ARTICLE

Discrete patterns of cortical thickness in youth with bipolar disorder differentially predict treatment response to quetiapine but not lithium

Wenjing Zhang [1], Yuan Xiao[1], Huaiqiang Sun[1], L. Rodrigo Patino[2], Maxwell J. Tallman[2], Wade A. Weber[2], Caleb M. Adler[2], Christina Klein[2], Jeffrey R. Strawn[2], Fabiano G. Nery[2], Qiyong Gong[1], John A. Sweeney[1,2], Su Lui[1] and Melissa P. DelBello[2]

The need for treatment response predictive biomarkers is being increasingly recognized in children and adolescents with psychiatric disorders. Structural gray matter abnormalities as a predictor of treatment outcome in pediatric bipolar disorder have not been systematically investigated, especially early in the illness course. With a prospective longitudinal study design, the present study enrolled 52 bipolar adolescents with no history of treatment with mood stabilizers or a therapeutic dose of antipsychotic drugs and 31 healthy controls. Patients were randomly assigned to treatment with quetiapine or lithium after pretreatment data collection. A hierarchical cluster analysis was performed using pretreatment cortical thickness data that identified two discrete patient subgroups. Compared to healthy subjects, patients in subgroup 1 ($n = 16$) showed widespread greater cortical thickness mainly across heteromodal cortex but also involving some regions of unimodal cortex, while those in subgroup 2 ($n = 36$) showed regional cortical thinning mainly in superior temporal and superior parietal regions. Patients within subgroup 1 showed a significantly higher response rate to quetiapine than those in subgroup 2 (100% vs 53%). No statistically significant difference was found in lithium response rate between the patient subgroups (63% vs 53%). Pretreatment clinical ratings and neuropsychological data did not differ across subgroups. Our findings suggest the existence of distinct and clinically relevant subgroups of pediatric bipolar patients, as defined by pattern of cortical thickness. These groups appear to differentially respond to antipsychotic treatment—notably with greater cortical thickness relative to controls predicting better treatment response.

INTRODUCTION

In individuals experiencing recent-onset bipolar disorder, choosing an effective pharmacological intervention can be challenging [1, 2]. Neurobiological heterogeneity of this disorder may account for the inconsistency of treatment outcomes [3, 4]. Thus, differentiating biologically discrete patient subgroups and identifying neural biomarkers relevant to treatment outcome is a promising research pathway to both facilitate mechanistic understanding of the illness and to develop individualized treatment strategies to improve outcomes.

Identifying neuroimaging biomarkers of treatment outcome in bipolar disorder has received increased attention in the past 2 decades [5, 6]. For example, pretreatment measurements of regional brain activation have been reported to predict effects of second-generation antipsychotics [7], while functional connectivity abnormalities may predict treatment response in bipolar patients treated with mood stabilizers [8]. Magnetic resonance spectroscopy (MRS) parameters have also been used with some success in predicting lithium treatment outcome [9], and post-treatment gray matter volume has been used to differentiate lithium responders and non-responders [10]. However, pretreatment brain morphometric measures (i.e., cortical mapping measures), which are more like trait-related features than functional measures [11], have been largely unexplored for this purpose.

To date, most studies of neuroimaging-based predictors of treatment outcome examined chronically treated patients or adults or mixed pediatric and adult patient samples. Studying patients early in their illness course who also have limited psychotropic exposure can minimize the confounding effects of illness course and medication on functional and structural brain measures. For example, effects of both lithium and antipsychotics on brain anatomy and function are well established [12–15]. Moreover, studying imaging biomarkers in pediatric patients is important because biomarkers may be differentially expressed in pediatric and adult patients [16]. However, no structural neuroimaging studies have investigated the utility of pretreatment anatomic scans for predicting treatment outcome in pediatric bipolar disorder. While previous studies determined the predictive potential of pretreatment measures in the whole patient sample, an alternative strategy is to first resolve neurobiological heterogeneity based on neuroanatomical features, as in previous studies of other disorders [17, 18], and then investigate the clinical relevance of using pretreatment MRI data to predict post-treatment clinical status in each identified patient subgroup. Conducting such studies as part of randomized clinical trials (RCT)

[1]Huaxi MR Research Center (HMRRC), Department of Radiology, The Center for Medical Imaging, West China Hospital of Sichuan University, 610041 Chengdu, China and [2]Department of Psychiatry and Behavioral Neuroscience, University of Cincinnati College of Medicine, Cincinnati, OH 45219, USA
Correspondence: Su Lui (lusuwcums@hotmail.com)

Discrete patterns of cortical thickness in youth with bipola…
W Zhang et al.

2257

is important to more readily interpret relations of pretreatment MRI parameters to treatment outcomes.

With these considerations in mind, we recruited a cohort of young patients with bipolar disorder who were early in their illness course into a prospective randomized clinical trial to investigate the potential of pretreatment neuroanatomic measures for predicting treatment outcome. A data-driven cluster analysis method was adopted to identify discrete patient subgroups using pretreatment quantitative cortical thickness measures, which are considered heritable and relatively stable structural brain characteristics [19]. After identifying discrete subgroups within the patient sample, we determined whether the subgroups differed in their response to specific pharmacotherapy.

MATERIALS AND METHODS
Participants
This study was approved by the University of Cincinnati Institutional Review Board. All study participants and their legal guardians provided written informed consent/assent after study procedures were fully explained. Fifty-two early course pediatric patients with bipolar I disorder (DSM-IV-TR criteria), and 31 healthy comparison subjects were recruited from the Cincinnati Children's Hospital Medical Center (CCHMC) and the University of Cincinnati Medical Center. Diagnoses of bipolar I disorder were confirmed by trained raters with established diagnostic reliability (kappa >0.9) via administration of the Washington University in St. Louis Kiddie Schedule of Affective Disorders and Schizophrenia (WASH-U-KSADS) [20]. Mood symptoms were rated using the Young Mania Rating Scale (YMRS) [21], Children's Depression Rating Scale-Revised (CDRS-R) [22], and Clinical Global Impressions-Severity (CGI-S) [23]. Parental socioeconomic status (SES) was evaluated by the Hollingshead-Redlich scale [24].

The age range for inclusion was 10–18 years. Patients were included if they were experiencing a manic or mixed episode, had a baseline YMRS score ≥20, and were less than 2 years from onset of bipolar disorder as defined by first mood episode. They had no prior psychiatric hospitalizations, no history of treatment with therapeutic doses of antipsychotic drugs, no history of treatment with mood stabilizers, and no psychotropic medication during the week (72 h for psychostimulants) prior to the MRI scanning and index psychiatric assessment. Patients could have had prior ADHD treatment or up to 3 months of prior antidepressant treatment, since excluding these patients would significantly limit the generalizability of our findings. Demographically matched healthy adolescents were recruited from the communities in which the bipolar participants resided, and were screened to ascertain the lifetime absence of psychiatric and neurological illness. They had no known history of affective or psychotic disorder among their first- or second-degree relatives. All participants were at Tanner stage III–V [25], in order to include only post-pubescent subjects and minimize brain changes associated with the onset of puberty [26].

The following exclusion criteria applied to both groups: (1) contraindication to MRI scanning (e.g., braces or claustrophobia); (2) IQ <70, as determined by the Wechsler Abbreviated Scale of Intelligence [27]; (3) a positive pregnancy test; (4) a history of 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); and (6) a lifetime DSM-IV-TR diagnosis of any pervasive developmental disorder.

Data acquisition at baseline
MRI data. MRI examinations were performed on a 4-T Varian Unity INOVA scanner with a 12-channel head coil. Earplugs and headphones were provided to block 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 4-T Varian scanner (T_{au} (magnetization preparation time) = 1.1 s, TR = 13 ms, TE = 5.3 ms, field of view = 256 × 192 × 192 mm, matrix = 256 × 192 × 96, flip angle = 20°, slice thickness = 2 mm. T1-weighted images of brain were inspected by two experienced neuroradiologists, and no scanning artifacts or gross brain abnormalities were observed in any participant.

Neurocognitive data. Three neuropsychological tests were administered prior to treatment from the Delis-Kaplan Executive Function System (D-KEFS) [28]: the trail making test, a verbal fluency test and the color-word interference test. Statistical analysis focused on number letter switching from the trail making test, letter fluency, and category fluency from the verbal fluency test, and inhibition scores from the color-word interference test.

Treatment procedures and post-treatment information
Following clinical evaluation and MRI scanning, patients were randomized, by an investigational pharmacist, to double-blind treatment with quetiapine or lithium, and evaluated weekly for 6 weeks. The randomization schedule was stratified by presence vs absence of ADHD, presence vs absence of psychosis, and the mood state (i.e., a mixed vs manic episode).

Quetiapine was initiated at 100 mg/day 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, and quetiapine/placebo as well as lithium/placebo capsules were identical. 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 physician monitoring trough lithium levels and making dose adjustments independent from treating physicians and clinical raters. However, blinded clinical tolerability rater dose adjustment recommendations took precedence over unblinded physician double-dummy dose adjustment recommendations. Acute treatment outcome was assessed using scores from the YMRS; responders were identified based on a ≥50% reduction in YMRS scores from baseline to end point [8, 29].

MRI data preprocessing
Cortical modeling and segmentation of structural MRI data were performed with FreeSurfer software (version 5.3.0, http://surfer.nmr.mgh.harvard.edu/), and details are in Supplementary Materials.

Feature extraction, cluster analysis, and cluster validation
The cortical thickness measures across the 34 cortical regions in each hemisphere from the Desikan/Killiany Atlas [30] were selected as structural features for the cluster analysis. We utilized a data-driven method of agglomerative clustering [31] to identify discrete homogeneous subgroups of patients based on their neuroanatomic scan data. As the existence and number of discrete subgroups of bipolar patients is unknown, hierarchical clustering was performed as it does not require an "a priori" decision about number of clusters [31, 32].

Agglomerative hierarchical clustering was performed using in-house Matlab code. Euclidean distance was used as the distance metric between subjects, and average distance between clusters was used as the linkage function [32]. In the cluster procedure, Euclidean distance was first calculated between subjects, and then pairs of subjects that were in close proximity were linked into binary clusters. The newly formed clusters were grouped into larger clusters in an iterative way until a hierarchical tree was formed.

Discrete patterns of cortical thickness in youth with bipola…
W Zhang et al.

2258

Table 1. Demographic and clinical characteristics of bipolar subgroups and healthy comparisons

Clinical and demographic variables	Bipolar subgroup 1 (N = 16) Mean (SD)	Bipolar subgroup 2 (N = 36)	Healthy controls (HC, N = 31)	Statistics	p	Post hoc analysis		
						Subgroup 1 vs HC	Subgroup 2 vs HC	Subgroup 1 vs subgroup 2
Age (years)	14.09 (1.71)	14.66 (1.76)	15.02 (1.69)	$F = 1.53$	0.22			
IQ	110.13 (12.11)	100.03 (11.68)	110.03 (11.45)	$F = 7.53$	<0.01*	0.99	<0.01*	0.01*
Parental SES	3.23 (1.01)	2.95 (1.05)	3.90 (1.01)	$F = 7.71$	<0.01*	0. 10	<0.01*	0.53
Duration of episode (days)	29.81 (36.51)	17.09 (27.04)	NA	$t = 1.26$	0.22			
CDRS-R (Baseline)	39.75 (7.02)	37.97 (8.33)	NA	$t = 0.74$	0.46			
CGI-S (Baseline)	4.94 (0.57)	5.03 (0.65)	NA	$t = 0.48$	0.64			
YMRS (Baseline)	28.06 (4.73)	28.94 (4.64)	NA	$t = 0.63$	0.53			
YMRS (End point)	10.31 (6.72)	14.94 (8.56)	NA	$t = 2.09$	0.04*			
Sex, N (%)								
Male	7 (44%)	13 (36%)	17 (55%)	$\chi^2 = 0.53$	0.77			
Female	9 (56%)	23 (64%)	14 (45%)					
ADHD comorbidity, N (%)	4 (25%)	16 (44%)	NA	$\chi^2 = 1.77$	0.18			
Current psychosis, N (%)	3 (19%)	7 (19%)	NA	$\chi^2 < 0.01$	0.95			
History of ADHD treatment, N (%)	6 (38%)	12 (33%)	NA	$\chi^2 = 0.09$	0.77			
History of antidepressant treatment, N (%)	2 (13%)	7 (19%)	NA	$\chi^2 = 0.60$	0.44			
History of antipsychotic treatment, N (%)	4 (25%)	12 (33%)	NA	$\chi^2 = 0.36$	0.55			
Therapeutics and response						**Effect size estimates (Cohen's d)**		
Dosage (mg)								
Lithium, mean (SD)	1237.50 (250.36)	1270.59 (492.14)		$t = 0.22$	0.83		−0.08	
Quetiapine, mean (SD)	462.50 (106.07)	505.88 (134.49)		$t = 0.80$	0.43		−0.36	
Response rate								
Lithium, N (%)	5/8 (63%)	9/17 (53%)		$\chi^2 = 0.20$	0.65		0.18	
Quetiapine, N (%)	8/8 (100%)	10/19 (53%)		$\chi^2 = 5.68$	0.02*		1.03	
General, N (%)	13/16 (81%)	19/36 (53%)		$\chi^2 = 3.39$	0.051		0.53	

ADHD attention deficit hyperactivity disorder, *CDRS-R* Children's Depression Rating Scale-Revised, *CGI-S* Clinical Global Impressions-Severity, *HC* healthy controls, *IQ* intelligence quotient, *NA* not available, *SD* standard deviation, *SES* socioeconomic status, *YMRS* Young Mania Rating Scale
*The *p* with statistical significance

Cluster quality and the optimal cluster number were determined using the Silhouette [33] and Dunn indices [34], which reflect the compactness and separation of clusters [35]. The Silhouette index reflects the compactness and separation of clusters; higher values indicate greater cluster delineation. The Dunn index is the ratio of the smallest distance between samples not in the same cluster to the largest intra-cluster distance. Larger values of Dunn index correspond to better cluster quality, and the number of clusters that maximizes the Dunn index is taken as the optimal number of clusters.

The stability of the cluster solution was tested using a bootstrap technique. The Jaccard coefficient was calculated as the similarity between resampled clusters with those derived from the primary clustering analysis to validate our clustering results [36]. Details of these procedures and analyses are presented in Supplementary Materials.

Statistical analysis
Primary hypothesis testing focused on determining whether identified subgroups of patients, based on their brain anatomic

scans, responded differently to quetiapine and lithium therapy. An analysis of variance (ANOVA) examined changes from pre- to post-treatment YMRS scores between patient subgroups (group-by-treatment). Chi-square tests compared response rate for different therapeutics. As the subgroups differed in IQ and parental SES, to ensure there were no complex confounding effects of these measures, we conducted a correlation analysis between IQ, parental SES, and the YMRS reduction (percentage) in the whole bipolar patient cohort, and the findings are presented in Supplementary Materials.

To identify regional pretreatment differences in cortical thickness profiles that contributed significantly to separating the participants with bipolar disorder into discrete subgroups, and the relationship of these measures to those of healthy controls, we compared the average thickness measures in each of the 68 brain regions of interest using an ANOVA with step-down post hoc tests for significant differences between controls and two patient subgroups. Age, sex, IQ, and parental SES were included as covariates. Testing for group differences in the 68 ANOVAs was performed with Bonferroni correction to control for
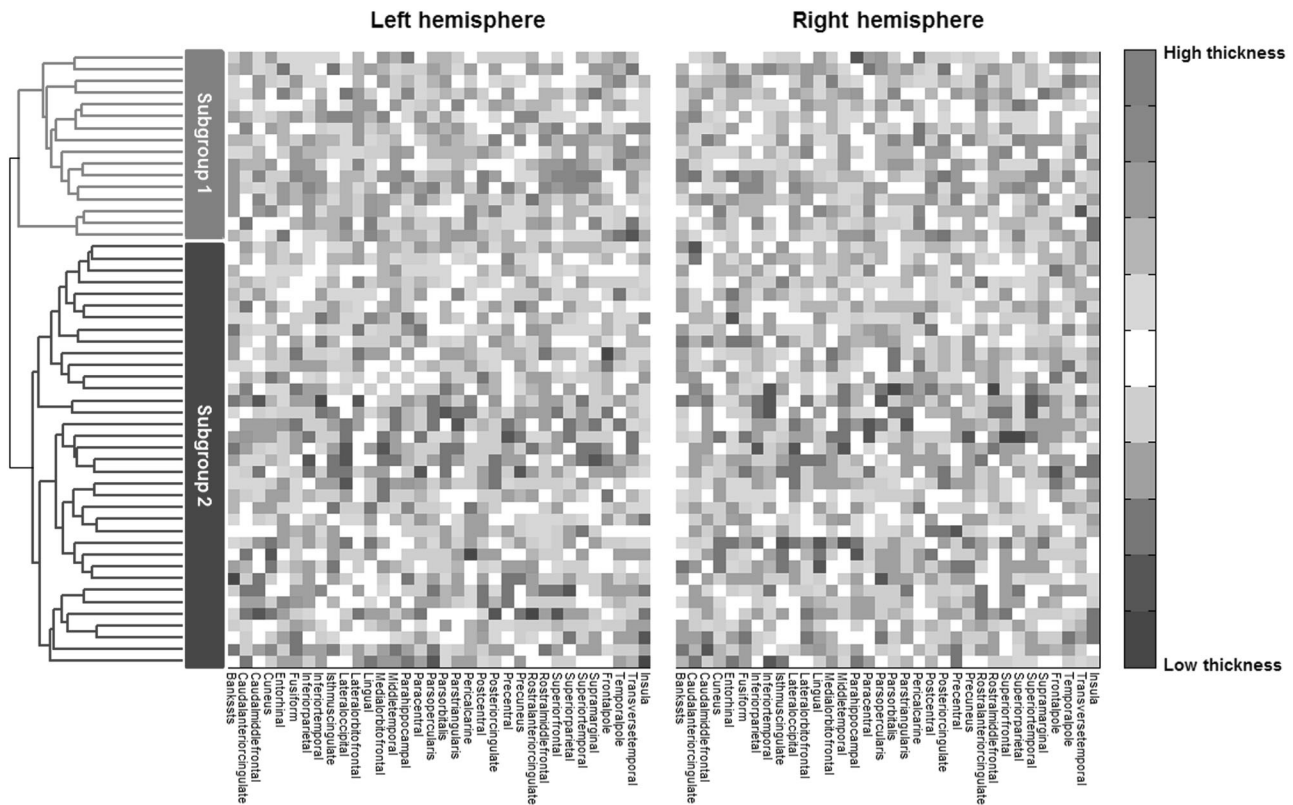
Discrete patterns of cortical thickness in youth with bipola…
W Zhang et al.

2259

Fig. 1 Dendrogram and heat map of the hierarchical clustering in youth with bipolar disorder based on cortical thickness measures

multiple comparisons. Post hoc pairwise analysis of regions with significant overall group differences was corrected with the false discovery rate (FDR) to preserve a $p < 0.05$ experiment-wise threshold.

In an exploratory analysis, to examine potential associations of clinical and cognitive variables with cortical morphometry in patients, partial correlation analyses were conducted between cortical thickness measures in regions with significant inter-group differences and psychopathologic and cognitive measures after covarying for age, sex, IQ, and parental SES. These analyses were done separately for each identified patient subgroup and are presented for heuristic purposes without type 1 error correction.

Given that aging effect might be interesting in adolescent patients, regression analyses using a linear model of age effects in relation to altered cortical thickness measures in each bipolar subgroup were conducted. These models were then compared to those of healthy controls to determine whether there was a significant differential rate of age-related change in each patient subgroup. Age-by-diagnosis interaction on cortical thickness differences across all the 68 regions in comparison to healthy controls was also calculated for each bipolar subgroup.

RESULTS
Demographics and clinical variables
Demographic characteristics and clinical features of study participants are presented in Table 1. Twenty patients had comorbid attention deficit hyperactivity disorder (ADHD), while ten subjects had psychosis. Among patients, 34.6% of participants had prior exposure to psychostimulant medications while 17.3% of them had up to 3 months of prior antidepressant treatment. A total of 30.7% of participants had less than ten lifetime doses of antipsychotic medication in less than 3 months that did not achieve therapeutic dosage.

With regard to randomized treatments in the longitudinal trial, 25 patients received lithium while the other 27 received quetiapine.

Hierarchical clustering
The results of hierarchical clustering of cortical thickness data are shown as a combination of dendrogram and heat map illustrations in Fig. 1. We evaluated the dendrogram from two to ten cluster solutions with Silhouette and Dunn indices. Both parameters reached their maximum (Silhouette index = 0.112 and Dunn index = 10.25) in the two cluster solution (see Supplementary Figure S1). The bootstrapping stability test also showed that the Jaccard coefficient achieved the highest value of 0.812 when the cluster number was 2 (see Supplementary Figure S2). Thus, the optimal and most stable number of discrete data structures that best represents the data for this patient group is two, and they are described as subgroups 1 and 2 below.

There were no statistically significant differences in age, sex, and parental SES between patient subgroups (see Table 1). Visual inspection of the two patient subgroups from the dendrogram and heat map illustrates that patients within subgroup 1 (16 patients, 30.8% of sample) had greater cortical thickness than patients comprising subgroup 2 (36 patients, 69.2% of sample). Statistical comparisons of MRI data of the two subgroups and controls are presented below.

Differences in clinical and cognitive ratings between patient subgroups
There were no significant differences between patient subgroups 1 and 2 in pretreatment YMRS, CGI-S, or CDRS-R scores (see Table 1), or in duration of current episode or number of prior mood episodes. The rates of ADHD comorbidity, current psychosis, and prior antidepressant, antipsychotic, or ADHD treatment did not differ between patient subgroups. However, there were

Discrete patterns of cortical thickness in youth with bipola…
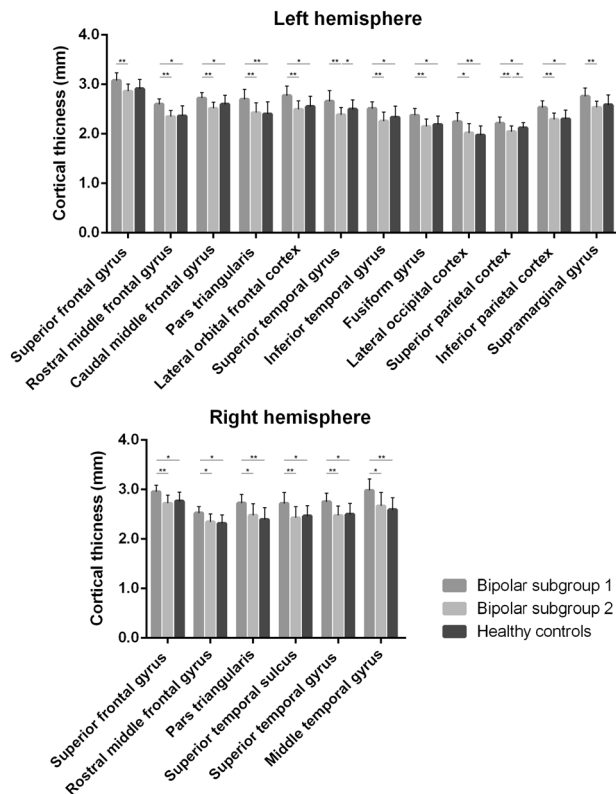W Zhang et al.

2260

Fig. 2 Region-wise cortical differences among patient subgroups and healthy controls within the 68 regions examined. * indicates regions with significant inter-group differences after post hoc analysis between groups

statistically significant differences in IQ score ($F = 7.53$, $p < 0.01$) and parental SES ($F = 7.71$, $p < 0.01$) among the two patient subgroups and healthy controls. IQ was lower in subgroup 2 compared to subgroup 1 ($p = 0.01$) and healthy controls ($p < 0.01$), while IQ scores did not differ between subgroup 1 patients and controls (see Table 1). Parental SES was lower in subgroup 2 compared to healthy controls ($p < 0.01$), but did not differ between subgroup 1 patients and healthy subjects. Pretreatment neurocognitive parameters did not statistically differ between patient subgroups, even after controlling for IQ and parental SES (see Supplementary Table S1).

Differences in treatment outcome between patient subgroups
In patient subgroup 1, eight patients were treated with lithium while the other eight were treated with quetiapine. In patient subgroup 2, 17 patients received lithium treatment while 19 patients received quetiapine. The medication dosage at the end point did not statistically differ between patient subgroups for either lithium ($p = 0.83$) or quetiapine ($p = 0.43$). However, for those treated with quetiapine, patients within subgroup 1 achieved a higher rate of treatment response relative to those in subgroup 2 (100% vs 53%, $p = 0.02$, effect size = 1.03). In lithium-treated patients, response rates did not significantly differ between subgroups (63% vs 53%, $p = 0.65$, effect size = 0.18).

Regardless of the therapeutics, there was a trend of higher general response rate for patients in subgroup 1 than patients in subgroup 2 (81% vs 53%, $p = 0.051$, effect size = 0.53). The ANOVA also revealed a greater decrease of general YMRS score from baseline to end point in subgroup 1 compared to subgroup 2 ($F = 3.95$, $p < 0.05$). But, as analyses above show these differences were primarily due to treatment effects in the quetiapine treatment group.

Regional differences in cortical thickness among patient subgroups and healthy controls
ANOVAs were conducted comparing the two patient subgroups and controls for each of the 68 ROI examined. This was done to identify the types and regional distribution of cortical thickness changes that discriminated participant groups. Significant group differences were seen in bilateral superior frontal gyrus, bilateral rostral middle frontal gyrus, bilateral pars triangularis, bilateral superior temporal gyrus, left fusiform gyrus, left inferior parietal cortex, left lateral occipital cortex, left lateral orbital frontal cortex, left caudal middle frontal gyrus, left superior parietal cortex, left supramarginal gyrus, left inferior temporal gyrus, as well as right superior temporal sulcus and right middle temporal gyrus ($p < 0.05$, Bonferroni corrected).

Post hoc pairwise comparisons showed that relative to healthy controls and patients in subgroup 2, patients in subgroup 1 showed thicker cortex in right superior frontal gyrus, bilateral rostral middle frontal gyrus, bilateral pars triangularis, right superior temporal gyrus, left inferior parietal cortex, left caudal middle frontal gyrus, left lateral orbital frontal cortex, left fusiform gyrus, left superior parietal cortex, left lateral occipital cortex, left inferior temporal gyrus, as well as right superior temporal sulcus and right middle temporal gyrus ($p < 0.05$, FDR corrected). Patients in this subgroup did not show any region with decreased cortical thickness relative to healthy comparisons.

Patients in subgroup 2, in contrast with healthy controls, displayed reduced cortical thickness in left superior temporal gyrus and left superior parietal cortex ($p < 0.05$, FDR corrected, see Supplementary Table S2, Table S3, and Fig. 2).

To confirm that the differential treatment outcome mainly derived from differences in cortical thickness not subcortical volumes, we also compared the volumes of bilateral thalamus, hippocampus, amygdala, caudate, putamen, pallidum, and accumbens areas between bipolar subgroups and found no significant inter-group differences. Responders and non-responders also did not differ in any subcortical volume measurement in either lithium or quetiapine treatment groups. Details are in Supplementary Materials.

Correlation between altered cortical thickness and clinical ratings in each patient subgroup
In patient subgroup 1, a small number of nominally significant associations were found with pretreatment cortical thickness data. Cortical thickness of left caudal middle frontal gyrus was negatively associated with CGI-S scores ($r = -0.66$, $p = 0.019$), while cortical thickness of left orbital frontal gyrus was positively associated with CGI-S scores ($r = 0.76$, $p = 0.004$). More importantly, pretreatment cortical thickness of left pars triangularis was positively associated with percent YMRS reduction ($r = -0.62$, $p = 0.032$). In patient subgroup 2, no significant correlations were found between cortical thickness of any region and psychopathologic ratings at baseline or end point.

Comparison of age-related changes among participant groups
In comparisons of age-related changes of altered cortical thickness in each patient subgroup, no significant difference was found between each patient subgroup and healthy controls, or between patient subgroups (see Supplementary Figure S3).

Age-by-diagnosis interaction on cortical thickness across all 68 regions of each bipolar subgroup in comparison to healthy controls also showed no significant findings (see Supplementary Table S4 and Table S5). Details are presented in Supplementary Materials.

DISCUSSION
The current study identified two distinct cortical thickness patterns in patients with pediatric bipolar disorder who were

Discrete patterns of cortical thickness in youth with bipola...
W Zhang et al.

2261

early in their illness course: compared to healthy controls, one group of patients exhibited widespread increases in thickness of the cortical mantle mainly in heteromodal association cortex but also involving some regions of unimodal cortex, while a second group showed regionally decreased cortical thickness in superior temporal and superior parietal regions. Subcortical volumes did not differ between these two patient subgroups. While the two patient subgroups did not show different acute illness severity or cognitive alterations, which is consistent with previous studies [8, 9], subgroup 1 exhibited better response to quetiapine relative to subgroup 2. Thus, in patients early in their illness course and with minimal prior treatment exposure, our findings indicate the existence of two neurobiologically distinct biotypes of youth with bipolar disorder, which may differentially respond to antipsychotic treatment. More importantly, as in previous efforts [4, 17, 18], we defined discrete groups of patients based on neurobiological features, and further demonstrated the clinical relevance of this group separation by showing differences in treatment outcome in the bipolar patient subgroups who shared similar clinical syndromal characteristics. The potential significance of our findings, then, is that neuroanatomic measures of cortical gray matter may provide clinically useful predictors of differential treatment response for individualizing treatment in youth with bipolar disorder.

While the clinical utility of structural brain alterations that may potentially guide differential therapeutics requires replication, the present study represents a promising step forward addressing the two major challenges that need to be addressed in such efforts. We both successfully resolved neurobiological heterogeneity and established the clinical relevance of these neurobiologically-based patient classifications for differential therapeutics. Although interest in biological heterogeneity in psychiatric syndromes has gained increasing attention across psychiatry research, most previous studies of neural biomarkers defined patient taxonomies according to clinical categorization schemes, and then analyzed neurobiological imaging data in clinically defined subgroups [37]. However, symptom-based subtyping strategies have been criticized for their instability over time, because patients share similar neural system pathology across subtypes and diagnostic categories, and for the concern that existing clinical classifications may not delineate patients with biologically distinct characteristics [4, 38, 39]. Our finding of distinct subgroups of pediatric bipolar patients classified by structural neuroimaging measures is promising in resolving clinically relevant neurobiological heterogeneity within the bipolar syndrome. Although the mechanisms for distinct patterns of gray matter alteration remain to be determined, our findings represent a significant step forward both in methodological approach as well as in providing findings potentially relevant to understanding etiopathological heterogeneity and treatment strategies for youth with bipolar disorder. This is an important procedure in Psychoradiology, an evolving subspecialty of radiology focusing on psychiatric disorders.

It is noteworthy that the patients with greater cortical thickness mainly across heteromodal cortex, somewhat more noteworthy in frontal cortex, had better short-term clinical responses to quetiapine than bipolar patients with regional cortical thinning. The greater cortical thickness could represent disorder-related synaptic remodeling with increased synaptic proliferation associated with reduced synaptic pruning. Increased synaptic proliferation is related to less reduction of NAA in brain [40, 41], while a previous MRS study indicated that young bipolar disorder patients who responded to quetiapine had more *N*-acetylaspartate (NAA) in frontal cortex compared to non-remitters at baseline [42]. Greater cortical thickness could also be related to increased regional activity, and baseline activation of prefrontal cortex has been associated with better treatment outcome [43].

Clinical responses to lithium did not differ between patient subgroups identified by cortical thickness. This is not surprising since neuroimaging markers for lithium response prediction have mainly be reported in subcortical regions [44, 45]; however, in our pediatric sample the two patient subgroups did not differ significantly in volumes of subcortical regions that were measured. The lack of difference between responders and non-responders for both medications in subcortical volumes supports the notion that it is the cortical gray matter, not subcortical regions, that maintains the predictive value for treatment outcomes for pediatric patients early in the illness course. Though mechanisms for such effects in relation to outcome of different medications need to be determined via future clinical and preclinical research, exploration of why those with this alteration were more responsive to antipsychotic medication could be important for future drug development and differential clinical therapeutics.

The pattern of increased cortical thickness may reflect a distinctive brain maturational alteration or more state-related factors related to illness onset. Our findings are consistent with prior reports that some young bipolar patients close to illness onset show gray matter enlargements in both cortical and subcortical regions [40, 46, 47]. However, chronic adult patients exhibited widespread cortical thinning [48, 49]. These findings suggested that the pattern of increased cortical thickness in ~1/3 of our patients may be specific to pediatric onset cases or a pattern that transitions over the course of illness in some patients. Notably, our previous study of untreated first-episode depression patients also found widespread increases of cortical thickness [50]. Thus, we note that the increased cortical thickness pattern in early-onset bipolar cases may prove to be a non-specific finding across affective disorders at illness onset. This may relate to shared mechanisms such as neuroinflammation [51, 52]. In the early stage of neuroinflammation, astrocytes, which constitute 90% cortical tissue volume, can be activated by proinflammatory cytokines and lead to cellular hypertrophy, astrocyte proliferation, process extension, and interdigitation, which can increase cortical thickness [53]. Preapoptotic osmotic changes or other neurodevelopmental factor leading to neuropil increases may also account for the observed pattern of increased cortical thickness in some patients.

While the pattern of increased cortical thickness may have particular importance for a mechanistic understanding of illness and individualizing therapeutics, morphometric measures in the majority of our patients revealed regional cortical thinning relative to healthy controls in superior temporal and superior parietal regions. Abnormalities in temporal and parietal cortex have long been appreciated as potential neural substrates for abnormal responses to emotional stimuli central to bipolar disorder [40, 48]. This pattern is consistent with previous evidence that a significant subgroup of bipolar patients, notably those with lower cognitive abilities, demonstrated this type of dystrophic gray matter alteration [54, 55]. Distinct neural system alterations are presumably more relevant for illness onset in such patients.

While this is among the first work to evaluate neurobiological predictors of differential treatment response in youth with bipolar disorder who were experiencing a mixed or manic episode, several limitations require consideration. First, patient subgroups differed in IQ, and patient subgroup 2 had lower parental SES than healthy controls. However, the group with the most distinguishing MRI trait (i.e., thicker cortical mantle) was well-matched to the healthy controls. Second, our analyses revealed no general intellectual or cognitive differences between the two patient subgroups. However, we define two patient subgroups with different types of abnormality, rather than with more or less abnormal MRI characteristics. Identifying distinguishing cognitive features of patient subgroups defined by MRI data are important and remain a target for future research with a more comprehensive neurocognitive assessment designed to identify such characteristics. Finally, only about one-third of our sample (*n* =

2262

16) was clustered into the subgroup with thicker neocortex, and that sample size is not large, especially when these patients are further divided into two treatment groups. Therefore, the 100% response rate in the quetiapine-treated group should be considered with caution due to the small sample size. Replication in a larger study is clearly needed as longitudinal studies of identified patterns of structural brain alterations and studies of adult-onset cases. Studies of animal models will be needed to resolve the mechanisms for the association between the imaging observations and treatment outcome.

With due considerations of the limitations discussed above, our identification of a distinct subgroup of youth with bipolar disorder, defined by widespread increase of the cortical thickness, may represent significant progress in understanding neurobiological heterogeneity within bipolar disorder and its potential relevance for therapeutic interventions. Future studies using other neuroimaging, neurophysiological, and genetic information, as well as preclinical research, may provide important insights into the specific neuropathological substrate of the atypically increased thickness of the cortical mantle in a significant subgroup of youth with bipolar disorder early in their illness course.

ADDITIONAL INFORMATION
Supplementary Information accompanies this paper at (https://doi.org/10.1038/s41386-018-0120-y).

Competing interests: Dr. Patino has received research support through the American Academy of Child and Adolescent Psychiatry through the Junior Investigator Award program. Dr. Adler has received research support from Johnson & Johnson, Merck, Forest, Otsuka, Purdue, Takeda, Pfizer, Shire, Sunovion, and SyneuRx. He is a consultant to Sunovion. Dr. Strawn has received research support from Edgemont, Forest Research Institute/Allergan, Lundbeck, Shire, Neuronetics, and the National Institute of Mental Health (NIMH and NIEHS). He receives royalties from Springer Publishing and UpToDate and has received material support from Assurex. Dr. Nery's spouse is an employee of Eli Lilly & Co. Dr. DelBello has received research support from Amarex, Johnson & Johnson, Pfizer, Otsuka, Shire, Sunovion, Supernus, and Lundbeck. She is a consultant to Akili, CMEology, Johnson & Johnson, Lundbeck, Neuronetics, Pfizer, Sunovion, Supernus, and Takeda. The remaining authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

REFERENCES
1. Dunner DL. Clinical consequences of under-recognized bipolar spectrum disorder. Bipolar Disord. 2003;5:456–63.
2. Ghaemi SN, Lenox MS, Baldessarini RJ. Effectiveness and safety of long-term antidepressant treatment in bipolar disorder. J Clin Psychiatry. 2001;62:565–9.
3. Hozer F, Houenou J. Can neuroimaging disentangle bipolar disorder? J Affect Disord. 2016;195:199–14.
4. Clementz BA, Sweeney JA, Hamm JP, Ivleva EI, Ethridge LE, Pearlson GD, et al. Identification of distinct psychosis biotypes using brain-based biomarkers. Am J Psychiatry. 2016;173:373–84.
5. Ketter TA, Wang PW. Predictors of treatment response in bipolar disorders: evidence from clinical and brain imaging studies. J Clin Psychiatry. 2002;63:21–5.
6. Lim CS, Baldessarini RJ, Vieta E, Yucel M, Bora E, Sim K. Longitudinal neuroimaging and neuropsychological changes in bipolar disorder patients: review of the evidence. Neurosci Biobehav Rev. 2013;37:418–35.
7. 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. 2011;216:485–99.
8. Wegbreit E, Ellis JA, Nandam A, Fitzgerald JM, Passarotti AM, Pavuluri MN, et al. Amygdala functional connectivity predicts pharmacotherapy outcome in pediatric bipolar disorder. Brain Connect. 2011;1:411–22.
9. Fleck DE, Ernest N, Adler CM, Cohen K, Eliassen JC, Norris M, et al. Prediction of lithium response in first-episode mania using the LITHium Intelligent Agent (LITHIA): pilot data and proof-of-concept. Bipolar Disord. 2017;19:259–72.
10. Moore GJ, Cortese BM, Glitz DA, Zajac-Benitez C, Quiroz JA, Uhde TW, et al. A longitudinal study of the effects of lithium treatment on prefrontal and subgenual prefrontal gray matter volume in treatment-responsive bipolar disorder patients. J Clin Psychiatry. 2009;70:699–705.
11. Gong Q, Lui S, Sweeney JA. A selective review of cerebral abnormalities in patients with first-episode schizophrenia before and after treatment. Am J Psychiatry. 2016;173:232–43.
12. Giakoumatos CI, Nanda P, Mathew IT, Tandon N, Shah J, Bishop JR, et al. Effects of lithium on cortical thickness and hippocampal subfield volumes in psychotic bipolar disorder. J Psychiatr Res. 2015;61:180–7.
13. Keedy SK, Rosen C, Khine T, Rajarethinam R, Janicak PG, Sweeney JA. An fMRI study of visual attention and sensorimotor function before and after antipsychotic treatment in first-episode schizophrenia. Psychiatry Res. 2009;172:16–23.
14. Keshavan MS, Bagwell WW, Haas GL, Sweeney JA, Schooler NR, Pettegrew JW. Changes in caudate volume with neuroleptic treatment. Lancet. 1994;344:1434.
15. Lui S, Li T, Deng W, Jiang L, Wu Q, Tang H, et al. Short-term effects of antipsychotic treatment on cerebral function in drug-naive first-episode schizophrenia revealed by "resting state" functional magnetic resonance imaging. Arch Gen Psychiatry. 2010;67:783–92.
16. Brooks JO 3rd, Vizueta N. Diagnostic and clinical implications of functional neuroimaging in bipolar disorder. J Psychiatr Res. 2014;57:12–25.
17. Drysdale AT, Grosenick L, Downar J, Dunlop K, Mansouri F, Meng Y, et al. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. Nat Med. 2017;23:28–38.
18. Sun H, Lui S, Yao L, Deng W, Xiao Y, Zhang W, et al. Two patterns of white matter abnormalities in medication-naive patients with first-episode schizophrenia revealed by diffusion tensor imaging and cluster analysis. JAMA Psychiatry. 2015;72:678–86.
19. Panizzon MS, Fennema-Notestine C, Eyler LT, Jernigan TL, Prom-Wormley E, Neale M, et al. Distinct genetic influences on cortical surface area and cortical thickness. Cereb Cortex. 2009;19:2728–35.
20. Geller B, Zimerman B, Williams M, Bolhofner K, Craney JL, DelBello MP, et al. Reliability of the Washington University in St. Louis Kiddie Schedule for Affective Disorders and Schizophrenia (WASH-U-KSADS) mania and rapid cycling sections. J Am Acad Child Adolesc Psychiatry. 2001;40:450–5.
21. Young RC, Biggs JT, Ziegler VE, Meyer DA. A rating scale for mania: reliability, validity and sensitivity. Br J Psychiatry. 1978;133:429–35.
22. Poznanski EO, Mokros HB. Children's depression rating scale, revised (CDRS-R). Los Angeles: Western Psychological Services;1996.
23. Shaffer D, Gould MS, Brasic J, Ambrosini P, Fisher P, Bird H, et al. A children's global assessment scale (CGAS). Arch Gen Psychiatry. 1983;40:1228–31.
24. Hollingshead AB, Redlich FC. Social class and mental illness: community study. Am J Public Health. 1958;97:1756–7.
25. Duke PM, Litt IF, Gross RT. Adolescents' self-assessment of sexual maturation. Pediatrics. 1980;66:918–20.
26. Blakemore SJ, Burnett S, Dahl RE. The role of puberty in the developing adolescent brain. Hum Brain Mapp. 2010;31:926–33.
27. Stano J. Wechsler abbreviated scale of intelligence. San Antonio, TX: The Psychological Corporation; 1999.
28. Delis DC, Kaplan E, Kramer JH. Delis-Kaplan executive function system (D-KEFS). San Antonio, TX: Psychological Corporation; 2001.
29. Schneider MR, Adler CM, Whitsel R, Weber W, Mills NP, Bitter SM, et al. The effects of ziprasidone on prefrontal and amygdalar activation in manic youth with bipolar disorder. Isr J Psychiatry Relat Sci. 2012;49:112–20.
30. Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage. 2006;31:968–80.
31. Johnson SC. Hierarchical clustering schemes. Psychometrika. 1967;32:241–54.
32. Kaufman L, Rousseeuw PJ (2009). Finding groups in data: an introduction to cluster analysis. Vol 344. New York: John Wiley & Sons;2009.
33. Rousseeuw PJ. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 1987;20:53–65.
34. Dunn JC. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. J Cybernetics. 1973;3:32–57.

Discrete patterns of cortical thickness in youth with bipola…
W Zhang et al.

2263

35. Sharan R, Maron-Katz A, Shamir R. CLICK and EXPANDER: a system for clustering and visualizing gene expression data. Bioinformatics. 2003;19:1787–99.

36. Hennig C. Cluster-wise assessment of cluster stability. Comput Stat Data Anal. 2007;52:258–71.

37. Abi-Dargham A, Horga G. The search for imaging biomarkers in psychiatric disorders. Nat Med. 2016;22:1248–55.

38. Keshavan MS, Clementz BA, Pearlson GD, Sweeney JA, Tamminga CA. Reimagining psychoses: an agnostic approach to diagnosis. Schizophr Res. 2013;146:10–6.

39. Walker E, Kestler L, Bollini A, Hochman KM. Schizophrenia: etiology and course. Annu Rev Psychol. 2004;55:401–30.

40. Adler CM, DelBello MP, Jarvis K, Levine A, Adams J, Strakowski SM. Voxel-based study of structural changes in first-episode patients with bipolar disorder. Biol Psychiatry. 2007;61:776–81.

41. Perry VH, O'Connor V. The role of microglia in synaptic stripping and synaptic degeneration: a revised perspective. ASN Neuro. 2010;2:e00047.

42. Adler CM, DelBello MP, Weber WA, Jarvis KB, Welge J, Chu WJ, et al. Neurochemical effects of quetiapine in patients with bipolar mania: a proton magnetic resonance spectroscopy study. J Clin Psychopharmacol. 2013;33:528–32.

43. Pavuluri MN, Passarotti AM, Lu LH, Carbray JA, Sweeney JA. Double-blind randomized trial of risperidone versus divalproex in pediatric bipolar disorder: fMRI outcomes. Psychiatry Res. 2011;193:28–37.

44. Hallahan B, Newell J, Soares JC, Brambilla P, Strakowski SM, Fleck DE, et al. Structural magnetic resonance imaging in bipolar disorder: an international collaborative mega-analysis of individual adult patient data. Biol Psychiatry. 2011;69:326–35.

45. Selek S, Nicoletti M, Zunta-Soares GB, Hatch JP, Nery FG, Matsuo K, et al. A longitudinal study of fronto-limbic brain structures in patients with bipolar I disorder during lithium treatment. J Affect Disord. 2013;150:629–33.

46. Lisy ME, Jarvis KB, DelBello MP, Mills NP, Weber WA, Fleck D, et al. Progressive neurostructural changes in adolescent and adult patients with bipolar disorder. Bipolar Disord. 2011;13:396–405.

47. Adler CM, Levine AD, DelBello MP, Strakowski SM. Changes in gray matter volume in patients with bipolar disorder. Biol Psychiatry. 2005;58:151–7.

48. Hanford LC, Nazarov A, Hall GB, Sassi RB. Cortical thickness in bipolar disorder: a systematic review. Bipolar Disord. 2016;18:4–18.

49. Hibar DP, Westlye LT, Doan NT, Jahanshad N, Cheung JW, Ching CRK, et al. Cortical abnormalities in bipolar disorder: an MRI analysis of 6503 individuals from the ENIGMA Bipolar Disorder Working Group. Mol Psychiatry. 2018;23:932–42.

50. Qiu L, Lui S, Kuang W, Huang X, Li J, Li J, et al. Regional increases of cortical thickness in untreated, first-episode major depressive disorder. Transl Psychiatry. 2014;4:e378.

51. Echeverria V, Grizzell JA, Barreto GE. Neuroinflammation: a therapeutic target of cotinine for the treatment of psychiatric disorders? Curr Pharm Des. 2016;22:1324–33.

52. Berk M, Kapczinski F, Andreazza AC, Dean OM, Giorlando F, Maes M, et al. Pathways underlying neuroprogression in bipolar disorder: focus on inflammation, oxidative stress and neurotrophic factors. Neurosci Biobehav Rev. 2011;35:804–17.

53. Liberto CM, Albrecht PJ, Herx LM, Yong VW, Levison SW. Pro-regenerative properties of cytokine-activated astrocytes. J Neurochem. 2004;89:1092–1100.

54. Lan MJ, Chhetry BT, Oquendo MA, Sublette ME, Sullivan G, Mann JJ, et al. Cortical thickness differences between bipolar depression and major depressive disorder. Bipolar Disord. 2014;16:378–88.

55. Rimol LM, Nesvag R, Hagler DJ Jr., Bergmann O, Fennema-Notestine C, Hartberg CB, et al. Cortical volume, surface area, and thickness in schizophrenia and bipolar disorder. Biol Psychiatry. 2012;71:552–60.