ORIGINAL ARTICLE

Distinctive Neuroanatomical Substrates for Depression in Bipolar Disorder versus Major Depressive Disorder

Koji Matsuo¹, Kenichiro Harada¹, Yusuke Fujita², Yasumasa Okamoto³, Miho Ota⁴, Hisashi Narita⁵, Benson Mwangi⁶, Carlos A. Gutierrez⁶, Go Okada³, Masahiro Takamura³, Hirotaka Yamagata¹, Ichiro Kusumi⁵, Hiroshi Kunugi⁴, Takeshi Inoue^{5,7}, Jair C. Soares⁶, Shigeto Yamawaki³ and Yoshifumi Watanabe¹

¹Division of Neuropsychiatry, Department of Neuroscience, Yamaguchi University Graduate School of Medicine, 1-1-1 Minamikogushi, Ube, Yamaguchi 755-8505, Japan, ²Division of Electrical, Electronic and Information Engineering, Graduate School of Sciences and Technology for Innovation, Yamaguchi University, Tokiwadai 2-16-1, Ube, Yamaguchi 755-8611, Japan, ³Department of Psychiatry and Neurosciences, Institute of Biomedical and Health Sciences, Hiroshima University, 1-2-3 Kasumi, Minami-ku, Hiroshima 734-8553, Japan, ⁴Department of Mental Disorder Research, National Institute of Neuroscience, National Center of Neurology and Psychiatry, 4-1-1 Ogawa-Higashi, Kodaira, Tokyo 187-8502, Japan, ⁵Department of Psychiatry, Hokkaido University Graduate School of Medicine, North 15, West 7, Kita-ku, Sapporo 060-8638, Japan, ⁶Department of Psychiatry, The University of Texas Health Science Center at Houston, TX 77030, USA and ⁷Department of Psychiatry, Tokyo Medical University, 6-7-1, Nishishinjuku, Shinjuku-ku, Tokyo 160-0023, Japan

Address correspondence to Koji Matsuo, Division of Neuropsychiatry, Department of Neuroscience, Yamaguchi University Graduate School of Medicine, 1-1-1 Minamikogushi, Ube, Yamaguchi 755-8505, Japan. Email: kmatsuo@yamaguchi-u.ac.jp

Abstract

No neuroanatomical substrates for distinguishing between depression of bipolar disorder (dBD) and major depressive disorder (dMDD) are currently known. The aim of the current multicenter study was to identify neuroanatomical patterns distinct to depressed patients with the two disorders. Further analysis was conducted on an independent sample to enable generalization of results. We directly compared MR images of these subjects using voxel-based morphometry (VBM) and a support vector machine (SVM) algorithm using 1531 participants. The VBM analysis showed significantly reduced gray matter volumes in the bilateral dorsolateral prefrontal (DLPFC) and anterior cingulate cortices (ACC) in patients with dBD compared with those with dMDD. Patients with the two disorders shared small gray matter volumes for the right ACC and left inferior frontal gyrus when compared with healthy subjects. Voxel signals in these regions during SVM analysis contributed to an accurate classification of the two diagnoses. The VBM and SVM results in the second cohort also supported these results. The current findings provide new evidence that gray matter volumes in the DLPFC and ACC are core regions in displaying shared and distinct neuroanatomical substrates and can shed light on elucidation of neural mechanism for depression within the bipolar/major depressive disorder continuum.

Key words: cingulate, frontal gyrus, support vector machine, unipolar depression, VBM

Introduction

Depression is a common mental disorder, characterized by depressive mood; loss of interest or pleasure; feelings of guilt or low self-worth; reduction of concentration, and alterations in sleep and appetite, leading to suicide-related behaviors in a severe state. Depression is often long-lasting or recurrent, substantially impairing an individual's ability to function at work or school, and significantly impacting society (World Health Organization; <http://www.who.int/topics/depression/en/>). However, it is not commonly straightforward to accurately diagnose depression including depressed state of major depressive disorder (dMDD) and bipolar disorder (dBD). In a survey of 600 patients with bipolar disorder (BD), 69% reported being misdiagnosed, with 60% instead receiving a major depressive disorder (MDD) diagnosis (Hirschfeld et al. 2003). In another study of 145 patients originally diagnosed with BD by primary care physicians or psychotherapists, 56% were later diagnosed with MDD (Zimmerman et al. 2010). Differential diagnosis of patients with acute depressive episodes of BD (dBD) versus MDD (dMDD) is an important clinical issue because their symptomatic criteria are the same, yet the pharmacotherapeutic strategies are quite different. Mood stabilizers such as lithium, antiepileptics, and atypical antipsychotics are the first choice for dBD (Yatham et al. 2013), while antidepressants are the first choice for dMDD (Bauer et al. 2013). However, if patients with dBD are administered antidepressants without any mood stabilizers because they have been incorrectly diagnosed with dMDD, this could potentially trigger a switch from the depressed phase to the manic phase (Yatham et al. 2013). The patients with dBD who were administered antidepressants without mood stabilizers would be at a risk of rapid cycling (Schnack et al. 2008) that is assumed to be treatment-resistant, resulting in an iatrogenic negative outcome (Schnack et al. 2008). Although clinicians and researchers have sought to identify biomarkers for differentiating between the two disorders, none are currently known.

Neuroimaging is a powerful tool for the potential identification of neuroanatomical biomarkers specific to dBD or dMDD, with substantial meta-analytic studies indirectly comparing patients with MDD or BD to healthy subjects (Koolschijn et al. 2009; Bora et al. 2010, 2012; Ellison-Wright and Bullmore 2010; Kempton et al. 2011; Selvaraj et al. 2012; Lai 2013). While a few structural neuroimaging studies have used whole-brain analysis techniques to directly compare the structural brain volume between patients with BD and those with MDD (Supplementary Table 1), the sample sizes were small and the results have thus far been inconclusive.

The Support Vector Machine (SVM) is a classifier in pattern recognition and has been utilized in various research fields and industries due to its characteristic flexibility and classification ability. SVM analyses have also been used in neuroimaging studies for mood disorders in order to classify patients and healthy subjects (Redlich et al. 2014; Rocha-Rego et al. 2014; Schnack et al. 2014; Mwangi et al. 2016) as well as a tool in making predictions of treatment response (Redlich et al. 2016). However, to our knowledge, no study has yet been done to classify neuroanatomical substrates of patients with dBD and dMDD in a large sample.

The aim of the current study was to identify neuroanatomical patterns distinct to patients with dBD or dMDD using voxel-based morphometry (VBM) and SVM, drawing on a large primary sample recruited from multiple sites and locations across Japan. To enable generalization of results by VBM and SVM, we also conducted further analysis on MR images from an independent

sample in the United States of America. We hypothesized that abnormal gray matter volume (GMV) in the prefrontal and limbic regions involved in emotional processing circuits and the pathophysiology of BD and MDD (Adler et al. 2006; Cardoso de Almeida and Phillips 2013; Savitz et al. 2013) would help discriminate between the 2 types of depression in VBM, classifying them with high accuracy in SVM. The prefrontal-limbic network is involved in the emotional processing and cognitive regulation, including attention and behavioral control (Price and Drevets 2012; Savitz et al. 2014). These regions also play a key role in the regulation of emotional, behavioral, endocrine, and innate immunological responses to stress (Price and Drevets 2010). Other networks have little evidence of involvement in such processing, in the context of mood disorders (Adler et al. 2006; Cardoso de Almeida and Phillips 2013; Savitz et al. 2013). The regions in the prefrontal-limbic network are associated with the clinical characteristics that may be used to discriminate between dBD and dMDD. For instance, the age at onset of illness, family history of BD, and multiple episodes of depression are discriminative characteristics (Bowden 2005; Mitchell et al. 2008) associated with reduced GMV in medial temporal, subcortical, and prefrontal regions (Matsuo et al. 2012a; Matsuo et al. 2012b; Matsubara et al. 2016). Manic symptoms manifest in BD but not in MDD, and one study has demonstrated an association between the number of manic episodes and reduced GMV of inferior prefrontal gyrus (Ekman et al. 2010). Although meta-analysis studies of VBM have revealed that patients with BD and MDD, compared with healthy subjects, show smaller GMV of similar regions of the prefrontal-limbic network, including the inferior prefrontal gyrus (Koolschijn et al. 2009; Bora et al. 2010, 2012; Ellison-Wright and Bullmore 2010; Selvaraj et al. 2012), we hypothesized that different structural abnormalities would be observed in similar depressed states of BD and MDD. We further hypothesized that these results would be confirmed/replicated in the independent, second cohort from the United States of America.

Materials and Methods

Participants

Participants in Japan included 596 patients with dMDD, 158 with dBD and 777 healthy participants, recruited from 3 university hospitals, 2 psychiatric hospitals, and 1 neuroscience institute with 4 physical sites (Table 1). Details of participant information are summarized in Supplementary Methods and Supplementary Table 2. In the Japanese sample, there were significant differences between groups in sex distribution ($\chi^2 = 31.042$, $P = 1.817 \times 10^{-7}$; effect size [ES] [Cramer's V] = 0.142) and intracranial volume ($F(2, 1528) = 39.503$, $P = 1.874 \times 10^{-17}$; ES [η_p^2] = 0.049) but not age ($F(2, 1528) = 0.367$, $P = 0.693$; $\eta_p^2 = 0.0005$). The effect sizes were relatively small. The US sample included 43 patients with dMDD, 36 with dBD, and 132 healthy participants from one university hospital in San Antonio, TX (Table 1). There was no significant difference between groups in sex distribution ($\chi^2 = 0.858$, $P = 0.651$; Cramer's $V = 0.064$), intracranial volume ($F(2, 208) = 1.904$, $P = 0.152$; $\eta_p^2 = 0.018$), or age ($F(2, 208) = 0.208$, $P = 0.217$; $\eta_p^2 = 0.015$). All participants in the present study were right handed. Mean (SD) score on the HDRS was 17.6 (6.00) ($n = 370$); mean duration of illness was 9.18 (8.71) years ($n = 375$); and the means of depressed, manic, and total number of episodes were 2.52 (3.60) ($n = 288$), 3.00 (8.81) ($n = 48$), and 3.10 (6.20) ($n = 288$), respectively, in the Japanese sample. These measures were only collected at some centers, while the severity of depression was

Table 1 Demographics and clinical characteristics of participants

Diagnosis	Japan			USA		
	dBD	dMDD	Healthy	dBD	dMDD	Healthy
No. of subjects	156	596	777	36	43	132
Age	44.0 (14.6)	45.1 (14.0)	45.0 (15.5)	40.3 (11.0)	38.4 (13.8)	36.0 (12.6)
Sex (Female, %)	79 (50.0)	299 (50.2)	499 (64.2) ^a	26 (72.2)	27 (62.8)	86 (65.2)
ICV	1333.48 (194.04)	1340.72 (202.34)	1234.63 (255.82) ^b	1313.08 (135.61)	1351.19 (134.90)	1356.62 (130.98)
HDRS	17.2 (6.1)	18.1 (6.0)		18.03 (5.91)	17.05 (4.12)	
No. of episodes						
Depressed	5.28 (6.92)	3.14 (4.13)		1.64 (4.08)	11.41 (31.65)	
Manic	3.06 (6.31)			8.00 (7.25)		
Total	7.11 (12.56)	2.16 (2.63)		5.42 (9.18)	11.41 (31.65)	
Duration of illness (y)	13.44 (10.19)	7.37 (7.47)				
Medicated (%)	96.8	90.9		55.6	35.7	
SSRI	27	231		10	6	
SNRI	9	103				
TCA	21	79				
Other ADs	18	110				
APs	77	131		3	6	
Li	54	47		2	0	
Other MSs	63	40		15	13	
Panic disorder	6	16		10	12	
GAD	2	30		9	9	
OCD	0	9		3	4	
Specific phobia	0	18		6	3	
Other comorbid anxieties	2	10		22	25	

Numbers represent the mean (SD) for each category. The number indicated for medications and comorbid illnesses is the number of subjects. dBD, depressed patients with bipolar disorder; dMDD, depressed patients with major depressive disorder; ICV, intracranial volume; HDRS, Hamilton Depression Rating Scale. SSRI, Selective Serotonin Reuptake Inhibitor; SNRI, Serotonin Noradrenaline Reuptake Inhibitor; TCA, tricyclic antidepressants; ADs, antidepressants; APs, antipsychotics; Li, lithium carbonate; MSs, mood stabilizers; GAD, generalized anxiety disorder; OCD, obsessive-compulsive disorder. In the US cohort, the 10 patients with dBD and 6 with dMDD that took antidepressants did so with one or more of those indicated. ^a $P = 1.817 \times 10^{-7}$ by Pearson's χ^2 and ^b $P = 1.874 \times 10^{-17}$ by analysis of variance for group difference in Japanese sample.

assessed clinically at others. This study was approved by the institutional review board at each site. Written informed consent was obtained from each participant after a complete description of the study was provided.

Magnetic Resonance Imaging

Brain images were collected using 8 MRI scanners. Details of MRI acquisition are summarized in Supplementary Methods. MR images were manually checked for quality and evaluated for abnormal findings at each site, then collectively rechecked at Yamaguchi University before analysis.

Image Analysis

All MR images were pooled at Yamaguchi University. Image preprocessing was performed using the VBM8 toolbox (<http://dbm.neuro.uni-jena.de/vbm/>) for SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK), running under MATLAB R2015a (MathWorks, Natick, MA, USA). All original images were manually aligned along the anterior-posterior commissure line. T1-weighted images were segmented to remove nonbrain tissues—including the scalp, skull, and dural venous sinuses—and imported into a format that could be used by the VBM8 algorithm. Segmented images were normalized in the Montreal Neurological Institute (MNI) space using the template assembled in VBM8 and smoothed with an 8-mm Gaussian filter. Brain regions were identified using automated anatomical labeling by WFU PickAtlas version 3.04 (<http://fmri.wfubmc.edu/software/PickAtlas>). All images show MNI coordinates.

Statistical Analysis of VBM

We analyzed gray matter images using a general linear model. We set the regions of interest (ROI) in the fronto-limbic structures involved in the pathophysiology of BD and MDD (Adler et al. 2006; Cardoso de Almeida and Phillips 2013; Savitz et al. 2013) to include 34 emotional processing regions within the frontal, subcortical, and medial temporal regions defined by the WFU PickAtlas version 3.04 (<http://fmri.wfubmc.edu/software/PickAtlas>). First, we evaluated the scanner effect using healthy subjects as in previous studies (Meda et al. 2008; Sarrazin et al. 2014) in order to eliminate any masking effects which might affect diagnosis. We applied an analysis of variance (ANOVA) model in SPM8 with gray matter images as a dependent variable; a scanner with 8 levels as an independent variable; and age, sex, and intracranial volume (ICV) as covariates. The result was treated as an image of exclusive mask to eliminate the scanner effect from subsequent statistical analyses of VBM. Then, we applied an ANOVA model with 3 diagnosis levels (dMDD, dBD, and healthy subjects) with age, sex, and ICV as covariates and scanner effects image treated as an exclusive mask. Voxel-wise F-tests and T-tests on diagnosis were performed with a threshold of $P < 0.05$ corrected for familywise error rate ($P_{FWE-corr}$) in SPM8. We calculated GMVs of regions with significance in this analysis using the *get_totals* script (<http://www0.cs.ucl.ac.uk/staff/G.Ridgway/>). We did not apply a statistical model of analysis of variance with scanner as a covariate for VBM analyses because a preliminary analysis showed that this model had similar results to the one without scanner as a covariate in our sample. Thus, we concluded that

the statistical model was not suitable to data analysis of MR imaging in this study in order to control for the effect of scanner.

Regarding regional GMVs with significance between patients with dBD and dMDD, we performed a partial correlation analysis between these GMVs and clinical variables, controlling for age, sex, and ICV in SPSS Statistics version 20 for Windows (IBM, Chicago, IL, USA); in which variables were the age of onset illness, the duration of illness, the number of depressed, manic, and total episodes, and the medication load. We referred to the number of antidepressants, antipsychotics, and mood stabilizers administered as the “medication load” because we did not obtain the dosage of psychotropic drugs administered to patients at the time of participation in the study in some sites. The results shown were Bonferroni-corrected.

SVM

SVM used 133,133 voxel signals of MR for each subject in the ROIs obtained from VBM analysis. SVM classified diagnoses using the voxel signals, age, sex, and scanner information. We performed SVM with the package “e1071 1.6.7” under R version 3.3.1 (Platform: x 86_64-redhat-linux-gnu, Running under: CentOS release 6.8).

Hard-margin SVM can classify linearly separable sets of data by two classes with separating hyperplanes. The formula is as follows:

$$f(\mathbf{x}) = \text{sgn}(\mathbf{w}^T \mathbf{x} + b)$$

where $\mathbf{x}$ represents features, $\mathbf{w}$ weight vector, and b bias. SVM creates a decision boundary with an optimal hyperplane to divide two classes in feature space. Weight vector and bias were decided to maximize the margin, which is the distance between the separating hyperplane and the closest training data. Classification function $f(\mathbf{x})$ outputs would be given “1” when the sum of dot product of features and weight vector on bias is more than zero and the outputs would be given “-1” when the sum is less than zero period-space. Weight vector and bias are decided during training by solving an optimization problem with the following constraint:

$$y_i(\mathbf{w}^T \mathbf{x}_i + b) \geq 1$$

$$\underset{\mathbf{w}, b}{\text{argmax}} \frac{1}{2} \|\mathbf{w}\|^2$$

where $\mathbf{x}_i$ belongs to one of two classes, with the corresponding labels $y_i = \pm 1$.

This study used soft-margin SVM to arrive at an optimal hyperplane. Soft-margin SVM adds an extra cost term with the slack variable ξ_i to the cost function of hard-margin SVM and calculates $\mathbf{w}$ and b to be minimized by the following formula:

$$\underset{\mathbf{w}, b}{\text{argmin}} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^N \xi_i$$

where $\mathbf{w}$ represents weight vector, b does bias, the second term represents the amount of distance from the sample to be misclassified by the separating hyperplane, and C represents the parameter for adjusting a trade-off between the maximum of the margin in the first term and mis-classification errors in the second term. The larger is C —the faster the recognition rate of the trained data—but larger C s may lead to overfitting and poor generalizability. With generalization in mind, optimal parameters

were decided by 10-fold cross-validation with 10^{-2} , 10^{-1} , 1, 10, and 10^2 as candidates for C in training data.

Feature Extraction

We performed feature reduction with two levels because the voxel data had too many features, some of which were potentially redundant. One level was domain-knowledge based—focused on the 34 specific neuroanatomical regions as in the VBM analysis—and the other was feature extraction using principal component analysis (PCA) by retaining only 90% of the variance. The variables were scaled to have unit variance. Then, we performed PCA with the function “prcomp” of the package “stats 3.4.1” under R version 3.3.1 (Platform: x 86_64-redhat-linux-gnu, Running under: CentOS release 6.8).

Performance Assessment

To evaluate generalizability, all samples were randomly divided into training and test samples (Supplementary Fig. 1a). We used 50% of the entire data set for training and the remaining 50% for testing.

Training data were used to design SVM, including feature extraction for each region by principal component analysis (PCA), which was then applied to test data. We also used age, sex, and scanner information as input predictors. Furthermore, we designed SVM using only features extracted from voxel data, age, and sex without scanner information to evaluate the effect of scanner information on SVM results. Classification performance of SVM was assessed by classification of test data.

These procedures were repeated 10 times using different data separation for training and test conditions. Performance assessments included accuracy %, sensitivity %, specificity %, and Diagnostic odds ratio (DOR).

We examined average coefficients of SVM weight vector $\mathbf{w}$ in each region on 10 trials. All voxels in each region have coefficients of SVM weight vectors and we calculated the l_2 norm of coefficients in each region as follows:

$$\|\mathbf{x}\|_2 = \sqrt{\sum_k x_k^2}$$

where $\mathbf{x}_k$ represents a coefficient of k th voxel in a region. Classification accuracies were tested for significance using a chi-square test as a permutation test (Fu et al. 2008; Redlich et al. 2014).

Generalization for the Independent Cohort

VBM

We set an ANOVA model with the 3 diagnosis levels (dMDD, dBD, and healthy subjects), and age, sex, and ICV as covariates in SPM8 in the US sample. Then, we analyzed the difference of diagnoses among dBD, dMDD, and healthy subjects in the fronto-limbic structures with and without the mask image that showed a significant contrast for diagnosis differences from the analysis of the Japanese sample.

SVM

We assessed the classification performance on US data using SVM which was trained on the Japanese samples to evaluate generalization ability. We used the Japanese analysis without the scanner information as US data were collected using one MRI scanner (Supplementary Fig. 1b).

Results

VBM

Effect of a Scanner

A significant effect of a scanner was demonstrated in a large area of the ROI with 2 clusters: the right amygdala/left thalamus (Coordinates of the voxel of maximal statistical significance, $x = 31.5$, $y = 4.5$, $z = -19.5$; $P_{FWE-corr} < 0.001$; $k = 83274$; $F = 258.7$) and the left middle frontal gyrus ($x = -33$, $y = 30$, $z = 21$; $P_{FWE-corr} = 0.001$; $k = 30$; $F = 6.8$) (Supplementary Fig. 2). The area with the effect of the scanner covered around 62.6% of all the ROI area. We treated the result as an exclusive mask for subsequent statistical analyses of VBM.

dBD Versus dMDD

Patients with dBD showed significantly reduced GMVs compared with those with dMDD in 4 clusters: bilateral dorsolateral prefrontal cortex (DLPFC) in middle frontal gyrus ($x = 30$, $y = 33$, $z = 48$; $k = 886$, $T = 6.60$, $P_{FWE-corr} = 4.80 \times 10^{-7}$ for the right side; $x = -25.5$, $y = 49.5$, $z = 31.5$; $k = 75$, $T = 5.46$, $P_{FWE-corr} = 2.78 \times 10^{-4}$ for the left side), right ACC ($x = 3$, $y = 49.5$, $z = 21$; $k = 55$, $T = 5.58$, $P_{FWE-corr} = 1.55 \times 10^{-4}$), and left superior medial frontal gyrus adjacent to the ACC ($x = -4.5$, $y = 49.5$, $z = 24$; $k = 98$, $T = 5.14$, $P_{FWE-corr} < 0.001$) (Fig. 1a). The difference of diagnosis for each scanner was significant and had a large effect size in all but one scanner (Supplementary Fig. 3a and Supplementary Table 3a). Thus, differences in regional GMVs between patients

with dBD and dMDD appeared to be valid across scanners. There was no region with significantly small GMVs in patients with dMDD compared with in those with dBD.

There was not a significant correlation between the 4 regional GMVs and clinical variables, including the duration of illness; the number of depressed, manic, and total episodes; or the number of antidepressants, antipsychotics, and mood stabilizers.

Patients Versus Healthy Subjects

Patients with dBD showed significantly reduced GMV compared with healthy subjects in the 2 clusters in the right ACC and middle frontal gyrus ($x = 3$, $y = 49.5$, $z = 19.5$; $k = 12318$, $T = 11.5$, $P_{FWE-corr} < 0.001$) and the left inferior frontal gyrus, triangular part ($x = -39$, $y = 10.5$, $z = 25.5$; $k = 44$, $T = 4.9$, $P_{FWE-corr} = 0.003$) (Fig. 1b). The difference of diagnosis for each scanner was significant and had a large effect size (Supplementary Fig. 3b and Supplementary Table 3b), thus the regional difference in GMVs by diagnosis appears valid across scanners.

Patients with dMDD showed significantly reduced GMVs compared with healthy subjects in three clusters, including the right middle frontal gyrus and left medial frontal gyrus ($x = 34.5$, $y = 49$, $z = 28.5$; $k = 12841$, $T = 9.9$, $P_{FWE-corr} < 0.001$); left inferior frontal gyrus, orbital and triangular part, and left insula ($x = -37.5$, $y = 28.5$, $z = 33$; $k = 348$, $T = 9.7$, $P_{FWE-corr} < 0.001$); and left middle frontal gyrus ($x = -24$, $y = -3$, $z = 48$; $k = 156$, $T = 4.8$, $P_{FWE-corr} = 0.005$) (Fig. 1c). The difference of diagnosis for each scanner in patients with dMDD and healthy subjects was

Figure 1. Difference in gray matter across diagnoses in voxel-based morphometry. Brain mapping of the differences in gray matter volume across diagnoses; depressed patients with bipolar disorder (dBD) versus major depressive disorder (dMDD) in voxel-based morphometry (VBM) (a), patients with dBD versus healthy subjects (b), and patients with dMDD versus healthy subjects (c). Red areas indicate reduced gray matter volume between diagnoses.

significant and had a large effect size (Supplementary Fig. 3c and Supplementary Table 3c). This observation suggests that the regional difference in GMVs by diagnosis appears valid across scanners.

There was no region with significantly small GMVs in healthy subjects compared with in patients with dBD or dMDD.

Post Hoc VBM Analysis

To ensure that the effect of scanner did not affect the results of VBM in the primary analysis, we additionally conducted a full-factorial model incorporating two factors with two levels of diagnosis (patients and healthy subjects) and a scanner with 8 levels, with age, sex, and ICV as covariates under SPM8. We tested the main effect of diagnosis, the main effect of scanners, and the interaction between diagnosis and scanners. In the results, the subcortical area with a significant main effect of scanners overlapped with that in the primary analysis (Supplementary Table 4). GMV of subcortical and medial temporal regions showed significant interactions between diagnosis and scanner, and these areas did not overlap with those having significant differences between dBD and dMDD in the primary analysis.

We also evaluated the effect size of the regional GMVs obtained from the primary VBM analysis under SPM8 using the design matrix with a full-factorial model in the post hoc VBM analysis. In the results, the effect size of the regions with significant differences across dBD, dMDD, and healthy subjects was larger in the diagnosis than in the scanner and in the interaction between diagnosis and scanner (Supplementary Fig. 4). The two post hoc VBM analyses suggested that the main findings in the primary VBM analysis had relatively little variance across scanners.

SVM

Cross-validation

dBD versus dMDD. SVM classified the two diagnoses with 63.4% accuracy, 69.7% sensitivity, 46.4% specificity, and 2.00 DOR. The bilateral middle and superior frontal gyrus and left inferior frontal gyrus, triangular part—which were also significant regions in the VBM results—contributed greatly to this classification (Fig. 2a). The probability of this classification was significant (χ^2 test: $\chi^2 = 64.7$, $df = 1$, $p = 8.74 \times 10^{-16}$). The contribution of features in regional SVM weights showed similar trends with and without the scanner information (Fig. 2a vs. Supplementary Fig. 5a), and the scanner features were placed at a low contribution (Fig. 2a), indicating that the 4 features of the scanner did not contribute to the classification in any substantial degree.

Patients Versus Healthy Subjects. The classification between patients with dBD and healthy subjects was 88.1% accuracy, 92.1% sensitivity, 73.4% specificity, and 32.20 DOR ($\chi^2 = 1542.31$, $df = 1$, $P < 0.01 \times 10^{-50}$) (Fig. 2b). The classification between patients with dMDD and healthy subjects was 75.9% accuracy, 78.1% sensitivity, 72.9% specificity, and 9.60 DOR ($\chi^2 = 1765.68$, $df = 1$, $P < 0.01 \times 10^{-50}$) (Fig. 2c). The contribution of features in regional SVM weights showed similar trends with and without the scanner information (Fig. 2b vs. Supplementary Fig. 5b for patients with dBD vs. healthy subjects; Fig. 2c vs. Supplementary Fig. 5c for patients with dMDD vs. healthy subjects) and the scanner features were placed at a low contribution (Fig. 2b and c), indicating that the feature of scanner had little influence on the SVM analysis.

Generalization for the Independent Cohort

VBM

VBM analysis that masked the areas with significance while analyzing the Japanese sample revealed that patients with dBD showed significantly decreased GMV in right DLPFC compared with those with dMDD in the US sample ($x = 25.5$, $y = 51$, $z = 34.5$, $k = 64$, $t = 3.36$, $P_{FWE-corr} = 0.03$) (Fig. 3). Patients with dBD showed reduced GMV compared with healthy subjects in one cluster ($x = 12$, $y = 55.5$, $z = -1.5$, $k = 387$, $t = 3.54$, $P_{uncorrected} = 0.0002$) that included the left medial orbital and superior medial frontal gyrus, and it did not reach significance after a multiple-comparison correction. There were no significant differences in any regional GMVs between patients with dMDD and healthy subjects during VBM analysis.

SVM

SVM classification for the two diagnoses in the US sample showed lower classification power than that in the Japanese sample: for dBD versus dMDD, accuracy 53.2%, specificity 58.1%, sensitivity 47.2%, and DOR 1.24 ($\chi^2 = 0.23$, $df = 1$, $P = 0.63$); for dBD versus healthy subjects, accuracy 58.3%, specificity 60.6%, sensitivity 50.0%, and DOR 1.54 ($\chi^2 = 1.31$, $df = 1$, $P = 0.25$); for dMDD versus healthy subjects, accuracy 54.3%, specificity 56.1%, sensitivity 48.8%, and DOR 1.22 ($\chi^2 = 0.31$, $df = 1$, $P = 0.58$).

Post Hoc Cross-validation of SVM in the Second Cohort

To determine the reason for lower classification power in the US sample, we also performed cross-validation for the US sample, as was done in the Japanese sample, to classify the diagnosis using SVM (Supplementary Fig. 1a). SVM classification for the two diagnoses was as follows: for dBD versus dMDD, accuracy 49.5%, specificity 57.5%, sensitivity 39.7%, and DOR 0.89 ($\chi^2 = 0.33$, $df = 1$, $P = 0.56$); for dBD versus healthy subjects, accuracy 72.1%, specificity 85.5%, sensitivity 17.6%, and DOR 1.26 ($\chi^2 = 0.97$, $df = 1$, $P = 0.33$); for dMDD versus healthy subjects, accuracy 63.6%, specificity 81.3%, sensitivity 12.8%, and DOR 0.64 ($\chi^2 = 4.14$, $df = 1$, $P = 0.04$). Accuracy scores in the cross-validation were lower in the US sample than in the Japanese sample.

We also tested a relationship in regional SVM weights between the US and Japanese sample. Pearson correlation coefficients were significant (Fig. 4): $r = 0.72$, $P < 0.001$ for dBD versus dMDD; $r = 0.68$, $P < 0.001$ for dBD versus healthy subjects; and $r = 0.73$, $P < 0.001$ for dMDD versus healthy subjects. The results suggest that the US and Japanese samples showed similar trends in regards to the rank of regional SVM weights and had the greatest SVM weights in the bilateral superior and middle frontal gyrus across the diagnoses (Supplementary Table 5).

Discussion

The primary findings in the current study are as follows: (1) GMVs of bilateral DLPFC and ACC involving medial frontal gyrus are different between patients with dBD and dMDD, (2) compared with healthy subjects, patients with the two disorders share small GMVs in the right ACC and left inferior frontal gyrus, (3) more severe abnormalities in certain regions were found in patients with dBD in VBM, and (4) these regions contribute to classification of the two diagnoses in SVM models. In part, the results from the independent US cohort also lent support to these conclusions. The findings indicate that GMVs in DLPFC and ACC are a core feature for shared and distinct brain

patients with BD have decreased GMVs in the ACC (Bora et al. 2010; Ellison-Wright and Bullmore 2010), insula (Bora et al. 2010; Ellison-Wright and Bullmore 2010; Selvaraj et al. 2012), and temporal pole (Selvaraj et al. 2012), while patients with MDD have decreased GMVs in the ACC (Koolschijn et al. 2009; Bora et al. 2012; Lai 2013) and right DLPFC (Bora et al. 2012). Prior studies suggest that severe outcomes of BD, including suicide attempts, intermediate onset, and multiple manic episodes, are associated with reduced GMV in the DLPFC (Penttila et al. 2009; Ekman et al. 2010; Benedetti et al. 2011; Bora et al. 2012; Abe et al. 2015). In a previous meta-analysis study, patients with MDD were shown to have reduced activation in

Figure 3. Results of voxel-based morphometry analysis in the independent second cohort. Reduced gray matter volume (GMV) in the right dorsolateral prefrontal cortex (red area) in depressed patients with bipolar disorder (dBD) compared with those with major depressive disorder (dMDD) (a). Plots of GMV for the right dorsolateral prefrontal cortex in patients with dBD and dMDD (b). The bars indicate mean GMV for each diagnostic group.

the DLPFC, ACC, medial and inferior prefrontal cortex, insula, and other fronto-temporal cortices (Fitzgerald et al. 2008). Also, regarding the direct comparison of BD and MDD in ACC volume, one study showed greater ACC GMVs in patients with BD compared to those with MDD (Redlich et al. 2014) while others did not demonstrate any differences between the two diagnoses (Almeida et al. 2013; Cai et al. 2015; Wise et al. 2017). Some neuropsychological studies suggest that patients with BD demonstrate worse performance in executive, attentional, and memory tasks, compared with patients with MDD (Maalouf et al. 2010; Cotrena et al. 2016; Szmulewicz et al. 2017). These cognitive functions are a readout of the function of DLPFC and ACC (Purves 2008; Vogt 2009), although a meta-analysis study showed that the patients with dBD and those with dMDD had similar behavioral performance, based on cognitive measures (Samame et al. 2017). The difference in GMVs of DLPFC and ACC, between patients with dBD and those with dMDD, is likely to contribute to the difference in cognitive deficits observed. Future studies to evaluate neurocognitive function in the patients with dBD and those with dMDD, combined with GMVs of the relevant brain regions, would be necessary to distinguish the functional and structural brain pathophysiology of dBD from that of dMDD. For the biological treatment of mood disorder, resting-state networks—including the DLPFC—predict ECT treatment outcomes for treatment-resistant patients with MDD (van Waarde et al. 2015). The DLPFC is also a major stimulation site in rTMS, and blood flow in the frontal regions at baseline may serve as a predictor of the patient's response to rTMS in those with depression (for review, Silverstein et al. 2015). A review of magnetic resonance spectroscopy demonstrated an abnormal neurochemical function of the DLPFC—as well as ACC—in patients with BD (Yildiz-Yesiloglu and Ankerst 2006). Furthermore, neuropathological studies have demonstrated that brain tissue of the DLPFC in patients with BD and MDD has reduced neuronal and glial density (Rajkowska et al. 1999, 2001). A very recent meta-analysis study of VBM in MDD and BD demonstrated reduced GMV in the right DLPFC in patients with MDD compared with those with BD (Wise et al. 2017). The investigators suggested that differential patterns of prefrontal GMV may potentially differentiate between these two disorders. Our contrasting findings may be due to methodological differences between the two studies. For instance, the former was a meta-analysis using publication studies of BD and MDD

that conducted indirect comparisons between the two disorders for effect size and various mood states (euthymic, depressed, and manic), while ours was a direct comparison of MR image data using VBM and SVM analysis and one mood state (depressed). Regarding direct comparison of MDD and BD, previous studies (de Azevedo-Marques Perico et al. 2011; Lan et al. 2014; Redlich et al. 2014; Cai et al. 2015) are in partial agreement with our results, with inconsistencies likely to be due to methodological differences (Supplementary Table 1). Of note, our results cannot necessarily provide an answer as to why different structural abnormalities would underlie similar phenotypic expression of depression in BD and MDD. In the future, a longitudinal study of GMV differences across various mood states (depressed, manic, or euthymic) in patients with BD and MDD might address the problem.

Generalization of the Results in the First Cohort

The VBM and SVM results in the independent second cohort were in partial agreement with those in the first cohort, although classification for the two disorders demonstrated lower accuracy in the second cohort than in the first. These weaker VBM and SVM accuracies may partly explain the minimal difference in GMVs for the fronto-limbic structures across diagnoses in the second cohort. VBM analysis did not show any significant differences in GMV for any region across diagnoses. SVM analysis in the cross-validation for the second cohort showed classification with much lower accuracies of diagnosis compared with the cross-validation analysis of the first cohort. These findings suggest that the MRI data in the second cohort itself presented a difficulty of classification and differentiation across diagnoses.

However, the rank of regional SVM weights showed very similar patterns for the first and second cohorts (Fig. 4), providing the evidence that the bilateral DLPFC in the middle frontal gyrus presents a strong and reliable contribution to classification.

Limitations

We should note some limitations in this study. First, we discuss the multiple-scanner effect in the VBM results. The difference in outputs during VBM analysis is assumed to reflect the differences in tissue classification caused by the imaging

Figure 4. Scatter plot of comparison between support vector machine weight averages in the first and second cohorts. dBD, depressed patients with bipolar disorder; dMDD, depressed patients with major depressive disorder.

characteristics of varying acquisition protocols (Tardif et al. 2010). The effect of a scanner in the current study shared around 60% of the ROIs with the two clusters of the right and left thalamus. These regions were consistent with those on the effect of a scanner in prior studies (Pardoe et al. 2008; Stonnington et al. 2008; Marchewka et al. 2014). We created a design for group-comparison analysis with a scanner as a covariate and evaluated the effect of a scanner among healthy subjects as in previous studies (Meda et al. 2008; Sarrazin et al. 2014). Then, we used the clusters with this effect as the masking template for comparison across diagnostic groups to eliminate the effect of the scanner from the results. We also compared the regional GMVs, for moderate and large effect sizes, with a significant difference in diagnosis for each scanner, demonstrating a similar GMV distribution across the diagnoses for each scanner (Supplementary Fig. 3 and Supplementary Table 3). The results of post hoc VBM analysis (Supplementary Table 4) also demonstrated that the GMV of

subcortical and medial temporal regions with a significant interaction between the diagnosis and scanner in the second analysis did not overlap with those regions of GMV that had a significant difference between dBD and dMDD in the primary analysis. Furthermore, the effect size of the diagnosis was larger than that of the scanner or the interaction between diagnosis and scanner, regarding the regions with a significant difference across diagnoses in the primary analysis (Supplementary Fig. 4). These results in the post hoc VBM analysis suggest that the findings in the primary VBM analysis had comparatively little variance across scanners. SVM analysis also supported the VBM results, which demonstrated very similar results for the model with- and without-scanner feature in these contrasts for diagnosis-comparison (dBD vs. dMDD, Fig. 2a and Supplementary Fig. 5a; dBD vs. healthy, Fig. 2b and Supplementary Fig. 5b; dMDD vs. healthy, Fig. 2c and Supplementary Fig. 5c). These findings are in at least indirect agreement with the supposition that the regional GMVs with

significance exert more influence on variance across diagnoses than across scanners; however, we cautiously note that we cannot exclude the possibility that effect of the scanner might be masking the GMV difference across diagnoses. Additionally, despite the lengthy period of MR image acquisition, we did not track software upgrade information for the MR scanners, and this may have had some effect on the results. Differences in partial volume effects, image intensity nonhomogeneity, electronic noise from the MRI apparatus, and signal noise may have had some effect on results (Stonnington et al. 2008). Assessment of scanner effects using a phantom, for instance, would be preferable in the current context, and more studies are likely necessary to evaluate intercenter acquisition quality in the future.

Second, we did not find a difference in GMV in the hippocampi and amygdalae across diagnoses, although these areas play crucial roles in emotional regulation and memory (Phillips et al. 2008a). Although GMVs in the areas showed a significant main effect of diagnosis, the confounding effects of multiple scanners obscured this result. Multiple previous meta-analytic neuroimaging studies of BD and MDD (Hamilton et al. 2008; Kempton et al. 2008; Hajek et al. 2009; Bora et al. 2010, 2012; Ellison-Wright and Bullmore 2010; Selvaraj et al. 2012; Lai 2013) observed such negative results of the medial temporal regions, while a few other meta-analyses of MDD have yielded positive results (Campbell et al. 2004; Videbech and Ravnkilde 2004; Koolschijn et al. 2009). Further, multisite MRI studies of patients with BD and MDD without any effect of scanner variance will be required to address this issue.

Third, because we sought to address the morphometric differences between the brains of patients with dBD and those with dMDD, using as large a sample size as possible, for the Japanese population, we did not control for participant demographics across sites: age or sex distribution, number of subjects with each disorder, diagnosis methodology, diagnosed subtypes of each disorder, or psychiatric family history of healthy subjects. In addition, we did not obtain complete background information for all patients, including duration of illness, medication load, and depression severity when we collected and pooled data for each site. Thus, we did not include depressive scores in the statistical model of VBM analysis, and the severity of depression may exert some influence on the results of VBM. While approximately 90% of patients were taking psychotropic medications at the time of study participation, information regarding the type of medication, dosage, or duration of use was unavailable. Of those patients with dBD, many took mood stabilizers with some taking antidepressants. Of those with dMDD, many took antidepressants with some taking mood stabilizers. These variables may affect brain volume results. Some studies have demonstrated that antidepressant administration increases dorsolateral prefrontal cortex GMV in patients with MDD (Smith et al. 2013) and that lithium affects cortical GM density (Bearden et al. 2007). Another study demonstrates long-term lithium treatment to be associated with increased GMV in the same areas for which decreased GMV is associated with attempted suicide (Benedetti et al. 2011). One review has suggested that medicated patients with BD show no significant effects of psychotropic medications on structural and functional neuroimaging measures (Phillips et al. 2008b). Thus, we must interpret our results with the appropriate caution, owing to possible effects related to the factors not included in the patient background information that was surveyed.

Fourth, we applied 10-fold-cross-validation with 5 candidates for C using training data for optimal parameters in the SVM analysis. We might have found more optimal parameters

by using a grid-search plot method. The 5 typical parameter candidates we used were confirmed to be appropriate in this study after exploratory testing using other parameters (data not shown). We applied PCA to do feature extraction for each region by retaining 90% of the variance, which was not optimized by cross-validation. We first analyzed this feature extraction by retaining 95% of the variance, which resulted in a slightly different classification performance from that in the original analysis by retaining 90% of the variance (data not shown). Thus, although better parameter optimization may potentially improve the classification performance of the SVM, we believe that our present results are still valid.

Fifth, we found that the cross-validation of SVM in the US sample showed very low sensitivities (0.18 for dBD vs. healthy and 0.13 for dMDD vs. healthy). This may be due to a substantial imbalance in sample size between patients and healthy subjects (36 dBD vs. 43 dMDD vs. 132 healthy subjects). Future studies with comparable numbers across three diagnoses in the US sample would be required to have further validation of results.

Finally, the results that the average SVM weights showed indicated a similar distribution between those of Japan and those of the United States of America (Fig. 4). It appeared as if larger weights were simply found in larger areas, but the reality may not be so straightforward. As the SVM analysis used signal intensity for voxels obtained from VBM analysis, larger weights in a region mean a greater average voxel signal and a larger number of voxels. For instance, it could illustrate that an ROI in both the right and left sides, which should contain a very similar number of voxels, were in different places (Supplementary Table 5). Thus, while larger areas tend to show larger weights, signal intensity per voxel should be taken into consideration when interpreting the results in Figure 4.

Conclusions

The results of the current study provide new evidence showing that reduced GMV in DLPFC and ACC structures is a shared pathophysiological characteristic of both dBD and dMDD, and that more severe abnormalities in certain regions can be used to identify distinct neuroanatomical substrates between the two disorders. The findings indicate that GMVs in DLPFC and ACC are a core feature for shared and distinct pathophysiology and can shed light on elucidation of neural mechanism for depression within the bipolar disorder/major depressive disorder continuum.

Supplementary Material

Supplementary data is available at *Cerebral Cortex* online.

Funding

Hokkaido site

This work was supported by a Grant-in-Aid for “Integrated research on neuropsychiatric disorders,” carried out under the Strategic Research Program for Brain Sciences (SRPBS) by the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT), and Japan Agency for Medical Research and development (AMED).

Tokyo site

This work was supported by an Intramural Research Grant (24-11) for Neurological and Psychiatric Disorders of NCNP (M.O. and H.K.)

and by a Grant-in-Aid for “Understanding of molecular and environmental bases for brain health” carried out under the SRPBS by the MEXT (H.K.).

Hiroshima site

This work was supported by Grants-in-Aid for “Integrated research on neuropsychiatric disorders” and “Brain-machine interface development” carried out under the SRPBS by the MEXT and AMED (O.Y.). A part of this work was also carried out under the Brain Mapping by Integrated Neurotechnologies for Disease Studies (Brain/MINDS) by the MEXT and AMED (S.Y.).

Yamaguchi site

This work was supported by a Grant-in-Aid for “Integrated research on neuropsychiatric disorders” carried out under the SRPBS by the MEXT and AMED (H.Y.) and KAKENHI from the Japan Society for the Promotion of Science (grant numbers 24591716 and 15K09832 to K.M.).

US site

This work was supported in part by the National Institute of Mental Health (grant number R01 085667 to J.C.S.), John S. Dunn Foundation, and Pat Rutherford, Jr. Endowed Chair in Psychiatry (J.C.S.).

Notes

Hokkaido site

T.I. has received honoraria from GlaxoSmithKline, Pfizer, Eli Lilly, Mitsubishi Tanabe Pharma, Mochida Pharmaceutical, Otsuka Pharmaceutical, Meiji Seika Pharma, Asahi Kasei Pharma, Shionogi, Dainippon Sumitomo Pharma, Takeda Pharmaceutical, Merck Sharp & Dohme (MSD), Eisai, and Yoshitomi Pharmaceutical. T.I. has also received research/grant support from Otsuka Pharmaceutical, Eli Lilly, Eisai, Mitsubishi Tanabe Pharma, AbbVie, Pfizer, Astellas Pharma, and Meiji Seika Pharma and is a member of the advisory boards of GlaxoSmithKline, Pfizer, Eli Lilly, Mochida Pharmaceutical, and Mitsubishi Tanabe Pharma. I.K. has received honoraria from Astellas Pharma, Chugai Pharmaceutical, Daiichi Sankyo, Dainippon Sumitomo Pharma, Eisai, Eli Lilly, Janssen Pharmaceutica, Kyowa Hakko Kirin, Meiji Seika Pharma, MSD, Nippon Chemiphar, Novartis, Ono Pharmaceutical, Otsuka Pharmaceutical, Pfizer, Tanabe Mitsubishi Pharma, Shionogi, and Yoshitomiyakuhi. I.K. has received research/grant support from AbbVie GK, Asahi Kasei Pharma, Astellas Pharma, Boehringer Ingelheim, Chugai Pharmaceutical, Daiichi Sankyo, Dainippon Sumitomo Pharma, Eisai, Eli Lilly, GlaxoSmithKline, Kyowa Hakko Kirin, Meiji Seika Pharma, MSD, Novartis, Ono Pharmaceutical, Otsuka Pharmaceutical, Pfizer, Takeda Pharmaceutical, Tanabe Mitsubishi Pharma, Shionogi, and Yoshitomiyakuhi and is a member of the advisory boards of Dainippon Sumitomo Pharma and Tanabe Mitsubishi Pharma.

Hiroshima site

S.Y. has received research support from Astellas Pharma, Pfizer, and MSD (all less than 10 000 USD/year) and has received honoraria from GlaxoSmithKline, Astellas Pharma, Daiichi Sankyo, Eli Lilly, Takeda Pharmaceutical, Otsuka Pharmaceutical, Eisai, Dainippon Sumitomo Pharma, Shionogi, Meiji Seika Pharma, Mochida Pharmaceutical, and Pfizer (all less than 5 000 USD/year).

Yamaguchi site

K.M. has received honoraria from GlaxoSmithKline, Mochida Pharmaceutical, and research support from Otsuka Pharmaceutical. Y.W. has received research support from MSD, GlaxoSmithKline, Eli Lilly, Yoshitomiyakuhi, Shionogi, Pfizer, Novartis, Janssen Pharmaceutica, Meiji Seika Pharma, Fujifilm RI Pharma, Takeda Pharmaceutical, Astellas Pharma, Dainippon Sumitomo Pharma, and Otsuka Pharmaceutical and received honoraria from MSD, Meiji Seika, Pfizer, Shionogi, Mochida Pharmaceutical, Eli Lilly, and AbbVie GK.

US site

J.C.S. has participated in research funded by Forest, Merck & Co., Bristol-Myers Squibb, and Glaxo Smith Kline and has been a paid speaker for Pfizer and Abbott Laboratories.

Acknowledgements

Hiroshima Site:

We thank all the members of the Hiroshima Depression Study (HDS) network for recruiting participants.

Yamaguchi Site:

We thank Drs. Toshio Matsubara, Toshio Watanuki, Koji Otsuki, Teruyuki Hobara, Naoko Higuchi, Fumihiro Higuchi, Tomohiko Shibata, Masayuki Nakano, and Mami Nakashima for recruiting participants. We also thank Dr. Ryota Hashimoto (Molecular Research Center for Children’s Mental Development, United Graduate School of Child Development, Osaka University) for assisting with Institutional Review Board procedures for multicenter studies and Dr. Kiyotaka Nemoto (Department of Psychiatry, University of Tsukuba) for technical assistance with the analysis of voxel-based morphometry. Also, we thank Gregory Wroblewski (Division of Neuroanatomy, Department of Neuroscience, Yamaguchi University Graduate School of Medicine) for proofreading the manuscript’s English usage. *Conflict of Interest:* None declared.

References

- Abe C, Ekman CJ, Sellgren C, Petrovic P, Ingvar M, Landen M. 2015. Manic episodes are related to changes in frontal cortex: a longitudinal neuroimaging study of bipolar disorder 1. *Brain*. 138:3440–3448.
- Adler CM, DelBello MP, Strakowski SM. 2006. Brain network dysfunction in bipolar disorder. *CNS Spectr*. 11:312–320.
- Almeida JR, Mourao-Miranda J, Aizenstein HJ, Versace A, Kozel FA, Lu H, Marquand A, LaBarbara EJ, Brammer M, Trivedi M, et al. 2013. Pattern recognition analysis of anterior cingulate cortex blood flow to classify depression polarity. *Br J Psychiatry*. 203:310–311.
- Bauer M, Pfennig A, Severus E, Whybrow PC, Angst J, Moller HJ, World Federation of Societies of Biological Psychiatry. Task Force on Unipolar Depressive D. 2013. World Federation of Societies of Biological Psychiatry (WFSBP) guidelines for biological treatment of unipolar depressive disorders, part 1: update 2013 on the acute and continuation treatment of unipolar depressive disorders. *World J Biol Psychiatry*. 14: 334–385.
- Bearden CE, Thompson PM, Dalwani M, Hayashi KM, Lee AD, Nicoletti M, Trakhtenbroit M, Glahn DC, Brambilla P, Sassi RB,

- et al. 2007. Greater cortical gray matter density in lithium-treated patients with bipolar disorder. *Biol Psychiatry*. 62:7–16.
- Benedetti F, Radaelli D, Poletti S, Locatelli C, Falini A, Colombo C, Smeraldi E. 2011. Opposite effects of suicidality and lithium on gray matter volumes in bipolar depression. *J Affect Disord*. 135:139–147.
- Bora E, Fornito A, Pantelis C, Yucel M. 2012. Gray matter abnormalities in Major Depressive Disorder: a meta-analysis of voxel based morphometry studies. *J Affect Disord*. 138:9–18.
- Bora E, Fornito A, Yucel M, Pantelis C. 2010. Voxelwise meta-analysis of gray matter abnormalities in bipolar disorder. *Biol Psychiatry*. 67:1097–1105.
- Bowden CL. 2005. A different depression: clinical distinctions between bipolar and unipolar depression. *J Affect Disord*. 84: 117–125.
- Cai Y, Liu J, Zhang L, Liao M, Zhang Y, Wang L, Peng H, He Z, Li Z, Li W, et al. 2015. Grey matter volume abnormalities in patients with bipolar I depressive disorder and unipolar depressive disorder: a voxel-based morphometry study. *Neurosci Bull*. 31:4–12.
- Campbell S, Marriott M, Nahmias C, MacQueen GM. 2004. Lower hippocampal volume in patients suffering from depression: a meta-analysis. *Am J Psychiatry*. 161:598–607.
- Cardoso de Almeida JR, Phillips ML. 2013. Distinguishing between unipolar depression and bipolar depression: current and future clinical and neuroimaging perspectives. *Biol Psychiatry*. 73:111–118.
- Cotrena C, Branco LD, Shansis FM, Fonseca RP. 2016. Executive function impairments in depression and bipolar disorder: association with functional impairment and quality of life. *J Affect Disord*. 190:744–753.
- de Azevedo-Marques Perico C, Duran FL, Zanetti MV, Santos LC, Murray RM, Scazufca M, Menezes PR, Busatto GF, Schaufelberger MS. 2011. A population-based morphometric MRI study in patients with first-episode psychotic bipolar disorder: comparison with geographically matched healthy controls and major depressive disorder subjects. *Bipolar Disord*. 13:28–40.
- Ekman CJ, Lind J, Ryden E, Ingvar M, Landen M. 2010. Manic episodes are associated with grey matter volume reduction - a voxel-based morphometry brain analysis. *Acta Psychiatr Scand*. 122:507–515.
- Ellison-Wright I, Bullmore E. 2010. Anatomy of bipolar disorder and schizophrenia: a meta-analysis. *Schizophr Res*. 117:1–12.
- Fitzgerald PB, Laird AR, Maller J, Daskalakis ZJ. 2008. A meta-analytic study of changes in brain activation in depression. *Hum Brain Mapp*. 29:683–695.
- Fu CH, Mourao-Miranda J, Costafreda SG, Khanna A, Marquand AF, Williams SC, Brammer MJ. 2008. Pattern classification of sad facial processing: toward the development of neurobiological markers in depression. *Biol Psychiatry*. 63:656–662.
- Hajek T, Kopecek M, Kozeny J, Gunde E, Alda M, Hoschl C. 2009. Amygdala volumes in mood disorders—meta-analysis of magnetic resonance volumetry studies. *J Affect Disord*. 115: 395–410.
- Hamilton JP, Siemer M, Gotlib IH. 2008. Amygdala volume in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. *Mol Psychiatry*. 13:993–1000.
- Hirschfeld RM, Lewis L, Vornik LA. 2003. Perceptions and impact of bipolar disorder: how far have we really come? Results of the national depressive and manic-depressive association 2000 survey of individuals with bipolar disorder. *J Clin Psychiatry*. 64:161–174.
- Kempton MJ, Geddes JR, Ettinger U, Williams SC, Grasby PM. 2008. Meta-analysis, database, and meta-regression of 98 structural imaging studies in bipolar disorder. *Arch Gen Psychiatry*. 65:1017–1032.
- Kempton MJ, Salvador Z, Munafo MR, Geddes JR, Simmons A, Frangou S, Williams SC. 2011. Structural neuroimaging studies in major depressive disorder. Meta-analysis and comparison with bipolar disorder. *Arch Gen Psychiatry*. 68:675–690.
- Koolschijn PC, van Haren NE, Lensvelt-Mulders GJ, Hulshoff Pol HE, Kahn RS. 2009. Brain volume abnormalities in major depressive disorder: a meta-analysis of magnetic resonance imaging studies. *Hum Brain Mapp*. 30:3719–3735.
- Lai CH. 2013. Gray matter volume in major depressive disorder: a meta-analysis of voxel-based morphometry studies. *Psychiatry Res*. 211:37–46.
- Lan MJ, Chhetry BT, Oquendo MA, Sublette ME, Sullivan G, Mann JJ, Parsey RV. 2014. Cortical thickness differences between bipolar depression and major depressive disorder. *Bipolar Disord*. 16:378–388.
- Maalouf FT, Klein C, Clark L, Sahakian BJ, Labarbara EJ, Versace A, Hassel S, Almeida JR, Phillips ML. 2010. Impaired sustained attention and executive dysfunction: bipolar disorder versus depression-specific markers of affective disorders. *Neuropsychologia*. 48:1862–1868.
- Marchewka A, Kherif F, Krueger G, Grabowska A, Frackowiak R, Draganski B, Alzheimer's Disease Neuroimaging I. 2014. Influence of magnetic field strength and image registration strategy on voxel-based morphometry in a study of Alzheimer's disease. *Hum Brain Mapp*. 35:1865–1874.
- Matsubara T, Matsuo K, Harada K, Nakano M, Nakashima M, Watanuki T, Egashira K, Furukawa M, Matsunaga N, Watanabe Y. 2016. Distinct and shared endophenotypes of neural substrates in bipolar and major depressive disorders. *PLoS One*. 11:e0168493.
- Matsuo K, Kopecek M, Nicoletti MA, Hatch JP, Watanabe Y, Nery FG, Zunta-Soares G, Soares JC. 2012a. New structural brain imaging endophenotype in bipolar disorder. *Mol Psychiatry*. 17:412–420.
- Matsuo K, Sanches M, Brambilla P, Soares JC. 2012b. Structural brain abnormalities in bipolar disorder. In: Strakowski SM, editor. *The bipolar brain. Integrating neuroimaging and genetics*. New York: Oxford University Press, Inc. p. 17–52.
- Meda SA, Giuliani NR, Calhoun VD, Jagannathan K, Schretlen DJ, Pulver A, Cascella N, Keshavan M, Kates W, Buchanan R, et al. 2008. A large scale (N=400) investigation of gray matter differences in schizophrenia using optimized voxel-based morphometry. *Schizophr Res*. 101:95–105.
- Mitchell PB, Goodwin GM, Johnson GF, Hirschfeld RM. 2008. Diagnostic guidelines for bipolar depression: a probabilistic approach. *Bipolar Disord*. 10:144–152.
- Mwangi B, Wu MJ, Cao B, Passos IC, Lavagnino L, Keser Z, Zunta-Soares GB, Hasan KM, Kapczynski F, Soares JC. 2016. Individualized prediction and clinical staging of bipolar disorders using neuroanatomical biomarkers. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 1:186–194.
- Pardoe H, Pell GS, Abbott DF, Berg AT, Jackson GD. 2008. Multi-site voxel-based morphometry: methods and a feasibility demonstration with childhood absence epilepsy. *Neuroimage*. 42:611–616.
- Penttilä J, Cachia A, Martinot JL, Ringuenet D, Wessa M, Houenou J, Galinowski A, Bellivier F, Gallarda T, Duchesnay E, et al. 2009. Cortical folding difference between patients with early-onset and patients with intermediate-onset bipolar disorder. *Bipolar Disord*. 11:361–370.

- Phillips ML, Ladouceur CD, Drevets WC. 2008a. A neural model of voluntary and automatic emotion regulation: implications for understanding the pathophysiology and neurodevelopment of bipolar disorder. *Mol Psychiatry*. 13:829–833–857.
- Phillips ML, Travis MJ, Fagioli A, Kupfer DJ. 2008b. Medication effects in neuroimaging studies of bipolar disorder. *Am J Psychiatry*. 165:313–320.
- Price JL, Drevets WC. 2010. Neurocircuitry of mood disorders. *Neuropsychopharmacol*. 35:192–216.
- Price JL, Drevets WC. 2012. Neural circuits underlying the pathophysiology of mood disorders. *Trends Cogn Sci*. 16:61–71.
- Purves D. 2008. Principles of cognitive neuroscience. Sunderland, MA: Sinauer Associates.
- Rajkowska G, Halaris A, Selemon LD. 2001. Reductions in neuronal and glial density characterize the dorsolateral prefrontal cortex in bipolar disorder. *Biol Psychiatry*. 49:741–752.
- Rajkowska G, Miguel-Hidalgo JJ, Wei J, Dilley G, Pittman SD, Meltzer HY, Overholser JC, Roth BL, Stockmeier CA. 1999. Morphometric evidence for neuronal and glial prefrontal cell pathology in major depression. *Biol Psychiatry*. 45:1085–1098.
- Redlich R, Almeida JJ, Grotegerd D, Opel N, Kugel H, Heindel W, Arolt V, Phillips ML, Dannlowski U. 2014. Brain morphometric biomarkers distinguishing unipolar and bipolar depression. A voxel-based morphometry-pattern classification approach. *JAMA Psychiatry*. 71:1222–1230.
- Redlich R, Opel N, Grotegerd D, Dohm K, Zaremba D, Burger C, Munker S, Muhlmann L, Wahl P, Heindel W, et al. 2016. Prediction of individual response to electroconvulsive therapy via machine learning on structural magnetic resonance imaging data. *JAMA Psychiatry*. 73:557–564.
- Rocha-Rego V, Jogia J, Marquand AF, Mourao-Miranda J, Simmons A, Frangou S. 2014. Examination of the predictive value of structural magnetic resonance scans in bipolar disorder: a pattern classification approach. *Psychol Med*. 44:519–532.
- Samame C, Szmulewicz AG, Valerio MP, Martino DJ, Strejilevich SA. 2017. Are major depression and bipolar disorder neuropsychologically distinct? A meta-analysis of comparative studies. *Eur Psychiatry*. 39:17–26.
- Sarrazin S, Poupon C, Linke J, Wessa M, Phillips M, Delavest M, Versace A, Almeida J, Guevara P, Duclap D, et al. 2014. A multicenter tractography study of deep white matter tracts in bipolar I disorder: psychotic features and interhemispheric disconnectivity. *JAMA Psychiatry*. 71:388–396.
- Savitz JB, Price JL, Drevets WC. 2014. Neuropathological and neuromorphometric abnormalities in bipolar disorder: View from the medial prefrontal cortical network. *Neurosci Biobehav Rev*. 42:132–147.
- Savitz JB, Rauch SL, Drevets WC. 2013. Clinical application of brain imaging for the diagnosis of mood disorders: the current state of play. *Mol Psychiatry*. 18:528–539.
- Schnack HG, Nieuwenhuis M, van Haren NE, Abramovic L, Scheewe TW, Brouwer RM, Hulshoff Pol HE, Kahn RS. 2014. Can structural MRI aid in clinical classification? A machine learning study in two independent samples of patients with schizophrenia, bipolar disorder and healthy subjects. *Neuroimage*. 84:299–306.
- Schneck CD, Miklowitz DJ, Miyahara S, Araga M, Wisniewski S, Gyulai L, Allen MH, Thase ME, Sachs GS. 2008. The prospective course of rapid-cycling bipolar disorder: findings from the STEP-BD. *Am J Psychiatry*. 165:370–377.
- Selvaraj S, Arnone D, Job D, Stanfield A, Farrow TF, Nugent AC, Scherk H, Gruber O, Chen X, Sachdev PS, et al. 2012. Grey matter differences in bipolar disorder: a meta-analysis of voxel-based morphometry studies. *Bipolar Disord*. 14:135–145.
- Silverstein WK, Noda Y, Barr MS, Vila-Rodriguez F, Rajji TK, Fitzgerald PB, Downar J, Mulsant BH, Vigod S, Daskalakis ZJ, et al. 2015. Neurobiological predictors of response to dorso-lateral prefrontal cortex repetitive transcranial magnetic stimulation in depression: a systematic review. *Depress Anxiety*. 32:871–891.
- Smith R, Chen K, Baxter L, Fort C, Lane RD. 2013. Antidepressant effects of sertraline associated with volume increases in dorsolateral prefrontal cortex. *J Affect Disord*. 146:414–419.
- Stonnington CM, Tan G, Kloppel S, Chu C, Draganski B, Jack CR Jr., Chen K, Ashburner J, Frackowiak RS. 2008. Interpreting scan data acquired from multiple scanners: a study with Alzheimer's disease. *Neuroimage*. 39:1180–1185.
- Szmulewicz AG, Valerio MP, Smith JM, Samame C, Martino DJ, Strejilevich SA. 2017. Neuropsychological profiles of major depressive disorder and bipolar disorder during euthymia. A systematic literature review of comparative studies. *Psychiatr Res*. 248:127–133.
- Tardif CL, Collins DL, Pike GB. 2010. Regional impact of field strength on voxel-based morphometry results. *Hum Brain Mapp*. 31:943–957.
- van Waarde JA, Scholte HS, van Oudheusden LJ, Verwey B, Denys D, van Wingen GA. 2015. A functional MRI marker may predict the outcome of electroconvulsive therapy in severe and treatment-resistant depression. *Mol Psychiatry*. 20:609–614.
- Videbech P, Ravnkilde B. 2004. Hippocampal volume and depression: a meta-analysis of MRI studies. *Am J Psychiatry*. 161:1957–1966.
- Vogt BA. 2009. Cingulate neurobiology and disease. Oxford: Oxford University Press.
- Wise T, Radua J, Via E, Cardoner N, Abe O, Adams TM, Amico F, Cheng Y, Cole JH, de Azevedo Marques Perico C, et al. 2017. Common and distinct patterns of grey-matter volume alteration in major depression and bipolar disorder: evidence from voxel-based meta-analysis. *Mol Psychiatry*. 22(10):1455–1463.
- Yatham LN, Kennedy SH, Parikh SV, Schaffer A, Beaulieu S, Alda M, O'Donovan C, Macqueen G, McIntyre RS, Sharma V, et al. 2013. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) collaborative update of CANMAT guidelines for the management of patients with bipolar disorder: update 2013. *Bipolar Disord*. 15:1–44.
- Yildiz-Yesiloglu A, Ankerst DP. 2006. Neurochemical alterations of the brain in bipolar disorder and their implications for pathophysiology: a systematic review of the in vivo proton magnetic resonance spectroscopy findings. *Prog Neuropsychopharmacol Biol Psychiatry*. 30:969–995.
- Zimmerman M, Ruggero CJ, Chelminski I, Young D. 2010. Psychiatric diagnoses in patients previously overdiagnosed with bipolar disorder. *J Clin Psychiatry*. 71:26–31.