0	0.58
Lithium model predicts quetiapine outcome	48.9	55.8	42.1	0.53

BAC, balanced accuracy; SEN, sensitivity; SPE, specificity; AUC, area under receiver operating characteristic curve.

Table 3. Top 10 morphometric features showing greatest contribution to baseline and 1-week change model

	Quetiapine		Lithium	
Rank	Measure	Region	Measure	Region
Baseline model				
1	CT	Cuneus L	CT	Inferior temporal L
2	Vol	Accumbens L	CT	Paracentral R
3	SA	Entorhinal L	CT	Isthmus L
4	SA	Temporal pole L	CT	Frontal pole L
5	Vol	Hippocampus L	CT	Inferior parietal R
6	CT	Parahippocampal L	CT	Postcentral L
7	SA	Inferior temporal L	SA	Pericalcarine R
8	CT	Cuneus R	CT	Posterior cingulate L
9	SA	Rostral anterior cingulate L	CT	Pars triangularis L
10	SA	Rostral middle frontal L	CT	Banks of superior temporal sulcus R
One-week change model				
1	CT	Cuneus L	SA	Precentral R
2	CT	Transverse temporal L	CT	Temporal pole L
3	SA	Pars opercularis L	CT	Superior frontal L
4	CT	Rostral middle frontal L	CT	Superior parietal L
5	SA	Postcentral L	SA	Precentral L
6	SA	Rostral middle frontal L	Vol	Hippocampus L
7	CT	Fusiform R	SA	Lateral orbitofrontal R
8	SA	Entorhinal R	SA	Superior temporal L
9	SA	Caudal anterior cingulate L	CT	Middle temporal L
10	CT	Superior parietal L	CT	Transverse temporal R

L, left; R, right; CT, cortical thickness; SA, surface area; Vol, volume.

https://doi.org/10.1017/S0033291722000757 Published online by Cambridge University Press

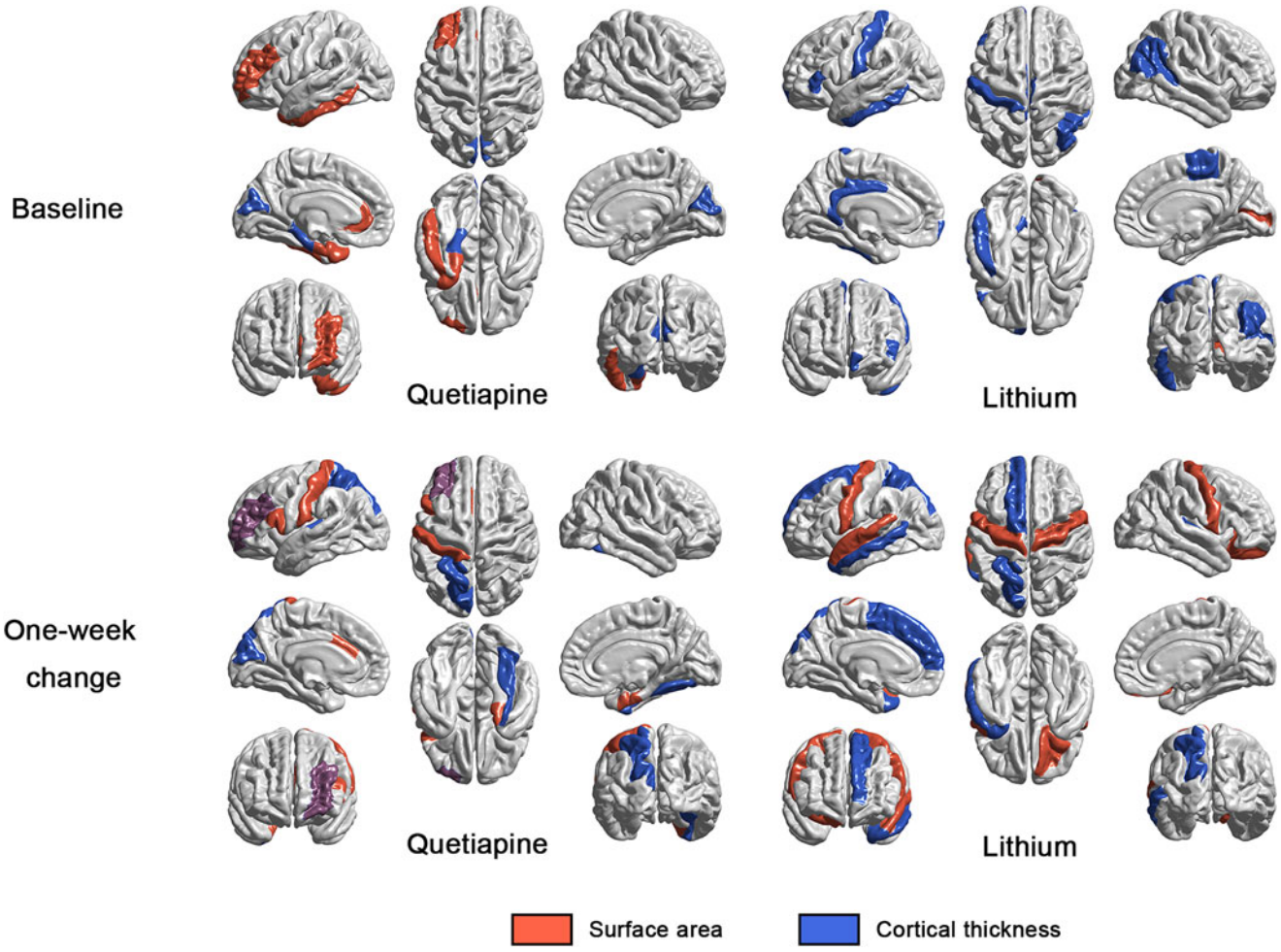


Fig. 2. Cortical regions of surface area and cortical thickness measures among top 10 morphometric features contributing to the non-linear dimensionality reduction. For each model, results were independently showed in both quetiapine and lithium group. Surface area measures are shown in red, and cortical thickness measures are shown in blue. If both cortical thickness and surface area of a single region exhibit top 10 contribution, this region will be shown in a hybrid purple color combining blue and red.

For the regionality of the most discriminative brain morphometric patterns, at baseline, the most discriminative regions for quetiapine were mainly located in temporal and subcortical regions, while frontal and parietal regions predicted treatment outcome better in lithium group. When considering morphometric changes at week 1, the regions contributing most to prediction mainly comprised frontal and temporal regions for both quetiapine and lithium. Previous research had reported disruption in fronto-temporal neural circuitry in remitted patients with BD (Robinson et al., 2009). Our finding may indicate that the acute morphometric changes in fronto-temporal neural circuitry may in part underlie therapeutic treatment effects for both quetiapine and lithium. Nevertheless, we should note the difference in sample size of these two medication groups which may impact precision in the estimation of contribution weight of features. Although our study is designed as a randomized clinical trial to reduce group differences as much as possible, and no significant confounds (i.e. age, sex, IQ, illness duration, YMRS, CDRS-R, etc.) were identified in our analyses, some other potential differences in patient samples to a degree can exist despite randomization. Replication is still needed to determine the extent that differences in patient samples impacted the current findings.

In addition to these general findings, our transferability test for the two treatments was also informative. The results showed that all models failed to reach comparable performance in the transferability test, indicating that the contributing features of the predictive models differed between treatment groups, both for baseline data and for treatment-related change at week 1. While predictive models for the two treatments had similar success, the quite limited transferability of drug-specific models suggests modest level of shared predictors, so that drug class-specific predictive models appear to be needed for treatment outcome prediction. Different changes in morphometric features at week 1 are perhaps not surprising given the markedly different pharmacology of the two study drugs, though the nature of the differences offers insight into regional differences in early drug impact on brain that are relevant to their therapeutic efficacy.

Following recent recommendations on overcoming methodological issues that can lead to inflated results (Arbabshirani, Plis, Sui, & Calhoun, 2017; Wolfers, Buitelaar, Beckmann, Franke, & Marquand, 2015; Woo, Chang, Lindquist, & Wager, 2017), we adopted two methodological precautions. First, to reduce the risk of overfitting and overly optimistic utility prediction estimates (Arbabshirani et al., 2017; Wolfers et al., 2015; Woo

https://doi.org/10.1017/S0033291722000757 Published online by Cambridge University Press

et al., 2017) for the high-dimensional neuroimaging data, we first applied DL technology to identify compact hierarchical features and achieve dimension reduction. Second, we examined region-level features with less noise and lower risk of overfitting than voxel-level data (Vieira et al., 2017). Other approaches for dimension reduction and development of classification algorithms may confirm and potentially refine models developed in the current study (Claude, Houenou, Duchesnay, & Favre, 2020; Collins & Moons, 2019).

Of note, there are several limitations in our current study. First, the models described in this paper will require replication in an independent dataset before any application in clinical decision making. Second, each patient in our study received either quetiapine or lithium monotherapy. Thus, our study can only predict response of each drug and identify predictors of response in two independent treatment groups. Information about prediction of individual-level preferential response to quetiapine or lithium cannot be provided, which should be interpreted with caution. Third, our study was not powered to examine the relationship between MRI features and outcome separately in different age groups. Fourth, the utility of predictive models is generalizable within drug class and remains to be evaluated. Fifth, additional outstanding questions include the ability of other imaging modalities, such as DTI and resting-state fMRI, to inform and supplement prediction models. In addition, it would be helpful to further evaluate the utility of our models for predicting treatment outcomes in adults and in individuals with a long history of illness and drug therapy.

In conclusion, the present study demonstrates that brain morphometric features may predict treatment response for two widely used medications for the treatment of youth with BD. Brain morphometric features provided similar overall prediction performance for both lithium and quetiapine therapies, though the features used to achieve that prediction varied between the two medications. In addition, our findings indicate that brain morphometric changes occurring after a single week of medication exposure are predictive of treatment outcome. These findings provide insight into brain morphometric features associated with the treatment outcome prediction in bipolar youth treated with quetiapine and lithium, and support the potential use of neuroanatomical scans as biomarkers for the optimization of treatment or personalized medication approach which offers potential for reducing the risk of a failed medication trial detectable only following a full course of treatment.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0033291722000757.

Financial support. This study was supported by the National Institute of Mental Health (NIMH) Grant (M.P.D., Grant No. 5R01MH080973) and the National Natural Science Foundation of China (Q.G., Grant No. 81621003).

Conflict of interest. Dr Strawn has received research support from the National Institutes of Health (NIMH/NIEHS/NICHD) as well as Allergan, Neuronetics, and Otsuka. He has received material support from and provided consultation to Myriad Genetics and receives royalties from the publication of two texts (Springer) and serves as an author for UpToDate and an Associate Editor for Current Psychiatry. He has spoken in CME presentations for Neuroscience Education Institute and CMEology. Finally, Dr Strawn also has provided consultation to the FDA and Intracellular Therapeutics. Dr Sweeney consults to VeriSci. Dr DelBello and Dr Adler are on the lecture bureau for Otsuka, and Dr Adler is on the lecture bureau for Janssen. Dr Patino and Dr DelBello have received research support from Acadia, Allergan, Janssen, Johnson and Johnson, Lundbeck, Otsuka, Pfizer, Sunovion, and Supernus, and Dr DelBello has provided consultation or advisory board services for Alkermes, Allergan, Assurex, CMEology, Janssen, Johnson and Johnson, Lundbeck, Neuronetics, Otsuka, Pfizer, Sunovion, and Supernus. Dr Adler has received research support from Merck, Forest, and Alkermes, and provided consultation for Janssen. All other authors declare that they have no competing interests.

References

Anand, A., Nakamura, K., Spielberg, J. M., Cha, J., Karne, H., & Hu, B. (2020). Integrative analysis of lithium treatment associated effects on brain structure and peripheral gene expression reveals novel molecular insights into mechanism of action. *Translational Psychiatry*, *10*(1), 103. doi: 10.1038/s41398-020-0784-z

Arbabshirani, M. R., Plis, S., Sui, J., & Calhoun, V. D. (2017). Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. *NeuroImage*, *145*(Pt B), 137–165. doi: 10.1016/j.neuroimage.2016.02.079

Bahrami, B., Seyedsadjadi, R., Babadi, B., & Noroozian, M. (2005). Brain complexity increases in mania. *Neuroreport*, *16*(2), 187–191. doi: 10.1097/00001756-200502080-00025

Berk, M., Dandash, O., Daglas, R., Cotton, S. M., Allott, K., Fornito, A., … Yücel, M. (2017). Neuroprotection after a first episode of mania: A randomized controlled maintenance trial comparing the effects of lithium and quetiapine on grey and white matter volume. *Translational Psychiatry*, *7*(1), e1011. doi: 10.1038/tp.2016.281

Brooks, J. O., 3rd, & Vizueta, N. (2014). Diagnostic and clinical implications of functional neuroimaging in bipolar disorder. *Journal of Psychiatric Research*, *57*, 12–25. doi: 10.1016/j.jpsychires.2014.05.018

Claude, L. A., Houenou, J., Duchesnay, E., & Favre, P. (2020). Will machine learning applied to neuroimaging in bipolar disorder help the clinician? A critical review and methodological suggestions. *Bipolar Disorder*, *22*(4), 334–355. doi: 10.1111/bdi.12895

Collins, G. S., & Moons, K. G. M. (2019). Reporting of artificial intelligence prediction models. *Lancet (London, England)*, *393*(10181), 1577–1579. doi: 10.1016/S0140-6736(19)30037-6

Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297. doi: 10.1007/bf00994018

Dandash, O., Yücel, M., Daglas, R., Pantelis, C., McGorry, P., Berk, M., & Fornito, A. (2018). Differential effect of quetiapine and lithium on functional connectivity of the striatum in first episode mania. *Translational Psychiatry*, *8*(1), 59. doi: 10.1038/s41398-018-0108-8

Desikan, R. S., Segonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., … Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, *31*(3), 968–980. doi: 10.1016/j.neuroimage.2006.01.021

Fernández, A., Al-Timemy, A. H., Ferre, F., Rubio, G., & Escudero, J. (2018). Complexity analysis of spontaneous brain activity in mood disorders: A magnetoencephalography study of bipolar disorder and major depression. *Comprehensive Psychiatry*, *84*, 112–117. doi: 10.1016/j.comppsych. 2018.03.015

Fleck, D. E., Ernest, N., Adler, C. M., Cohen, K., Eliassen, J. C., Norris, M., … Strakowski, S. M. (2017). Prediction of lithium response in first-episode mania using the LITHium Intelligent Agent (LITHIA): Pilot data and proof-of-concept. *Bipolar Disorder*, *19*(4), 259–272. doi: 10.1111/bdi.12507

Geller, B., Zimerman, B., Williams, M., Bolhofner, K., Craney, J. L., DelBello, M. P., & Soutullo, C. (2001). Reliability of the Washington University in St. Louis Kiddie Schedule for Affective Disorders and Schizophrenia (WASH-U-KSADS) mania and rapid cycling sections. *Journal of the American Academy of Child and Adolescent Psychiatry*, *40*(4), 450–455. doi: 10.1097/00004583-200104000-00014

Germana, C., Kempton, M. J., Sarnicola, A., Christodoulou, T., Haldane, M., Hadjulis, M., … Frangou, S. (2010). The effects of lithium and anticonvulsants on brain structure in bipolar disorder. *Acta Psychiatrica Scandinavica*, *122*(6), 481–487. doi: 10.1111/j.1600-0447.2010.01582.x

Giedd, J. N., Blumenthal, J., Jeffries, N. O., Castellanos, F. X., Liu, H., Zijdenbos, A., … Rapoport, J. L. (1999). Brain development during childhood and adolescence: A longitudinal MRI study. *Nature Neuroscience*, *2*(10), 861–863. doi: 10.1038/13158

Golland, P., & Fischl, B. (2003). Permutation tests for classification: Towards statistical significance in image-based studies. *Information Processing in Medical Imaging*, *18*, 330–341. doi: 10.1007/978-3-540-45087-0_28

Hazlett, H. C., Gu, H., Munsell, B. C., Kim, S. H., Styner, M., Wolff, J. J., … Statistical, A. (2017). Early brain development in infants at high risk for autism spectrum disorder. *Nature*, *542*(7641), 348–351. doi: 10.1038/nature21369

Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *Science*, *313*(5786), 504–507. doi: 10.1126/science.1127647

Hirschfeld, R. M., Keck, P. E., Jr., Kramer, M., Karcher, K., Canuso, C., Eerdekens, M., & Grossman, F. (2004). Rapid antimanic effect of risperidone monotherapy: A 3-week multicenter, double-blind, placebo-controlled trial. *American Journal of Psychiatry*, *161*(6), 1057–1065. doi: 10.1176/appi.ajp.161.6.1057

Hozer, F., Sarrazin, S., Laidi, C., Favre, P., Pauling, M., Cannon, D., … Houenou, J. (2020). Lithium prevents grey matter atrophy in patients with bipolar disorder: An international multicenter study. *Psychological Medicine*, *51*(7), 1201–1210. doi: 10.1017/S0033291719004112.

Keck, P. E., Marcus, R., Tourkodimitris, S., Ali, M., Liebeskind, A., & Saha, A., … Aripiprazole Study Group (2003). A placebo-controlled, double-blind study of the efficacy and safety of aripiprazole in patients with acute bipolar mania. *American Journal of Psychiatry*, *160*(9), 1651–1658. doi: 10.1176/appi.ajp.160.9.1651

Keck, P. E., Jr., Welge, J. A., Strakowski, S. M., Arnold, L. M., & McElroy, S. L. (2000). Placebo effect in randomized, controlled maintenance studies of patients with bipolar disorder. *Biological Psychiatry*, *47*(8), 756–761. doi: 10.1016/s0006-3223(99)00309-1

Keshavan, M. S., Bagwell, W. W., Haas, G. L., Sweeney, J. A., Schooler, N. R., & Pettegrew, J. W. (1994). Changes in caudate volume with neuroleptic treatment. *Lancet*, *344*(8934), 1434. doi: 10.1016/s0140-6736(94)90599-1

Ketter, T. A., Miller, S., Dell'Osso, B., & Wang, P. W. (2016). Treatment of bipolar disorder: Review of evidence regarding quetiapine and lithium. *Journal of Affective Disorder*, *191*, 256–273. doi: 10.1016/j.jad.2015.11.002

Ketter, T. A., & Wang, P. W. (2002). Predictors of treatment response in bipolar disorders: Evidence from clinical and brain imaging studies. *Journal of Clinical Psychiatry*, *63*(Suppl. 3), 21–25. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/11908918.

Kim, J., Calhoun, V. D., Shim, E., & Lee, J. H. (2016). Deep neural network with weight sparsity control and pre-training extracts hierarchical features and enhances classification performance: Evidence from whole-brain resting-state functional connectivity patterns of schizophrenia. *NeuroImage*, *124*(Pt A), 127–146. doi: 10.1016/j.neuroimage.2015.05.018

Kim, T. T., Dufour, S., Xu, C., Cohen, Z. D., Sylvia, L., Deckersbach, T., … Nierenberg, A. A. (2019). Predictive modeling for response to lithium and quetiapine in bipolar disorder. *Bipolar Disorder*, *21*(5), 428–436. doi: 10.1111/bdi.12752

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. doi: 10.1038/nature14539

Lei, D., Li, W., Tallman, M. J., Patino, L. R., McNamara, R. K., Strawn, J. R., … DelBello, M. P. (2021). Changes in the brain structural connectome after a prospective randomized clinical trial of lithium and quetiapine treatment in youth with bipolar disorder. *Neuropsychopharmacology*, *46*(7), 1315–1323. doi: 10.1038/s41386-021-00989-5

Lei, D., Pinaya, W. H. L., van Amelsvoort, T., Marcelis, M., Donohoe, G., Mothersill, D. O., … Mechelli, A. (2020). Detecting schizophrenia at the level of the individual: Relative diagnostic value of whole-brain images, connectome-wide functional connectivity and graph-based metrics. *Psychological Medicine*, *50*(11), 1852–1861. doi: 10.1017/s0033291719001934

Lewinsohn, P. M., Klein, D. N., & Seeley, J. R. (2000). Bipolar disorder during adolescence and young adulthood in a community sample. *Bipolar Disorder*, *2*(3 Pt 2), 281–293. doi: 10.1034/j.1399-5618.2000.20309.x

Lewinsohn, P. M., Seeley, J. R., Buckley, M. E., & Klein, D. N. (2002). Bipolar disorder in adolescence and young adulthood. *Child and Adolescent Psychiatric Clinics of North America*, *11*(3), 461–475, vii. doi: 10.1016/s1056-4993(02)00005-6

Lim, C. S., Baldessarini, R. J., Vieta, E., Yucel, M., Bora, E., & Sim, K. (2013). Longitudinal neuroimaging and neuropsychological changes in bipolar disorder patients: Review of the evidence. *Neuroscience and Biobehavioral Reviews*, *37*(3), 418–435. doi: 10.1016/j.neubiorev.2013.01.003

Luo, C., Xu, H., & Li, X. M. (2005). Quetiapine reverses the suppression of hippocampal neurogenesis caused by repeated restraint stress. *Brain Research*, *1063*(1), 32–39. doi: 10.1016/j.brainres.2005.09.043

Moore, G. J., Bebchuk, J. M., Parrish, J. K., Faulk, M. W., Arfken, C. L., Strahl-Bevacqua, J., & Manji, H. K. (1999). Temporal dissociation between lithium-induced changes in frontal lobe myo-inositol and clinical response in manic-depressive illness. *American Journal of Psychiatry*, *156*(12), 1902–1908. doi: 10.1176/ajp.156.12.1902

Moore, G. J., Cortese, B. M., Glitz, D. A., Zajac-Benitez, C., Quiroz, J. A., Uhde, T. W., … Manji, H. K. (2009). A longitudinal study of the effects of lithium treatment on prefrontal and subgenual prefrontal gray matter volume in treatment-responsive bipolar disorder patients. *Journal of Clinical Psychiatry*, *70*(5), 699–705. doi: 10.4088/JCP.07m03745

Nord, M., Nyberg, S., Brogren, J., Jucaite, A., Halldin, C., & Farde, L. (2011). Comparison of D_2 dopamine receptor occupancy after oral administration of quetiapine fumarate immediate-release and extended-release formulations in healthy subjects. *International Journal of Neuropsychopharmacology*, *14*(10), 1357–1366. doi: 10.1017/s1461145711000514

Passarotti, A. M., Sweeney, J. A., & Pavuluri, M. N. (2011). Fronto-limbic dysfunction in mania pre-treatment and persistent amygdala over-activity post-treatment in pediatric bipolar disorder. *Psychopharmacology*, *216*(4), 485–499. doi: 10.1007/s00213-011-2243-2

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., … Chintala, S. J. a. e.-p. (2019). PyTorch: An imperative style, high-performance deep learning library. arXiv:1912.01703. Retrieved from https://arxiv.org/abs/1912.01703.

Patino, L. R., Klein, C. C., Strawn, J. R., Blom, T. J., Tallman, M. J., Adler, C. M., … DelBello, M. P. (2021). A randomized, double-blind, controlled trial of lithium versus quetiapine for the treatment of acute mania in youth with early course bipolar disorder. *Journal of Child and Adolescent Psychopharmacology*, *31*(7), 485–493. doi: 10.1089/cap.2021.0039

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., … Louppe, G. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research*, *12*, 2852–2830.

Perlis, R. H., Miyahara, S., Marangell, L. B., Wisniewski, S. R., Ostacher, M., DelBello, M. P., … Investigators, S.-B. (2004). Long-term implications of early onset in bipolar disorder: Data from the first 1000 participants in the systematic treatment enhancement program for bipolar disorder (STEP-BD). *Biological Psychiatry*, *55*(9), 875–881. doi: 10.1016/j.biopsych.2004.01.022

Pinaya, W. H., Gadelha, A., Doyle, O. M., Noto, C., Zugman, A., Cordeiro, Q., … Sato, J. R. (2016). Using deep belief network modelling to characterize differences in brain morphometry in schizophrenia. *Scientific Reports*, *6*, 38897. doi: 10.1038/srep38897

Rakic, P. (2009). Evolution of the neocortex: A perspective from developmental biology. *Nature Review Neuroscience*, *10*(10), 724–735. doi: 10.1038/nrn2719

Robinson, J. L., Bearden, C. E., Monkul, E. S., Tordesillas-Gutierrez, D., Velligan, D. I., Frangou, S., & Glahn, D. C. (2009). Fronto-temporal dysregulation in remitted bipolar patients: An fMRI delayed-non-match-to-sample (DNMS) study. *Bipolar Disorders*, *11*(4), 351–360. doi: 10.1111/j.1399-5618.2009.00703.x

Tamnes, C. K., Herting, M. M., Goddings, A. L., Meuwese, R., Blakemore, S. J., Dahl, R. E., … Mills, K. L. (2017). Development of the cerebral cortex across adolescence: A multisample study of inter-related longitudinal changes in cortical volume, surface area, and thickness. *Journal of Neuroscience*, *37*(12), 3402–3412. doi: 10.1523/JNEUROSCI.3302-16.2017

Vieira, S., Gong, Q. Y., Pinaya, W. H. L., Scarpazza, C., Tognin, S., Crespo-Facorro, B., … Mechelli, A. (2020). Using machine learning and structural neuroimaging to detect first episode psychosis: Reconsidering the evidence. *Schizophrenia Bulletin*, *46*(1), 17–26. doi: 10.1093/schbul/sby189

Vieira, S., Pinaya, W. H., & Mechelli, A. (2017). Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications. *Neuroscience and Biobehavioral Reviews*, *74*(Pt A), 58–75. doi: 10.1016/j.neubiorev.2017.01.002

https://doi.org/10.1017/S0033291722000757 Published online by Cambridge University Press

Vošahlíková, M., Roubalová, L., Brejchová, J., Alda, M., & Svoboda, P. (2021). Therapeutic lithium alters polar head-group region of lipid bilayer and prevents lipid peroxidation in forebrain cortex of sleep-deprived rats. *Biochimica et Biophysica Acta – Molecular and Cell Biology of Lipids*, *1866*(9), 158962. doi: 10.1016/j.bbalip.2021.158962

Wade, B. S., Joshi, S. H., Njau, S., Leaver, A. M., Vasavada, M., Woods, R. P., … Narr, K. L. (2016). Effect of electroconvulsive therapy on striatal morphometry in major depressive disorder. *Neuropsychopharmacology*, *41*(10), 2481–2491. doi: 10.1038/npp.2016.48

Wegbreit, E., Ellis, J. A., Nandam, A., Fitzgerald, J. M., Passarotti, A. M., Pavuluri, M. N., & Stevens, M. C. (2011). Amygdala functional connectivity predicts pharmacotherapy outcome in pediatric bipolar disorder. *Brain Connectivity*, *1*(5), 411–422. doi: 10.1089/brain.2011.0035

Wolfers, T., Buitelaar, J. K., Beckmann, C. F., Franke, B., & Marquand, A. F. (2015). From estimating activation locality to predicting disorder: A review of pattern recognition for neuroimaging-based psychiatric diagnostics. *Neuroscience and Biobehavioral Reviews*, *57*, 328–349. doi: 10.1016/j.neubiorev.2015.08.001

Woo, C. W., Chang, L. J., Lindquist, M. A., & Wager, T. D. (2017). Building better biomarkers: Brain models in translational neuroimaging. *Nature Neuroscience*, *20*(3), 365–377. doi: 10.1038/nn.4478

Yang, J., Lei, D., Qin, K., Pinaya, W. H. L., Suo, X., Li, W., … Gong, Q. (2021). Using deep learning to classify pediatric posttraumatic stress disorder at the individual level. *BMC Psychiatry*, *21*(1), 789. doi: 10.1186/s12888-021-03503-9.

Young, R. C., Biggs, J. T., Ziegler, V. E., & Meyer, D. A. (1978). A rating scale for mania: Reliability, validity and sensitivity. *British Journal of Psychiatry*, *133*, 429–435. doi: 10.1192/bjp.133.5.429

Zhang, W., Xiao, Y., Sun, H., Patino, L. R., Tallman, M. J., Weber, W. A., … DelBello, M. P. (2018). Discrete patterns of cortical thickness in youth with bipolar disorder differentially predict treatment response to quetiapine but not lithium. *Neuropsychopharmacology*, *43*(11), 2256–2263. doi: 10.1038/s41386-018-0120-y

Zhu, Z., Lei, D., Qin, K., Suo, X., Li, W., Li, L., … Gong, Q. (2021). Combining Deep learning and graph-theoretic brain features to detect posttraumatic stress disorder at the individual level. *Diagnostics*, *11*(8), 1416. doi: 10.3390/diagnostics11081416.

https://doi.org/10.1017/S0033291722000757 Published online by Cambridge University Press