Clinical utility of a short resting-state MRI scan in differentiating bipolar from unipolar depression


**Objective:** Depression in bipolar disorder (BipD) requires a therapeutic approach that is from treating unipolar major depressive disorder (UniD), but to date, no reliable methods could separate these two disorders. The aim of this study was to establish the clinical validity and utility of a non-invasive functional MRI-based method to classify BipD from UniD.

**Method:** The degree of connectivity (degree centrality or DC) of every small unit (voxel) with every other unit of the brain was estimated in 22 patients with BipD and 22 age, gender, and depressive severity-matched patients with UniD and 22 healthy controls. Pattern classification analysis was carried out using a support-vector machine (SVM) approach.

**Results:** Degree centrality pattern from 8-min resting fMRI discriminated BipD from UniD with an accuracy of 86% and diagnostic odds ratio of 9.6. DC was reduced in the left insula and increased in bilateral precuneus in BipD when compared to UniD. In this sample with a high degree of uncertainty (50% prior probability), positive predictive value of the DC test was 79%.

**Conclusion:** Degree centrality maps are potential candidate measures to separate bipolar depression from unipolar depression. Test performance reported here requires further pragmatic evaluation in regular clinical practice.

**Significant outcomes**
- An 8-min resting scan approach can successfully discriminate bipolar depression from unipolar depression with high degree of accuracy (86%).
- Bipolar depression and unipolar depression have 70% dissimilarity of illness-specific connectivity profile and only 30% similarity of transdiagnostic dysconnectivity.
- A reduction in the centrality of insula along with an increase in centrality of precuneus uniquely relates to the pathophysiology of bipolar depression.

**Limitations**
- Lacks external validation set, so results are based on a single dataset collected at the same site.
- A number of patients were on mood stabilizer and/or antidepressants at the time of scan.

---


1. Mental Health Center, West China Hospital of Sichuan University, Chengdu, Sichuan, China
2. State Key Laboratory of Biotherapy, Psychiatric Laboratory, West China Hospital, Sichuan University, Chengdu, Sichuan, China
3. West China Brain Research Center, West China Hospital of Sichuan University, Chengdu, China
4. Robarts Research Institute & The Brain and Mind Institute, University of Western Ontario, London, Ontario, Canada
5. Department of Psychiatry, University of Western Ontario, London, ON, Canada
6. Lawson Health Research Institute, London, ON, Canada

Key words: bipolar depression; unipolar depression; degree centrality; diagnostic accuracy; pattern classification

Tao Li, The Mental Health Center and the Psychiatric Laboratory, West China Hospital, Sichuan University, 29# Dianxin Nanjie, Wuhou District, Chengdu, Sichuan 610041, China.
E-mail: litaohx@scu.edu.cn

Lena Palaniyappan, Prevention & Early Intervention Program for Psychoses (PEPP), 2A:638, LHSC-VH, 800 Commissioners Road, London, ON N6A 5W9, Canada.
E-mail: lpalaniy@uwo.ca

Accepted for publication April 18, 2017
Introduction

Depression is a common clinical manifestation in both bipolar disorder and major depressive disorder, but the two disorders require distinct treatment planning and service delivery (1, 2). The misdiagnosis of BipD as UniD leads to inappropriate prescription, iatrogenic mood instability and manic switching to mania, increased suicidality, and a profoundly negative personal impact, and consequently, higher economic burden (3, 4). However, only one-fifth of patients with bipolar disorder who are experiencing a depressive episode receive the correct diagnosis within the first year of seeking treatment (5). There are no tools available to a clinician that can help separate these two disorders when a patient presents for the first time with an episode of depression. This is one area of psychiatry where diagnostic tools are direly needed, and if found, could make a significant impact on clinical services.

Certain clinical characteristics such as early age of onset, presence of psychotic symptoms, positive family history of bipolar disorder, and mood instability increase the likelihood of diagnosing BipD in a depressed subject (6, 7). Concerted efforts to enable early diagnosis of BipD have yielded rating scales that include many of these clinical features, for example, the Screening Assessment of Depression Polarity (8), the Hypomania Checklist (9), the Mood Disorder Questionnaire (10), and the Bipolar Spectrum Diagnostic Scale (11). Despite the demonstrated validity, scales based on clinical descriptors continue to lack sufficient specificity to reduce diagnostic uncertainty and support clinical decision-making. In recent years, several promising results have identified notable differences in the neural circuitry of BipD and UniD using neuroimaging (12). This has raised the possibility that brain connectivity profiling could provide a means to diagnostic separation of these two disorders.

Neuroimaging investigations of the distinction between BipD and UniD have indicated distributed differences in connectivity that includes medial temporal (amygdala), subcortical, orbitofrontal, medial prefrontal, lateral prefrontal, and anterior cingulate regions (13–16). While the effect sizes of individual regional differences have not been large enough to be of diagnostic use, the application of multivariate pattern classification has demonstrated success in separating BipD from UniD, with diagnostic accuracies varying between 69% and 92% achieved in prior studies using various tools such as quantitative EEG (80.2% accuracy) (17), arterial spin labeling (81% accuracy) (18), emotional face processing task performance (90% accuracy) (19), using thickness and area measures from structural imaging (74.3% accuracy) (20), combined structural and functional data from preselected networks (69% accuracy) (21), or atlas-based parcellations (92.1% accuracy) (22). The above studies use specific tools (e.g., ASL, task-based fMRI) or use sophisticated analytical approaches (e.g., surface-based morphometry) that are either labor-intensive or not immediately available to many clinical centers.

A number of factors are crucial in determining the diagnostic utility of a promising biomarker for BipD. The marker must be easy to obtain, preferably non-invasive, based on a widely available instrument, cost-effective and should possess a high diagnostic accuracy. In particular, it should demonstrate notable precision when applied to patients that present with high levels of diagnostic uncertainty, that is, during a depressive episode with no mixed or manic features. Analytical validation when facing high clinical uncertainty is a crucial step that can help filtering out poorly performing markers using cross-sectional studies, before investing on long-term follow-up efforts aimed at establishing predictive validity and cost-effectiveness using a controlled diagnostic trial design.

In this study, we use resting-state fMRI wherein the confound of task performance and cognitive strategy is eliminated. Prior resting fMRI studies have shown promising differences in the connectivity involving the insula, prefrontal cortex, and amygdala, although the results are not consistent, with some studies suggesting higher connectivity (23, 24) and others suggesting reduced connectivity in bipolar depression (25). To resolve this, we use an intuitive index of the state of connectivity of the various spatial units (voxels in fMRI) of the brain. We count the number of connections that each voxel makes with every other voxel in the brain during resting-state fMRI and use this voxelwise degree centrality (DC) for pattern classification analysis. DC can be readily obtained using fully automated routines using ~5 min of resting fMRI; it is not constrained by a priori regional selection or network definitions and thus demonstrates high reliability (26). DC can be obtained without manual estimation or image editing—features aiding its wide usage in quantifying global connectivity states (26, 27); thus, it is well suited for
multivariate pattern recognition tests aimed at diagnostic separation. DC is sensitive to the diagnostic status of MDD and bipolar disorder (28–33). In a prior study, we have shown that DC patterns (‘maps’) differ significantly between bipolar disorder with psychosis and schizophrenia, highlighting its diagnostic potential (31). However, to date, no study directly compared the degree centrality maps of individual in BipD and UniD.

Aims of the study

In this study, we used degree centrality maps to separate bipolar depression from unipolar depression using pattern classification approach. In addition to deriving traditional measures of test performance (accuracy, diagnostic odds ratio, and predictive values), we also derived the number needed to diagnose for the degree centrality maps and quantified the overall reduction in uncertainty achieved using this approach in clinical practice.

Material and methods

Patients presenting with an episode of depression and seeking treatment were recruited from the Mental Health Center of West China Hospital of Sichuan University. None of patients with BipD but some patients with UniD were first episode. Subjects were rated on the 17-item Hamilton Depression Rating Scale (HDRS) and Young Mania Rating Scale (YMRS) at the time of the baseline scan. To increase the baseline diagnostic uncertainty, only the patients with HDRS score >17 and YMRS score <6 were recruited in this study. A diagnosis of with BipD (n = 22) or UniD (n = 22) was based on the Structured Clinical Interview for DSM-IV (SCID-IP) (34). Patients with other mental diseases, nervous system diseases, severe physical diseases, personality disorders, and abuse of alcohol and drugs were excluded. The total defined daily dose (DDD) (35) was calculated separately for antipsychotics, mood stabilizers, and antidepressants. Healthy controls (n = 22), who were interviewed using the Structured Clinical Interview for DSM-IV, non-patient edition (36), were recruited from the same area. The short version of the seven-subtest (information, arithmetic, digital symbol, digital span test, block design, picture completion, and similarities) Chinese Revised Version of the Wechsler Adult Intelligence Scale (WAIS-RC) (37) was used to test the verbal intelligence quotient (VIQ), performance intelligence quotient (PIQ), and intelligence quotient (IQ) of participants.

All participants were Han Chinese and right-handed. All subjects signed an informed consent form prior to participation in the study. This study was approved by the Ethics Committee of Sichuan University and was conducted according to the Helsinki Declaration.

MRI data acquisition

All scanning was performed on a 3.0 T MR scanner (Achieva; Philips, Amsterdam, the Netherlands) using an eight-channel phased-array head coil. Foam padding and earplugs were used to minimize head movement and scanner noise. During scanning, participants were often reminded to remain motionless with eyes closed, without falling asleep, and without thinking of anything in particular (confirmed by subjects immediately after the experiment).

High-resolution T1 images were acquired by 3D magnetization-prepared rapid gradient-echo sequence as follows: repetition time 8.37 ms, echo time 3.88 ms, flip angle 7°, in-plane matrix resolution 256 × 256, field of view 24 × 24 cm², and number of slices 188.

A total of 240 volumes of echo-planar images were obtained axially with a gradient-echo echo-planar imaging sequence with the following parameters: repetition time 2000 ms, echo time 3.711 ms, flip angle 7°, in-plane matrix resolution 256 × 256, field of view 256 × 256 mm², and number of slices 38. None of the participants had more than 2 mm maximum displacement in x, y, or z and 2° of angular motion during the whole MRI scan. For each participant, the resting-state functional magnetic resonance imaging (rsfMRI) scanning lasted for 8 min and 6 s, and 240 volumes were obtained.

Resting-state functional imaging preprocessing

rsfMRI image processing was carried out using Statistical Parametric Mapping (SPM8, http://www.fil.ion.ucl.ac.uk/spm) and Data Processing Assistant for Resting-State fMRI (DPARSF) (38). The first 10 time points were removed to allow the fMRI signal to reach steady state. Raw rsfMRI images were first slice time corrected and realigned and were subsequently unwrapped to correct for susceptibility-by-movement interaction. One-way analysis of variance used to estimate the percent of framewise displacement >0.5 mm (39) in group of patients with bipolar depression, group of patients with unipolar depression, and group of healthy controls, and no significant head movement
difference was found \((F = 0.726, \ P = 0.488)\) among the three groups (40). The data were scrubbed based on a measure of average framewise displacement (FD) with \((FD \leq 0.5 \ \text{mm})\) using a nearest neighbor interpolation approach to further correct for movement artifacts (39). Next, each image volume was spatially normalized using T1 image of DATEL segment, and nuisance covariates including Friston 24 motion parameters, cerebrospinal fluid (CSF), and white matter signals were regressed out (41, 42). All images were linearly detrended and bandpass-filtered (0.01–0.10 Hz) to eliminate high-frequency physiological noise (43). The head motion and rotation of all the subjects (maximal motion between volumes in each direction, and rotation about each axis) are \(<2 \ \text{mm} \ \text{and} \ <2^\circ\).

Degree centrality calculation

Degree centrality measure was calculated using DPARSF. For each voxel, the BOLD time course was extracted and correlated with every other voxel in the brain. In line with previous studies (27, 31), the correlation matrix was binarized by thresholding at \(r > 0.25\) before counting the number of connections to generate voxelwise DC. For each subject, a map with DC values for every voxel was obtained. These maps were normalized to \(z\) maps using the mean value and standard deviation within the whole gray matter mask (26). Finally, the resulting DC maps were smoothed with a Gaussian kernel (full-width half maximum \(= 6 \ \text{mm}\)) to enable group comparisons.

Statistical analysis

Demographic characteristics. Chi-square test for categorical data, one-way ANOVA, and independent two-sample \(t\)-test for continuous variables were used to evaluate differences in demographic characteristics among three groups or between two patients group with threshold \(P < 0.05\) in spss 18.0.

Classification of bipolar depression and unipolar depression. A linear kernel support-vector classifier (SVM) implemented in the Pattern Recognition for Neuroimaging Toolbox version 2.0 (PRoNTo) (http://www.mlnl.cs.ucl.ac.uk/pronto) (44) was used to classify bipolar depression and unipolar depression using DC maps.

A linear kernel was used with default settings and the parameter \(C = 1\), in line with which is recommended for high-dimensional data and relatively small sample sizes (45, 46). To validate and assess performance of the classifier, cross-validation was performed in a double-nested loop. The inner loop was used using a leave one subject per group out. The outer loop for assessing the model’s performance also used a leave one subject per group out cross-validation approach, which indicated the classifier is trained on all subjects except one. Statistical significance was estimated by performing permutation tests (1000 iterations) using \(P < 0.05\). We report the area under the ROC (receiver operating characteristics) curve [AUC] as the measure of accuracy, as this is more robust than the numerical estimation when employing learning algorithms (47). We also derived specificity, sensitivity, positive predictive value (PPV), negative predictive value (NPV), predictive summary index (PSI), number needed to diagnose (NND), diagnostic odds ratio (DOR), and proportionate reduction in uncertainty (PRU) (48–50) as described in the supplement from the resulting values (Table S1). PSI provides a measure of gain in certainty when a diagnostic tool is used in a target population (48). NND is the number of patients required to be tested in order to correctly diagnosis in one person using a Bayesian measure of diagnostic certainty (48). DOR is a composite indicator of test performance that summarizes sensitivity and specificity without being affected by variations in disease prevalence, and the higher the value, the more superior is the discriminatory performance (50).

From the SVM features, corresponding weight images were computed and extracted to visualize contributing brain areas. As SVM analysis uses a multivariate approach, all voxels in the gray matter mask carry a certain weight value signifying its contribution to classification; therefore, discriminations are based on the global spatial pattern, and local inferences should not be made in regards to the weights. However, for the sake of clarity when visualizing the regions with relatively higher contributions to the decision function, in line with our prior study (51), a threshold of 30% of the maximum positive and negative weight values was set, presenting the spatial pattern of the regions that most contributed to the group discrimination. Note that the decision function for the statistical discrimination is independent of this visualization threshold. These maps are illustrated in Fig. 1.

Differences in degree centrality among groups. To determine the core cortical hub architecture, we identified the significant clusters with high DC in healthy controls using one sample \(t\)-test with the threshold of family error wise (FWE) corrected.
type-1 error rate of 5% at a voxelwise threshold of $P < 0.001$. A full-factor design with sex, age, and education years as covariates was used to compare the DC maps among the three groups using SPM 12, and a significant difference was set as the threshold of FWE-corrected type-1 error rate of 5% at a voxelwise threshold of $P < 0.001$.

To exclude the effect of duration of illness, additional two-sample $t$-test with sex, age, education years, and duration of illness as covariates was used to compare the DC maps between BipD group and UniD group in SPM 12, and a significant difference was set as the threshold of FWE-corrected type-1 error rate of 5% at a voxelwise threshold of $P < 0.001$. Result was showed in Data S1.

Degree of overlap between bipolar depression and unipolar depression. To analyze the overlap abnormalities found in BipD and UniD, a conjunction measure called Dice coefficient of similarity (DCS) (52) was used in this study. First, an intersection (overlap) mask and a combination (union) mask for the contrasts BipD vs controls and UniD vs controls were derived with an uncorrected threshold of $P < 0.001$. Then, the

Fig. 1. The weight images (per voxel) for classification model based on degree centrality maps. Discrimination maps at a threshold of 30% of the maximum positive and negative weight values contributed to classify bipolar depression from unipolar depression (a), bipolar depression from healthy controls (b), and unipolar depression from healthy controls (c). It superimposed onto a standard brain template provided by MRICron.
DCS was calculated between the two groups (DSC = spatial extent of the intersection map /2* combination mask). A DCS value of 100% means that both the disorders have perfect spatial agreement in the distribution of abnormalities across the brain (31).

Effects of medication and clinical characteristic on degree centrality. To test whether the differences in DC maps between BipD and UniD were related to the medication, number of depressive episodes, duration of current episode, and total duration of illness, two partial correlation analyses were processed in spss 18.0. First, partial correlation analysis was used to test the relationship between the value extracted from the clusters showing significant differences in the contrast of BipD vs UniD and the total DDD of all medication in 23 patients, total DDD of antipsychotic in 11 patients, total DDD of mood stabilizer in 10 patients, total DDD of antidepressants in 16 patients, with number of depressive episodes, duration of current episode, total duration of illness, age, gender, and education years as covariates. Second, partial correlation analysis was also used to test the relationship between the value extracted from the clusters showing significant differences in the contrast of BipD vs UniD and number of depressive episodes, duration of current episode, total duration of illness with age, gender, and education years as covariates. A significant correlation was set at a threshold $P < 0.05$, Bonferroni-corrected. In addition, the direct multiple regression analysis was processed between DC and four types DDD with age, gender, and education years as covariates using voxel-based analysis on SPM12, and a significant correlation was set as the threshold of FWE-corrected type-1 error rate of 5% at a voxelwise threshold of $P < 0.001$.

Results

Demographic and clinical characteristics

The demographic characteristics of the subjects are shown in Table 1. No significant differences were found in age, gender, years of education, VIQ, PIQ, and IQ among three groups. No significant differences were found in total duration of illness and HDRS score between BipD group and UniD group, while BipD has shorter duration of current episode, higher number of depressive episodes, total DDD, and YMRS score than UniD (Table 1).

Classification

Support-vector machine using DC maps was able to differentiate BipD from UniD with an accuracy of 86% ($P = 0.002$), BipD from HC with an accuracy of 81% ($P = 0.007$); UniD from HC with an accuracy of 94% ($P = 0.001$) (Table 2, Figure S1). The specificity, sensitivity, PPV, NPV, PSI, NND, and DOR to classify participants of three groups also showed in Table 2. The proportionate reduction in uncertainty for positive test for various hypothetical prevalence values is listed in Table S2 and plotted in Figure S2. The spatial loading of the maximum positive and negative weight values (thresholded at 30%) inciting each region’s contribution to the diagnostic separation is shown in Fig. 1. The prediction plot for SVM-based classification is shown in Figure S3.

Differences in degree centrality among groups

Voxelwise comparisons revealed significant differences among the groups in the DC of several brain regions. Compared with UniD, patients with BipD had lower DC in left insula and increased DC in bilateral precuneus extending to middle and postcingulate gyrus, and left declive of the cerebellum. Both BipD and UniD groups had lower DC in bilateral fusiform, lingual gyrus, precuneus, and higher DC in orbital frontal cortex when compared to healthy controls. When compared to healthy controls, only patients with UniD had lower DC in bilateral cerebellar regions and higher DC in bilateral insular, right inferior temporal, and caudate regions (Fig. 2, Table S3). The F-test of the differences in DC map among three groups is shown in Figure S4.

Degree of overlap between bipolar depression and unipolar depression

The DCS test revealed only 30.17% overlap in the topography of degree centrality abnormalities between BipD and UniD (Table S4), indicating that voxelwise DC maps of the two disorders had a large proportion of unique, disease-specific information (69.83%).

Effects of medication and clinical characteristic on degree centrality

No statistically significant correlations were found between different degree centrality in BipD vs UniD and total DDD of all medicine, total DDD of antipsychotic, total DDD of mood stabilizer, and total DDD of antidepressants;
different degree centrality in BipD vs UniD was also not significantly correlated to number of depressive episodes, duration of current episode, total duration of illness (Table S5). Voxel-based analyses of multiple regression showed the total DDD of mood stabilizer was positively correlated to right angular region ($T = 5.8934$, voxel $= 28$; peak MNI coordinates, $x = 48$, $y = -54$, $z = 30$) at a threshold of FWE-corrected type-1 error rate of 5% at a voxelwise threshold of $P < 0.001$ (Figure S5).

**Discussion**

To our knowledge, this is the first study to use a whole-brain connectivity index to discriminate bipolar from unipolar depression. We report three main findings: First, the DC maps successfully discriminate the two disorders with high degree of accuracy despite the high degree of diagnostic uncertainty in the samples matched for their depression severity. The accuracy of 86% (bootstrapped $P = 0.002$), combined with DOR of 9.64, improves the pretest probability of 50% to a post-test probability of 79%, thus reducing nearly 58% of uncertainty associated with diagnosing BipD. Second, the whole-brain connectivity profile of the two disorders shows some shared, possibly transdiagnostic features (30% similarity) although a large proportion of the centrality profile appears to be illness-specific (70% dissimilarity). This emphasizes that despite the phenotypic similarity of depressive episodes, neurobiological differences between these two disorders are large enough to be exploited for clinical separation. Finally, we note that a reduction in the centrality of insula along with an increase in centrality of precuneus uniquely relates to the pathophysiology of bipolar disorder.

The discriminatory accuracy of 86% is higher than several (17, 18, 21, 53) previous pattern recognition studies separating UniD and BipD, while two other studies have reported better performance of 90–91% by utilizing multimodal data or task-based fMRI (19, 22). We deliberately matched our two patient groups for the severity of depression, thus eliminating any artifactual classification based on differences in symptom severity. Only one prior study had utilized a similar approach (19). Our exclusive focus on the depressive phase possibly contributed to a reduction in heterogeneity which can lead to poor fitting of classifier in pattern recognition (54). DOR is a useful index for comparing classifier performance reported in a variety of clinical settings, and 9.64 of DOR in

### Table 1. Demographics and clinical information

<table>
<thead>
<tr>
<th></th>
<th>Bipolar_depression ($n = 22$)</th>
<th>Unipolar_depression ($n = 22$)</th>
<th>Control ($n = 22$)</th>
<th>df</th>
<th>$F$/$T$/$\chi^2$ Value</th>
<th>$P$ Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>28.73 (10.11)</td>
<td>27.68 (8.65)</td>
<td>28.27 (9.55)</td>
<td>42</td>
<td>0.068</td>
<td>0.93</td>
</tr>
<tr>
<td>Years of education</td>
<td>14.09 (3.07)</td>
<td>12.73 (3.65)</td>
<td>14.68 (3.15)</td>
<td>42</td>
<td>2.03</td>
<td>0.72</td>
</tr>
<tr>
<td>Gender (male/female)</td>
<td>9/13</td>
<td>9/13</td>
<td>2</td>
<td>2</td>
<td>0.12</td>
<td>0.94</td>
</tr>
<tr>
<td>VIQ</td>
<td>112.42 (12.13)</td>
<td>105.32 (18.81)</td>
<td>111.82 (12.76)</td>
<td>2</td>
<td>1.42</td>
<td>0.25</td>
</tr>
<tr>
<td>PIQ</td>
<td>105.52 (15.41)</td>
<td>101.74 (15.38)</td>
<td>104.73 (12.71)</td>
<td>2</td>
<td>0.23</td>
<td>0.90</td>
</tr>
<tr>
<td>IQ</td>
<td>106.52 (13.98)</td>
<td>104.11 (17.85)</td>
<td>109.59 (12.22)</td>
<td>2</td>
<td>0.92</td>
<td>0.40</td>
</tr>
<tr>
<td>Duration of current episode (months)</td>
<td>2.14 (1.39)</td>
<td>4.50 (3.02)</td>
<td>2.05 (1.39)</td>
<td>2</td>
<td>29.52</td>
<td>3.34</td>
</tr>
<tr>
<td>Total duration of illness (months)</td>
<td>49.32 (53.75)</td>
<td>23.36 (35.52)</td>
<td>42</td>
<td>1.89</td>
<td>0.066</td>
<td></td>
</tr>
<tr>
<td>No. of depressive episodes</td>
<td>2.18 (0.96)</td>
<td>1.50 (0.80)</td>
<td>2</td>
<td>42</td>
<td>2.56</td>
<td>0.014</td>
</tr>
<tr>
<td>Mood stabilizer (Yes/No)</td>
<td>10/12</td>
<td>0/22</td>
<td>2</td>
<td>1</td>
<td>12.94</td>
<td>0.000</td>
</tr>
<tr>
<td>Antidepressants (Yes/No)</td>
<td>10/12</td>
<td>1/21</td>
<td>2</td>
<td>1</td>
<td>9.818</td>
<td>0.002</td>
</tr>
<tr>
<td>Total DDD (IU)</td>
<td>19.244</td>
<td>12.25</td>
<td>2</td>
<td>21</td>
<td>1.145</td>
<td>0.265</td>
</tr>
<tr>
<td>HDRS score</td>
<td>20.82 (3.11)</td>
<td>20.62 (2.97)</td>
<td>42</td>
<td>0.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>YMRS score</td>
<td>2.95 (3.48)</td>
<td>0.32 (1.49)</td>
<td>28.45</td>
<td>3.26</td>
<td>0.003</td>
<td></td>
</tr>
</tbody>
</table>

VIQ, verbal intelligence quotient; PIQ, performance intelligence quotient; IQ, intelligence quotient; YMRS, Young Mania Rating Scale; HDRS, Hamilton Depression Rating Scale; df, degrees of freedom; No, number; mean (standard variance); DDD, defined daily dose.

### Table 2. Classification results for support-vector machine using degree centrality map

<table>
<thead>
<tr>
<th></th>
<th>BipD vs UniD</th>
<th>BipD vs Control</th>
<th>UniD vs Control</th>
<th>Accuracy</th>
<th>ROC (%)</th>
<th>Sensitivity (%)</th>
<th>Specificity (%)</th>
<th>Positive Predictive value (%)</th>
<th>Negative Predictive value (%)</th>
<th>$P$ value</th>
<th>PSI</th>
<th>DOR</th>
<th>NND</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive Predictive value (%)</td>
<td>78.95</td>
<td>72.00</td>
<td>90.48</td>
<td>88.96</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative Predictive value (%)</td>
<td>73.91</td>
<td>76.19</td>
<td>90.48</td>
<td>88.96</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td>0.002</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

BipD, bipolar depression; UniD, unipolar depression; ROC, receiver operator curve; PSI, predictive summary index; DOR, diagnostic odds ratio; NND, number needed to diagnosis.
classifying BipD from UniD in present study suggests that the likelihood of DC-based MRI test being ‘positive’ in a patient with BipD is ~10 times higher than the same result being reported in a patient with UniD. Consequently, resting-state DC mapping has a potential to complement clinical practice. NND of 2 indicates that only two subjects would be required to undergo the DC-based resting fMRI test to correctly diagnose one person (48). These indirectly indicate a high probability for DC-based test to be cost-effective when used to differentiate BipD from UniD.

We noted an accuracy of 94% for discriminating UniD from healthy controls; this is much higher than what has been achieved using other modalities of neuroimaging (55). While such a distinction has less translational value than separating BipD from UniD, the test performance measures for UniD vs HC comparison highlight the ability of DC maps to extract illness-specific connectivity profiles rather accurately. As the occurrence of isolated depressive symptoms is fairly common among otherwise healthy individuals, it is worth employing DC maps to discriminate subjects with subthreshold depression (that may not require antidepressant treatment) from those who may require more active interventions.

In our sample, the DC maps of depressed subjects with UniD varied notably from those with BipD, with Dice coefficient of similarity being much smaller (30%) than that would be obtained if there was complete overlap in the DC pattern of the two groups. From the multivariate weight maps based on the decision function classifying UniD and BipD, the dissimilarity is evident in left insula, bilateral precuneus, middle and postcingulate gyrus, and the declive of cerebellum. The pattern of reduced insular centrality and increased precuneus and cingulate gyrus centrality in BipD at rest is interesting, given the attention modulating role of insula, a central node in the Salience Network, and the introspective role of precuneus.
Our findings of the loss of insula centrality and increase in precuneus centrality in BipD compared to UniD align with several task fMRI studies noting reduced insula activity in bipolar disorder (56, 57), and reduced functional connectivity of precuneus in UniD (58, 59). In terms of the triple network model (60), if this pattern of dysconnectivity relates to everyday cognitive functions, this represents a relative inability to swiftly switch between introspective and external information, especially when task-related demands are placed in BipD. We speculate that this may explain the somewhat higher cognitive burden seen in bipolar disorder compared to unipolar disorder (61).

We also observed some shared connectivity features in both BipD and UniD when compared to healthy controls. Decreased DC in primary visual cortex, fusiform/lingual gyrus, and increased DC in orbital frontal cortex were noted in both BipD and UniD group, consistent with prior reports (62–68). The reduced frequency of functional connectivity of fusiform and lingual gyrus with rest of the brain supports the notion of sensory processing deficits in both BipD and UniD, which may be related to the range of cognitive and emotional symptoms (e.g., anhedonia, dull perception, apathy) seen in depressive states (69). Increased DC in both UniD and BipD was found in orbital frontal cortex (OFC). Ochsner and Gross hypothesized that OFC regulates emotions through a top-down outcome-based appraisal process (70). We speculate that an increase in information flow from or to OFC may indeed compensate for the deficits in integrating sensory information from defective fusiform and lingual regions during a depressive state.

A number of limitations must be taken into account when considering our results. First, our results are based on a single dataset collected at the same site; we lack external validation set. Nevertheless, our ‘test’ data have been validated using a jackknife leave-one-out cross-validation procedure, which has been shown to be robust and thus widely used in MRI-based pattern classification studies (71). Second, a number of our patients were on mood stabilizer and/or antidepressants at the time of scan. We addressed the possible confounding effect of medications by quantifying the relationship between dose and DC maps in each group, for each class of medication used. We noted a significant relationship between DC of angular gyrus and mood-stabilizer dose among patients with BipD, but the spatial location of this association was small, and non-overlapping with the contrast between BipD and UniD. A detailed analysis of the effect of drug dosages to the DC value of the clusters that emerged as significantly different between BipD and UniD produced no notable associations, increasing our confidence that the observed results are not attributable to treatment differences. Finally, patients included in this study were not from a first-episode cohort with the greatest clinical need for a diagnostic test, but included patients in who a definite diagnosis of bipolar disorder could be made using clinical interview based on SCID. Selecting cases with well-defined group membership is essential at this stage of concurrent analytical validation using a supervised test algorithm; given the strong diagnostic signal reported here, in future longitudinal studies estimating predictive validity can be planned.

Substantial degree of uncertainty that surrounds making a diagnosis of bipolar disorder in depressed subjects can be addressed using a well-tolerated, readily accessible, safe and brief fMRI scan. If this test can be used in a sequence with other widely available clinical scales of bipolarity, the overall gain in diagnostic confidence could surpass the utility of highly variable clinical intuition. We call for further focused efforts in this direction.

Acknowledgements

This work was partly supported by National Nature Science Foundation of China (91332205 to TL, 81501174 to MLL); 1.3.5 Project for Disciplines of Excellence, West China Hospital, Sichuan University (ZY2016203, ZY2016103); Burke Family Fund, Schulich School of Medicine, University of Western Ontario (LP); Opportunities Fund, Academic Medical Organization of South Western Ontario (LP).

Declaration of interest

LP received Travel Support from Magstim Limited (2014) and in the last 2 years has held shares of Shire Inc. and Glaxo Smith Kline in his/spousal pension funds.

References

5. Hirschfeld RM, Lewis L, Vornik LA. Perceptions and impact of bipolar disorder: how far have we really come?


Differentiate bipolar from unipolar depression


Supporting Information

Additional Supporting Information may be found in the online version of this article:

Figure S1. The receiver operating characteristic (ROC) curve for the classification model based on degree centrality maps.

Figure S2. The proportionate reduction in uncertainty scores (PRU) for diagnosing a bipolar depression in a depressive patient as a result of carrying out a MRI test, according to the estimated likely prevalence of bipolar depression in samples of patients with a depressive episode.

Figure S3. Prediction plot for classification modelled by support vector machine using degree centrality maps.

Figure S4. The difference in degree centrality maps among the 3 groups (ANOVA contrast from main effects of group in full-factorial design in SPm software).

Figure S5. Effects of mood stabilizer on degree centrality. Table S1. Test performance measures.
Table S2. The impact of positive and negative MRI test results on the PRU of a depressive patient is bipolar depression, according to the prior probability (likely clinical prevalence) of them having a positive diagnosis.
Table S3. Degree centrality differences between patients and controls.
Table S4. Dice coefficients of spatial overlap between the bipolar depression and unipolar depression in contrasts with healthy controls.
Table S5. The correlation between different degree centrality in BipD vs UniD and total defined daily dose (DDD).
Data S1. Additional result for differences degree centrality between BipD group and UniD group.