◆ Human Brain Mapping 35:1630–1641 (2014) ◆

Unsupervised Classification of Major Depression Using Functional Connectivity MRI

Ling-Li Zeng,[1] Hui Shen,[1] Li Liu,[2] and Dewen Hu[1]*

[1]*College of Mechatronics and Automation, National University of Defense Technology, Changsha, Hunan, People's Republic of China*
[2]*Department of Psychiatry, First Affiliated Hospital, China Medical University, Shenyang, Liaoning, People's Republic of China*

Abstract: The current diagnosis of psychiatric disorders including major depressive disorder based largely on self-reported symptoms and clinical signs may be prone to patients' behaviors and psychiatrists' bias. This study aims at developing an unsupervised machine learning approach for the accurate identification of major depression based on single resting-state functional magnetic resonance imaging scans in the absence of clinical information. Twenty-four medication-naive patients with major depression and 29 demographically similar healthy individuals underwent resting-state functional magnetic resonance imaging. We first clustered the voxels within the perigenual cingulate cortex into two subregions, a subgenual region and a pregenual region, according to their distinct resting-state functional connectivity patterns and showed that a maximum margin clustering-based unsupervised machine learning approach extracted sufficient information from the subgenual cingulate functional connectivity map to differentiate depressed patients from healthy controls with a group-level clustering consistency of 92.5% and an individual-level classification consistency of 92.5%. It was also revealed that the subgenual cingulate functional connectivity network with the highest discriminative power primarily included the ventrolateral and ventromedial prefrontal cortex, superior temporal gyri and limbic areas, indicating that these connections may play critical roles in the pathophysiology of major depression. The current study suggests that subgenual cingulate functional connectivity network signatures may provide promising objective biomarkers for the diagnosis of major depression and that maximum margin clustering-based unsupervised machine learning approaches may have the potential to inform clinical practice and aid in research on psychiatric disorders. *Hum Brain Mapp 35:1630–1641, 2014.* © 2013 Wiley Periodicals, Inc.

Key words: major depression; functional connectivity magnetic resonance imaging; unsupervised classification; resting-state; subgenual cingulate; maximum margin clustering

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

Contract grant sponsor: National Basic Research Program of China; Contract grant numbers: 2011CB707802 and 2013CB329401; Contract grant sponsor: National High-tech Program of China; Contract grant number: 2012AA011601; Contract grant sponsor: National Science Foundation of China; Contract grant number: 61003202; Contract grant sponsors: Hunan Provincial Innovation Foundation for Postgraduate; Contract grant number: CX2011B015; Contract grant sponsor: Graduate Innovation Fund of National University of Defense Technology; Contract grant number: B110304.

*Correspondence to: Dewen Hu, College of Mechatronics and Automation, National University of Defense Technology, Changsha, Hunan 410073, China. E-mail: dwhu@nudt.edu.cn

Received for publication 22 October 2012; Revised 19 January 2013; Accepted 7 February 2013

DOI: 10.1002/hbm.22278

Published online 24 April 2013 in Wiley Online Library (wileyonlinelibrary.com).

INTRODUCTION

In recent years, interest in exploring the brain variance between patients with psychiatric disorders and healthy controls using machine learning methods based on neuroimaging data has increased [Oquendo et al., 2012; Pereira et al., 2009]. Potential neuroimaging-based biomarkers that highlight the pathophysiology of psychiatric disorders can be found by investigating disease-related structural and functional brain abnormalities [Kawasaki et al., 2007; Shen et al., 2010a; Uddin et al., 2012; Zeng et al., 2012]. There are two main approaches within machine learning: supervised and unsupervised learning [Bishop, 2006; Orrù et al., 2012]. Supervised learning techniques can map two or more sets of observations with two or more categories based on the labeled data, while unsupervised learning techniques can be used to determine how the data are organized when a priori labeling information is unavailable. To the best of our knowledge, most of the previous classification studies of brain imaging have focused on supervised machine learning techniques, which require the prior diagnostic labeling of prototype subjects. Increasing evidence has demonstrated that supervised classification approaches that are dependent on prior clinical behavioral scores may be prone to inter-user bias [Hardoon et al., 2007]. First, different patients are likely to have different ways of expressing symptoms and complaints, which may affect the clinicians' interpretations of the patients' behavior [Buckner et al., 2007]. Second, some evidence indicates that clinicians might be vulnerable to engaging in personal or stereotyped thinking, which could affect important aspects of diagnoses [Buckner et al., 2007; Caplan and Cosgrove, 2004; Sue, 1998], despite their best efforts to control prejudice due to diverse cultures, languages, values, customs, beliefs and practices. Hence, it is important to develop an unsupervised classification framework that is independent of prior clinical diagnoses to escape the above biases [Bishop, 2006]; to date, however, little attention has been paid to unsupervised classification approaches for disease state prediction from neuroimaging data.

Unsupervised machine learning approaches can discover homogeneous groups (called clusters) of similar samples within the data and flexibly determine the complex and unknown distribution or pattern of data within the input space without prior knowledge [Bishop, 2006]. Therefore, the unsupervised classification of brain imaging data may provide more objective neuroimaging evidence for the diagnosis of psychiatric disorders compared to supervised classification based on clinical scores. C-means and normalized cuts (Ncut) are two widely used clustering algorithms [Shi and Malik, 2000]. C-means performs best if the data have a Gaussian distribution, but it cannot be guaranteed that brain-imaging data are Gaussian in feature space. The Ncut algorithm is very robust to outliers and exhibits favorable performance relative to other graph clustering methods [Shen et al., 2010b; Shi and Malik, 2000], but it does not always lead to particularly good solutions [Luxburg, 2007]. Maximum margin clustering (MMC), a recently proposed powerful clustering algorithm, is very robust to the data distribution and outperforms other popular algorithms, including both C-means and Ncut, in two-class clustering analysis [Ding et al., 2009; Wang et al., 2010; Xu et al., 2005; Xu and Schuurmans, 2005]. This study attempted to develop an MMC-based unsupervised machine learning approach for accurately identifying major depressive disorder (MDD) on the basis of single resting-state functional magnetic resonance imaging (fMRI) scans in the absence of clinical information.

MDD, which is characterized by persistent depressed mood or anhedonia and cognitive deficits [APA, 2000], has been ranked as a leading cause of living on disability worldwide [Eisch and Petrik, 2012; Tanti and Belzung, 2010]. To date, the pathophysiology of MDD remains unclear, and the diagnosis of MDD is based largely on history, self-reported symptoms and clinical signs. Therefore, the investigation of objective neurobiological markers is significant for both diagnostic systems and treatment decisions.

The core depressive symptoms may be associated with brain network dysfunction [Drevets et al., 2008; Mayberg, 2003]; this suggested association has prompted a wealth of resting-state functional connectivity MRI (rs-fcMRI) studies. The subgenual anterior cingulate cortex (sACC), which is located beneath the genu of the corpus callosum and corresponds primarily to Brodmann area 25 (BA25) and the caudal portions of BA32 and BA24 [Johansen-Berg et al., 2008], has become a critical region of interest (ROI) within the distributed brain networks that mediate depressive symptoms [Greicius et al., 2007; Mayberg, 2003]. Many previous studies have implicated the sACC as a focus of dysfunction in MDD [Botteron et al., 2002; Coryell et al., 2005; Davey et al., 2012; Drevets et al., 2002; Greicius et al., 2007; Kennedy et al., 2001; Mayberg et al., 2000; Osuch et al., 2000]. All these findings imply that resting-state functional connections of the sACC can provide potential useful information to advance our current understanding of the pathophysiology of MDD and may be promising features in identifying depression.

The sACC is functionally and cytoarchitectonically distinct from the pregenual anterior cingulate cortex (pACC) [Johansen-Berg et al., 2008; Margulies et al., 2007; McCormick et al., 2006; Vogt et al., 1995; Walter et al., 2009], although the pACC may also play an important role in the pathophysiology of MDD [McCormick et al., 2006; Walter et al., 2009]. In fact, the two regions are activated during the presentation of different emotional and cognitive stimuli [Kelley et al., 2002; Kross et al., 2009; Walter et al., 2009], and this distinction is also reflected by hyperactivity in the sACC and hypoactivation in the pACC, to some extent, in depression [Walter et al., 2009]. However, functional parcellation of the perigenual cingulate comprising the sACC and pACC has not been investigated. Thus, the

TABLE 1. Characteristics of the participants in this study

Variable	Patient	Control	*P* value
Sample size	24	29	
Gender (M/F)	8/16	9/20	0.86[a]
Age (yr)	31.83 ± 10.99	33.62 ± 10.29	0.54[b]
Education (yr)	11.71 ± 3.13	11.00 ± 3.12	0.66[b]
Number of previous episodes	1.63 ± 0.77		
Duration of current episode (months)	5.33 ± 6.29		
Hamilton Depression Rating Scale (HDRS)	26.42 ± 5.22 (18–38)	4.25 ± 1.02 (3–6)	
Hamilton Anxiety Rating Scaling (HAMA)	20.29 ± 5.25 (8–30)	3.55 ± 0.91 (2–5)	
Clinical Global Impression Scale-Severity (CGI-S)	5.92 ± 0.65 (5–7)		

[a]Pearson χ^2 test.
[b]Two-sample *t*-test.

characterization of the variability of the perigenual cingulate functional connectivity in healthy subjects is a first step in evaluating patient variability. In this study, we used an MMC algorithm to perform a functional connectivity-based parcellation of the perigenual cingulate region and then used the functional connectivity of the subregions from the parcellation as classification features to distinguish the depressed patients from the healthy controls.

MATERIALS AND METHODS

Subjects

The study's participants included 24 patients diagnosed with MDD from the outpatient clinic at the First Affiliated Hospital of China Medical University and 29 healthy volunteers, matched for gender, age and education, who were recruited via local recruitment advertisements in the hospital and in the neighboring communities (Table 1). All of the subjects were right-handed and native Chinese speakers. The depressed patients met the criteria for a current episode of unipolar recurrent major depression based on the DSM-IV (Diagnostic and Statistical Manual of Mental Disorders-IV) criteria [APA, 2000]. Using the Structured Clinical Interview for DSM-IV [First et al., 1995], the diagnosis was confirmed by clinical psychiatrists. The exclusion criteria included acute physical illness, substance abuse or dependence, a history of a head injury resulting in loss of consciousness and a major psychiatric or neurological illness other than depression. The patients abstained from caffeine, nicotine and alcohol prior to the scanning session. All patients were medication-naive at the time of the scan, and the depressive symptoms were assessed with the 17-item Hamilton Depression Rating Scale (HDRS) [Hamilton, 1960], Hamilton Anxiety Rating Scale (HAMA) [Hamilton, 1959], and Clinical Global Impression Scale-Severity (CGI-S) [Guy, 1976] (Table 1). The healthy volunteer subjects were studied under identical conditions. This study was approved by the Ethics Committee of China Medical University, and all participants gave written informed consent.

Imaging Acquisitions and Data Preprocessing

All resting-state fMRI data were acquired using a 1.5-Tesla GE SIGNA scanner (GE Medical System, Milwaukee, WI). A total of 245 volumes of echo planar images were obtained axially (repetition time/echo time = 2,000/50 ms, thickness/gap = 5/1.5 mm, field of view = 240 × 240 mm, flip angle = 90°, matrix = 64 × 64, and slices = 20).

All of the fMRI images were preprocessed using SPM5 (http://www.fil.ion.ucl.ac.uk/spm). For each subject, the first five volumes of the scanned data were discarded for magnetic saturation. The remaining 240 volumes were corrected by registering and reslicing for head motion, and all subjects had less than 2.5 mm translation in the *x*, *y*, or *z*-axis and 2° of rotation in each axis. Next, the volumes were normalized to the standard template (Montreal Neurological Institute). The resulting images were detrended to eliminate linear trends and then temporally filtered with a band-pass filter (0.01–0.08 Hz). No spatial smoothing was performed to avoid the possible degradation of classification ability [Kriegeskorte et al., 2006]. Finally, the time courses were further corrected by regressing out head motion parameters, global signals, white matter and cerebrospinal fluid average signals, i.e., these nuisance variables were used as regressors in a general linear regression model to remove the associated variance [Desjardins et al., 2001; Fox et al., 2009]. The residuals of these regressions were the set of time series used for functional connectivity analyses. We evaluated the functional connectivity between each time series pair using Pearson correlation coefficients. In addition, all correlation coefficients were converted to z-scores by applying the Fisher *r*-to-*z* transformation [Zar, 1996].

Maximum Margin Clustering

Clustering analysis is one of the most popular unsupervised machine learning techniques [Jain et al., 1999; Madhulatha, 2012]. In this study, we used an unsupervised MMC algorithm to perform the functional connectivity-based parcellation of the perigenual cingulate to

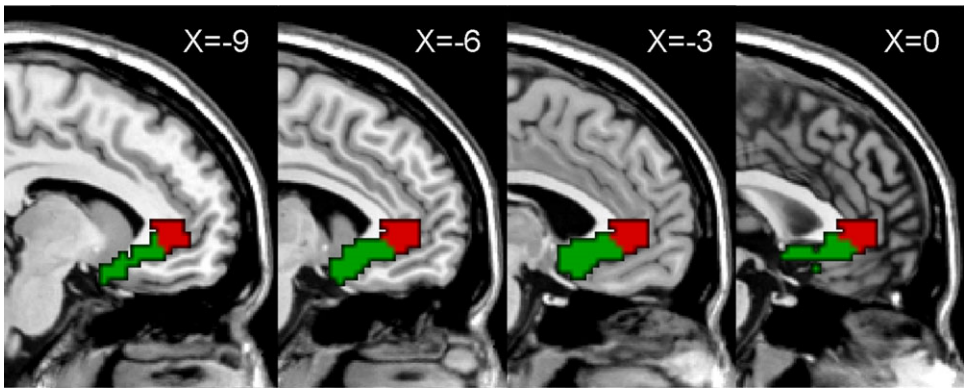

Figure 1.
Location of the perigenual cingulate and its functional connectivity-based parcellation defining two distinct subregions: the subgenual (sACC, green) and pregenual anterior cingulate cortex (pACC, red). [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

distinguish depressed patients from healthy controls on the basis of single resting-state fMRI scans and in the absence of clinical information. Inspired by the success of the large margin criterion in a support vector machine (SVM), Xu et al. [2005] proposed the use of the maximum margin principle for clustering, referred to as MMC, which simultaneously learns the optimal hyperplane and cluster labels and is very robust to various data distributions [Xu and Schuurmans, 2005; Xu et al., 2005]. The goal of an SVM is to find the linear discriminant that maximizes the minimum misclassification margin based on labeled training data, but the goal of MMC is to find labeling that results in a large margin classifier. However, the original mathematical model of two-class MMC is a mixed integer program [Li et al., 2009]. To make it more tractable and to avoid the problem of suffering from local minima, MMC can be solved globally via convex semi-definite programming (SDP) relaxation. Recently, Li et al. [2009] proposed a novel and scalable convex optimization method, LG-MMC, which maximized the margins of opposite clusters via "Label Generation," to alleviate the computational burden of SDP. Their experiments showed that LG-MMC achieves promising clustering performance. In this study, we used the LG-MMC package (linear kernel, LGMMC_V2; http://lamda.nju.edu.cn/datacode/LGMMC.htm) to solve the optimization problem.

Unsupervised Parcellation of the Perigenual Cingulate

Individual connectivity-based parcellation was performed using the LG-MMC algorithm, and then population-based parcellation was achieved using a standard majority-voting scheme based on the individual parcellation results. First, according to the previous studies [Johansen-Berg et al., 2008; Mayberg et al., 2005], we defined the ROI of the perigenual cingulate with the free software WFU_PickAtlas (version 2.0, http://www.ansir.wfubmc.edu) on the basis of its ACC region excluding the dorsal component, BA25 region, and the caudal portions of BA32 and BA24 regions, as shown in Figure 1. Second, resting-state functional connectivity was calculated between each voxel within the perigenual cingulate and all other voxels in the brain for each healthy subject, allowing a feature vector to be extracted from the functional connectivity map of each voxel within the perigenual cingulate. Taking each voxel within the perigenual cingulate as a sample, two-class clustering analysis was performed for each healthy subject. In the LG-MMC algorithm, the slack parameter was set as *Inf*, and the class balance parameter was set as approximately one fifth of the total voxel number in the perigenual cingulate. The unsupervised connectivity-based parcellation defined two subregions with distinct resting-state functional connectivity patterns within the perigenual cingulate for each healthy subject. To avoid the curses of data dimensionality in clustering analysis [Bellman, 1961; Guyon and Elisseeff, 2003], principal component analysis (PCA) was used to reduce the dimensionality of the feature space. Finally, to resolve disagreements in the labeling of the 29 healthy subjects coming from the same voxel within the perigenual cingulate, we applied a standard majority-voting scheme in which the label predicted most frequently among the 29 healthy subjects was assumed as the label for the given voxel. A population-based parcellation into two subregions was generated in standard brain space.

Unsupervised Classification of Major Depression

We developed an MMC-based unsupervised machine learning approach to cluster the depressed patients and the healthy controls into two groups based only on rs-fcMRI data. The feature space for classification was spanned by the functional connectivity between each

subregion in our parcellation of the perigenual cingulate and all other voxels. The LG-MMC package, described previously, was used. As in a previous study [Xu et al., 2005], the following two widely used algorithms were also used: C-means and Ncut [Shi and Malik, 2000]. In all algorithms, we set the number of clusters to 2 for classifying the depressed patients and the healthy controls. In addition, there were certain parameters that needed to be set in the LG-MMC algorithm: (1) the slack parameter was selected in a range of [0.001, 0.01, 0.1, 1, 10, 100, *Inf*]; and (2) the class balance parameter was set by grid search from [6, 15] with granularity 1. In addition, the dimensionality of principal component subspace was set by grid search from [2, 52] with granularity 1. All of the algorithms would be reported with the best parameter setting.

To evaluate the consistency between the clustering labels and clinical diagnostic labels of the subjects, we defined clustering consistency as similar to clustering accuracy [Wang et al., 2010], which can be used to discover one-to-one relationships between clusters and clinical classes and can measure the extent to which each cluster contains data points from the corresponding class. Clustering consistency sums up the entire matching degree between all pair class clusters. Clustering consistency can be computed as

$$\text{Consistency} = \frac{1}{N} \max_{C_k, L_m} \left(\sum_{C_k, L_m} T(C_k, L_m) \right) \tag{1}$$

where C_k denotes the k_{th} ($k \in \{1, 2\}$) cluster in the final results, and L_{m} is the diagnostic m_{th} ($m \in \{1, 2\}$) group (i.e., patient group and control group). $T(C_k, L_m)$ is the number of samples that belong to group m and are assigned to cluster k. Consistency is the maximum sum of $T(C_k, L_m)$ for all pairs of clusters and groups, and these pairs have no overlaps.

In addition, we used a leave-one-out cross-validation (LOOCV) strategy to estimate the generalization ability of our MMC-based unsupervised classification method at an individual subject level. To better understand the performance of the MMC-based classifiers, two other popular linear classification methods, SVM with linear kernel functions and linear discriminant analysis (LDA), were applied to the same data [Bishop, 2006; Vapnik, 1995]. The overall proportion of samples whose assigned labels were consistent with the diagnostic labels was evaluated by classification consistency, with a greater classification consistency resulting in better generalization ability.

Reconstruction of the Most Discriminative Features

A linear MMC-based classifier in nature is a linear projection, and the absolute value of an element in the projective vector represents the contribution of the corresponding feature to the discriminative score. Thus, these values can be used to measure the discriminative power of the functional connectivity features, called feature weights. Given a two-class data set $X^0 = \{X^0_i\} \in \mathrm{R}^{D \times N}$, where $X^0_i \in \mathrm{R}^D$ are samples (feature vectors in this study), it was assumed that the value of d ($d < D$) was given. Using PCA, all sample data were projected onto d-dimensional subspace spanned by the principle components, and the projective matrix $U \in \mathrm{R}^{D \times d}$ was obtained:

$$X_i = U^{\mathrm{T}}\left(X^0_i - \bar{X}^0_i\right) \tag{2}$$

where $\overline{X}^0_i$ was the mean vector of X^0_i and $X_i \in \mathrm{R}^d$ was the output vector of PCA. Usually, U consists of d eigenvectors with the first d largest eigenvalues of the covariance matrix.

In brief, an MMC algorithm with linear kernel function seeks an optimal projective direction by maximizing the margin. Theoretically, the d-dimensional feature vector X of each sample is projected to a one-dimensional discriminative score y by the inner product operation:

$$y = \langle X, W \rangle - y_0 = W^{\mathrm{T}}X - y_0 \tag{3}$$

where $W \in \mathrm{R}^d$ is the optimal projective vector and $y_0 \in \mathrm{R}$ is a classification threshold.

In fact, we can combine Eq. (2) with Eq. (3) in the clustering procedure, and then an equivalent result is obtained:

$$y_i = W^{\mathrm{T}}\left(U^{\mathrm{T}}(X^0_i - \bar{X}^0_i)\right) - y_0 = (UW)^{\mathrm{T}}\ (X^0_i - \bar{X}^0_i) - y_0 \tag{4}$$

Suppose $\Lambda = UW = (\lambda_1, \ldots, \lambda_D)^{\mathrm{T}}$; then, the absolute value of the element λ_j ($j = 1, \ldots, D$) represents the contribution of the corresponding feature to the discriminative score, which denotes the feature weight. Thus, a greater feature weight means a greater contribution of this feature to the discriminative score, and the features with the greatest weights can be considered to be the most discriminative features. Because the feature weights might differ slightly from fold to fold of LOOCV, the final feature weights were averages from all folds of LOOCV. The top 5% of features with the greatest weights were extracted as the most discriminative features. Finally, a cluster size threshold of 10 voxels was used to extract the anatomical regions with the highest discriminative power.

RESULTS

Functional Connectivity-Based Parcellation of the Perigenual Cingulate

A population-based parcellation of the perigenual cingulate into two subregions was generated in the standard brain space (Fig. 1). Interestingly, the two clusters

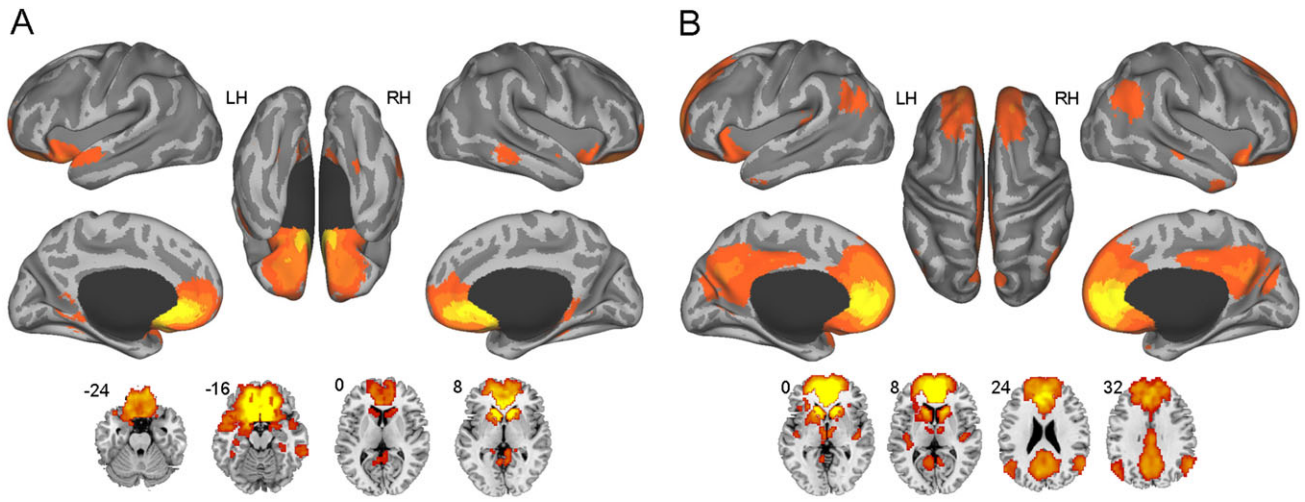

Figure 2.

Resting-state functional connectivity maps of the subgenual (**A**) and pregenual (**B**) anterior cingulate cortex (one-sample *t*-test, False Discovery Rate corrected, $P < 0.05$, cluster size >20 voxels). Relative to the subgenual cingulate, the pregenual cingulate is additionally connected with certain default mode regions, mainly including the precuneus/posterior cingulate cortex and bilateral angular gyri. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

represented the two typical portions of the perigenual cingulate, including the sACC (Fig. 1, green) and pACC (Fig. 1, red). We selected the sACC and pACC regions from the parcellation as seeds to calculate the functional connectivity with all the other voxels in the brain for each subject, resulting in resting-state functional connectivity maps of the two regions as shown in Figure 2 (one-sample *t*-test, False Discovery Rate corrected, $P < 0.05$, cluster size >20 voxels). It was observed that the sACC was most strongly connected with the medial prefrontal cortex (mPFC, BA11/10/24/32/25), basal ganglia, temporal lobe (BA38) and cerebellum (Supporting Information Table S1). By contrast, and in addition to these areas, the pACC is also strongly connected with the precuneus/posterior cingulate cortex (BA7/31/23), bilateral angular gyri (BA39/40), and thalamus (Supporting Information Table S2), all of which are key default mode regions. There are some common areas between the two functional connectivity maps, but there are also apparent differences between the two subregions that may drive the parcellation.

Classification Results

Using functional connectivity maps of the sACC, LG-MMC achieved better performance (92.5% clustering consistency) than Ncut (83.0%) or C-means (100 runs: 82.1 ± 5.1%) (Table 2). We also clustered the depressed patients and controls based on the functional connectivity maps of the pACC, resulting in the clustering consistencies of 84.9%, 81.1%, and 72.9 ± 3.8% using MMC, Ncut, and C-means, respectively.

Considering the promising clustering performance of MMC, we tested its generalization ability using an LOOCV strategy. The MMC-based classifiers achieved an individual-level classification consistency of 92.5% using the functional connectivity maps of the sACC, which was similar to the results of SVM classifiers but better than the results of LDA classifiers (Table 3). Using the functional connectivity maps of the pACC, the MMC classifiers via LOOCV achieved a classification consistency of 83.0%, whereas the SVM and LDA classifiers both achieved a classification consistency of 81.1%.

Most Discriminative Functional Connectivity Networks in Major Depression

The most discriminative sACC-based functional connectivity network was extracted, as shown in Figure 3A.

TABLE 2. Comparison of clustering performance using different algorithms and different seed regions

		Clustering consistency		
Seed	Algorithm	Patient (%)	Control (%)	Total (%)
Subgenual cingulate	MMC	100	86.2	92.5
	Ncut	70.8	93.1	83.0
	C-means	84.6 ± 7.0	80.0 ± 7.1	82.1 ± 5.1
Pregenual cingulate	MMC	87.5	82.8	84.9
	Ncut	66.7	93.1	81.1
	C-means	62.0 ± 7.3	82.0 ± 3.3	72.9 ± 3.8

MMC, Maximum Margin Clustering; Ncut, Normalized Cuts Spectral Clustering.

TABLE 3. Classification results in leave-one-out cross-validation using the functional connectivity maps of the subgenual and pregenual anterior cingulate cortex

Seed	Algorithm	Classification consistency		
		Patient (%)	Control (%)	Total (%)
Subgenual cingulate	MMC	100	86.2	92.5
	SVM	95.8	89.7	92.5
	LDA	95.8	82.8	88.7
Pregenual cingulate	MMC	83.3	82.8	83.0
	SVM	87.5	75.9	81.1
	LDA	66.7	93.1	81.1

LDA, Linear Discriminant Analysis; MMC, Maximum Margin Clustering; SVM, Support Vector Machine.

These connections were primarily located between the sACC and certain cortical regions and limbic structures (Table 4). The cortical regions primarily included the ventrolateral prefrontal cortex (vlPFC, BA47/11), superior temporal gyri (BA21/22), and vmPFC (BA10/32/24), while the limbic areas mainly comprised the bilateral lentiform nucleus, right hippocampus, bilateral insula, bilateral thalamus, brainstem and right amygdala.

The most discriminative functional connections of the pACC were primarily located between the pACC and the central default mode regions, i.e., the medial prefrontal cortex (mPFC, BA10/11/32), precuneus/posterior cingulate cortex (BA7/23/31), orbitofrontal cortex (BA11),

TABLE 4. Regions with the most discriminative functional connectivity of the subgenual anterior cingulate cortex in major depression

Regions[a]	Side	BA	Voxels	MNI coordinates (mm)		
				x	*y*	*z*
vlPFC	L	47	685	−32	31	−15
	R	47/11	252	32	30	−14
Lentiform nucleus	L		247	−27	3	−7
	R		255	24	6	−7
Hippocampus	R		166	24	−42	−6
Superior temporal gyrus	L	22	152	−49	−14	−6
	R	22/21	101	52	−9	−9
Insula	L	13	122	−35	11	−10
	R	13	88	33	22	−11
vmPFC/ACC	R	10/32/24	107	6	64	11
	L	10/32/24	68	−13	64	11
Thalamus	R		89	15	−25	−3
	L		74	−15	−8	−3
Superior frontal gyrus	R	10	39	22	60	14
	L	10	20	−20	57	5
Brainstem	R		32	14	−24	4
	L		31	−5	−23	−2
Calcarine gyrus	R	19	24	25	−51	4
Amygdala	R		16	24	−8	−14

[a]These regions were identified by setting the threshold to the top 5% most discriminative functional connectivity with cluster size >10 voxels. ACC, anterior cingulate cortex; BA, Brodmann area; MNI, Montreal Neurological Institute; vlPFC, ventrolateral prefrontal cortex; vmPFC, ventromedial prefrontal cortex; L, left; R, right.

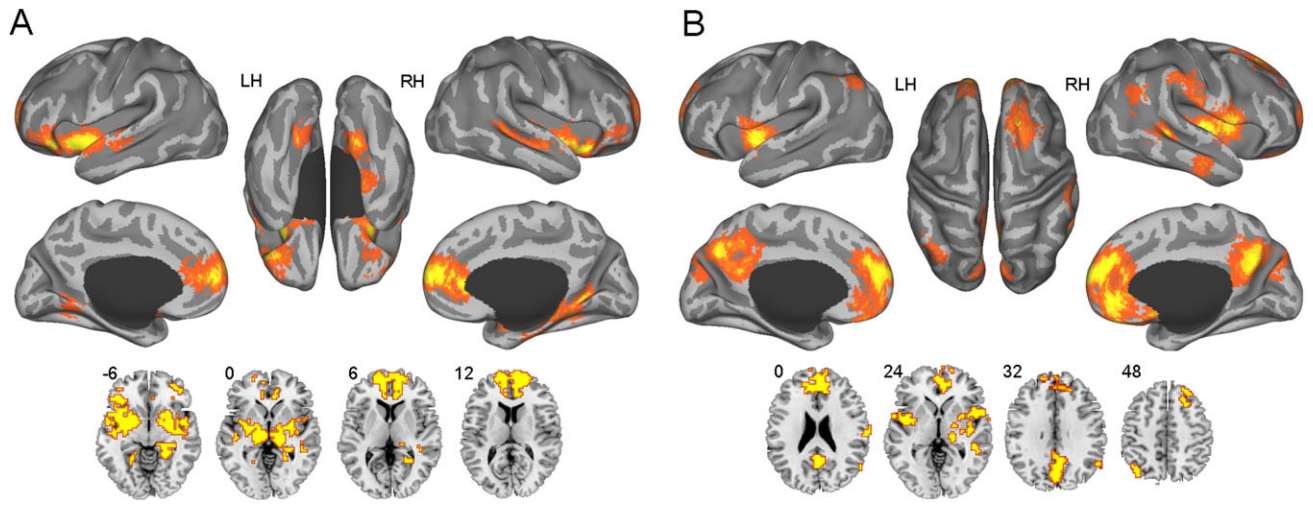

Figure 3.

The most discriminative functional connectivity networks of the subgenual (**A**) and pregenual (**B**) anterior cingulate cortex. The voxels with the most discriminative subgenual cingulate functional connectivity are located primarily within the prefrontal lobe, some limbic areas and the temporal lobe, while the voxels with the most discriminative pregenual cingulate functional connectivity are located primarily within the medial prefrontal cortex, precuneus/posterior cingulate cortex, insula, bilateral angular gyri and temporal lobe. [Color figure can be viewed in the online issue, which is available at wileyonlinelibrary.com.]

TABLE 5. Regions with the most discriminative functional connectivity of the pregenual anterior cingulate cortex in major depression

Regions[a]	Side	BA	Voxels	MNI coordinates (mm)		
				x	*y*	*z*
mPFC/ACC	L/R	10/11/32	697	−3	57	12
Precuneus/posterior cingulate	L/R	7/31/23	347	1	−70	39
Superior frontal gyrus	R	8/9	155	24	36	53
Insula	R	13	138	41	−3	0
	L	13	57	−45	0	0
Orbitofrontal cortex	L	11	75	−14	42	−24
	R	11	44	11	29	−21
Rolandic oper gyrus	R	44	71	49	6	2
Middle temporal gyrus	R	21	45	45	−34	1
Superior temporal gyrus	L	22	35	−60	−6	2
Thalamus	R		28	18	−21	0
Angular gyrus	L	40	27	−39	−69	48
	R	39	22	57	−61	25
Infeiror temporal gyrus	R	21	27	60	−12	−29
supramarginal gyrus	R	1	25	66	−17	21
Putamen	R		23	32	4	0

[a]These regions were identified by setting the threshold to the top 5% most discriminative functional connectivity with cluster size >10 voxels. ACC, anterior cingulate cortex; BA, Brodmann area; MNI, Montreal Neurological Institute; mPFC, medial prefrontal cortex; L, left; R, right.

bilateral angular gyri (BA39/40), and thalamus (see Table 5 and Fig. 3B). The most discriminative functional connections also included those between the pACC and the right superior frontal gyrus (BA8/9), bilateral insula (BA13), and temporal lobe (BA21/22) (Table 5 and Fig. 3B).

DISCUSSION

This study demonstrates that the subgenual and pregenual portions of the ACC have distinct functional connectivity patterns and that the sACC functional connectivity could achieve a promising clustering and classification consistency of 92.5% in identifying major depression using the unsupervised MMC algorithm, whereas the pACC functional connectivity reached only relatively low consistencies. Moreover, the most discriminative functional connectivity network related to the sACC primarily included the prefrontal lobe, limbic areas and temporal lobe, while the most discriminative pACC-related network mainly encompassed the insula and some default mode regions.

Functional Connectivity-Based Parcellation of the Perigenual Cingulate

In the current study, major differences between the resting-state functional connectivity patterns of the pregenual and subgenual portions of human ACC were observed. Although there was some overlap between the functional connectivity patterns of the two subregions, the pACC was also strongly connected with the precuneus/posterior cingulate cortex and bilateral angular gyri. It is likely that these apparent differences drove the functional connectivity-based parcellation. Note that our parcellation of the perigenual cingulate is similar to the anatomical connectivity-based parcellation in Johansen-Berg et al. [2008] and the structural MRI-based parcellation in McCormick et al. [2006], although we performed our parcellation using functional connectivity. In addition, we used a C-means algorithm to perform the parcellation [Zhang and Li, 2012], which yielded results similar to those of MMC. However, the group consistency [Shen et al., 2010b] of the MMC results (group consistency = 0.3392) was better than that of the C-means results (group consistency = 0.3861), indicating that MMC achieved better agreement between parcellations across all subjects (a smaller group consistency value indicates better agreement between parcellations across all subjects) [Shen et al., 2010b].

In addition, this study demonstrated that the sACC exhibited higher discriminative power than the pACC in identifying major depression, indicating that the parcellation benefited the identification of major depression. Briefly, our results further confirm that there are two distinct subregions within the perigenual cingulate and suggest, in addition to the distinct anatomical connectivity maps [Johansen-Berg et al., 2008], that the two subregions may also have distinct resting-state functional connectivity patterns [Margulies et al., 2007]. By characterizing the variability of functional connectivity of the perigenual cingulate, the current study may have important clinical implications for studies of functional connectivity network variability in depressed patients.

Unsupervised Classification of Major Depression

Using the functional connectivity maps of the sACC, the MMC-based unsupervised machine learning approach achieved group-level clustering consistency of 92.5%. The current results demonstrate that this unsupervised approach reliably captured the intrinsic disease-related functional connectivity pattern underlying the rs-fcMRI data. In our previous study, we used supervised SVM to identify major depression based on whole-brain functional connectivity [Zeng et al., 2012], and several other brain imaging studies have also attempted to distinguish depressed patients from healthy controls using supervised methods [Costafreda et al., 2009; Craddock et al., 2009; Fang et al., 2012; Fu et al., 2008; Mwangi et al., 2012]. Mourão-Miranda et al. [2011] used one-class SVM to

investigate whether patterns of fMRI responses to sad facial expressions in depressed patients would be classified as outliers in relation to patterns of healthy controls. Supervised classification requires the prior diagnosis of all subjects based on clinical signs and symptoms; thus, these methods may be prone to the subjects' behaviors and self-reported symptoms, as well as clinical psychiatrists' bias [Buckner et al., 2007; Caplan and Cosgrove, 2004; Sue, 1998]. In contrast, unsupervised classification methods can perform classification without any prior labeling knowledge; thus, these methods can avoid the above biases [Bishop, 2006], and provide relatively objective biomarkers. To the best of our knowledge, the current study was the first time unsupervised machine learning methods have been used to identify psychiatric disorders based on neuroimaging data.

In the current study, MMC outperformed both the C-means and Ncut algorithms to some extent, which is consistent with previous studies [Ding et al., 2009]. In C-means algorithms, it is supposed that the samples of each class exhibit a hyperspherical or Gaussian distribution in the high-dimension feature space and that these samples are located symmetrically around the center of the hypersphere. However, due to the diversity of psychiatric patients, the neuroimaging data of the subjects are not likely to be hyperspherical or symmetrical in feature space. The Ncut algorithm is robust to outliers and exhibits favorable performance relative to other graph clustering methods [Shen et al., 2010b; Shi and Malik, 2000], but it does not always lead to particularly good solutions [Luxburg, 2006]. The MMC algorithm, as a scalable and global optimization method, learns the optimal hyperplane with the maximum margin between two clusters, ignoring the internal data distribution structure of each cluster. Thus, the MMC algorithm is very robust to the data distribution. In this study, it was conceivable that our MMC-based unsupervised approach achieved higher clustering consistencies than both C-means and Ncut.

We used an LOOCV strategy to estimate the generalization ability of the unsupervised MMC method and obtained an individual-level classification consistency of 92.5% based on the functional connectivity maps of the sACC. The results are equivalent to, or better than, the results of supervised SVM and LDA classifiers, indicating that the MMC-based unsupervised machine learning approach has promising generalization ability and the potential to inform clinical practice and research, as it can make accurate predictions about brain scans at an individual subject level.

To highlight the promising contribution of the sACC to the classification, we also tested other seed regions, including the posterior cingulate cortex within the default network, the anterior lateral prefrontal cortex within the frontoparietal network, and the motor cortex within the motor network [Van Dijk et al., 2012], obtaining classification consistencies of 73.6%, 62.3%, and 60.4%, respectively. Obviously, these classification results are not as good as the results obtained by using the sACC as the seed and are a further indication of the vital role of the sACC in the pathophysiology of major depression.

Spatial smoothing may degrade classification performance [Kriegeskorte et al., 2006]. We performed smoothing with 2-, 4-, 6-, and 8-mm FWHM widths, obtaining classification consistencies of 90.6%, 90.6%, 88.7%, and 84.9%, respectively. Spatial smoothing may blur the subtle information hidden in the fMRI data and reduce the discriminative power of the features.

Most Discriminative Functional Connectivity Networks in Major Depression

The most discriminative functional connections of the sACC were located primarily between the sACC and the prefrontal lobe, limbic areas and temporal lobe. The prefrontal lobe referred to in this study included mainly the vlPFC and vmPFC. Abnormal functional coupling of the sACC and some parts of the prefrontal lobe has been found in numerous previous studies [Greicius et al., 2007; Mayberg et al., 1999] and may be responsible for some aspects of emotional dysregulation in depression. For example, using positron emission tomography (PET), Mayberg et al. [1999] reported a significantly altered correlation between the sACC and the right dorsal prefrontal cortex in depressed patients. Greicius et al. [2007] demonstrated increased connectivity between the sACC and the default network, including the medial prefrontal and orbitofrontal cortex. Our findings share certain features with these results. The limbic areas, including the striatum, hippocampus, insula, thalamus, brainstem and amygdala, are fundamentally involved in critical aspects of motivational, affective and emotional behaviors [LeDoux, 2000]. As pivotal nodes in the limbic-cortical dysregulation model in depression [Drevets et al., 2008; Mayberg, 2003], the limbic structures may mediate the vegetative and somatic aspects of the disorder, including sleep, appetite, libido and endocrine disturbances [Mayberg, 1997; Sheline et al., 2010]. Here, altered functional connectivity of the sACC with these regions may account for at least a portion of the complex of cognitive and emotional deficits in depression. In addition, the functional connectivity between the sACC and superior temporal gyri exhibited high discriminative power in classification. The superior temporal gyri are involved in the perception of emotions in facial stimuli [Radua et al., 2010]; thus, altered functional connectivity between the sACC and this region may result in the social avoidance observed in depressed patients through the reinforcement of negative cognitive models. We speculate that this region should also be included in the limbic-cortical dysregulation model in depression.

The most discriminative functional connections related to the pACC were found primarily between the pACC and the default mode regions, i.e., the mPFC, precuneus/posterior cingulate cortex, orbitofrontal cortex, bilateral

angular gyri, and thalamus. In addition, the functional connections between the pACC and the right superior frontal gyrus, bilateral insula, and temporal lobe also exhibited high discriminative power. In a previous study, abnormalities in the functional connectivity between the pACC and the bilateral dorsomedial prefrontal cortex were found in major depression using univariate statistical analyses [Sheline et al., 2010]. In the current study, using multivariate pattern analyses, we extended the brain areas with aberrant coupling of the pACC to include the key regions in the default network, similar to the pACC-related limbic-cortical dysregulation model in depression [Mayberg, 2003]. The default network is known to be involved in self-referential activity [Greicius et al., 2003; Raichle, 2001], and abnormality of the default network in depression has been reported in several previous studies [Greicius et al., 2007; Sheline et al., 2009; Zhu et al., 2012]. Our results may provide new evidence for the importance of the default network in the pathophysiology of major depression and suggest that abnormal functional connectivity between the pACC and the default network may be a major depressive trait. Discriminative functional connections were also found between the pACC and bilateral insula, in line with the previous studies [Horn et al., 2010]. Altered functional coupling of the pACC and the insula may play an important role in major depressive symptoms, such as anhedonia and impaired emotion processing [Horn et al., 2010].

Limitations and Future Directions

There are several limitations to this study. First, our classification results of depressed patients and healthy controls may be negatively influenced by the small sample size, scanner variability and the lack of large independent datasets. Therefore, it is important to confirm the classification results with larger sample size and multicenter imaging data in the future. Second, due to the limitations of the small sample size, we performed the characterization of perigenual cingulate connectivity patterns and the classification of depression using the same group of healthy controls. The use of the same healthy control participants in both arms of the experiment is a potential confound that may influence the classification results, although LOOCV was used to address this possibility. The impact of this issue on classification performance remains to be determined. Further studies with independent datasets to characterize the perigenual cingulate connectivity patterns and classify depressed patients will allow us to avoid this potential confound. Third, although our unsupervised classification results are promising, additional brain imaging evidence, including structural abnormalities, is needed to develop synthesized biomarkers for a more reliable clinical diagnosis of depression. In addition, although linear clustering algorithms were used here to allow the reconstruction of the most discriminative features, some other advanced unsupervised machine learning techniques, including low-dimensional manifold embedding and nonlinear kernel methods, may have the potential to improve the classification performance [Koutsouleris et al., 2009; Shen et al., 2010a].

ACKNOWLEDGMENTS

The authors thank the volunteers and patients for their participation in the study and the three anonymous referees for their insightful comments and suggestions.

REFERENCES

APA (2000): Diagnostic and Statistical Manual of Mental Disorders, 4th ed. Washington, DC: American Psychiatric Press.

Bellman RE (1961): Adaptive Control Processes: A Guided Tour. Princeton, New Jersey: Princeton University Press.

Bishop CM (2006): Pattern Recognition and Machine Learning. New York: Springer.

Botteron KN, Raichle ME, Drevets WC, Heath AC, Todd RD (2002): Volumetric reduction in left subgenual prefrontal cortex in early onset depression. Biol Psychiatry 51:342–344.

Buckner JD, Castro Y, Holm-Denoma JM, Joiner TE, editors (2007): Mental Health Care for People of Diverse Backgrounds. Abingdon, Oxon: Radcliffe Publishing Ltd.

Caplan PJ, Cosgrove L, editors (2004): Bias in Psychiatric Diagnosis. Oxford: Jason Aaronson.

Coryell W, Nopoulos P, Drevets WC, Wilson T, Andreasen NC (2005): Subgenual prefrontal cortex volumes in major depressive disorder and schizophrenia: Diagnostic specificity and prognostic implications. Am J Psychiatry 162:1706–1712.

Costafreda SG, Chu C, Ashburner J, Fu CHY (2009): Prognostic and diagnostic potential of the structural neuroanatomy of depression. PLoS One 4:e6353.

Craddock RC, Holtzheimer III PE, Hu XP, Mayberg HS (2009): Disease state prediction from resting state functional connectivity. Magn Reson Med 62:1619–1628.

Davey CG, Harrison BJ, Yücel M, Allen NB (2012): Regionally specific alterations in functional connectivity of the anterior cingulate cortex in major depressive disorder. Psychol Med 42(10): 2071–2081.

Desjardins AE, Kiehl KA, Liddle PF (2001): Removal of confounding effects of global signal in functional MRI analyses. Neuroimage 13:751–758.

Ding SL, Van Hoesen GW, Cassell MD, Poremba A (2009): Parcellation of human temporal polar cortex: A combined analysis of multiple cytoarchitectonic, chemoarchitectonic, and pathological markers. J Comp Neurol 18:595–623.

Drevets WC, Bogers W, Raichle ME (2002): Functional anatomical correlates of antidepressant drug treatment assessed using PET measures of regional glucose metabolism. Eur Neuropsychopharmacol 12:527–544.

Drevets WC, Price JL, Furey ML (2008): Brain structural and functional abnormalities in mood disorders: implications for neurocircuitry models of depression. Brain Struct Funct 213:93–118.

Eisch AJ, Petrik D (2012): Depression and hippocampal neurogenesis: A road to remission? Science 338:72–75.

Fang P, Zeng L-L, Shen H, Wang L, Li B, Liu L, Hu D (2012): Increased cortical-limbic anatomical network connectivity in

major depression revealed by diffusion tensor imaging. PLoS One 7:e45972.

First MB, Spitzer RL, Gibbon M (1995): Structured Clinical Interview for DSM-IV Axis 1 Disorder-Patient Edition(SCID-I/P). New York: New York State Psychiatric Institute.

Fox MD, Zhang DY, Snyder AZ, Raichle ME (2009): The global signal and observed anticorrelated resting state brain networks. J Neurophysiol 101:3270–3283.

Fu CHY, Mourao-Miranda J, Costafreda SG, Khanna A, Marquand AF, Williams SCR, Brammer MJ (2008): Pattern classification of sad facial processing:Toward the development of neurobiological markers in depression. Biol Psychiatry 63:656–662.

Greicius MD, Flores BH, Menon V, Glover GH, Solvason HB, Kenna H, Reiss AL, Schatzberg AF (2007): Resting-state functional connectivity in major depression: Abnormally increased contributions from subgenual cingulate cortex and thalamus. Biol Psychiatry 62:429–437.

Greicius MD, Krasnow B, Reiss AL, Menon V (2003): Functional connectivity in the resting brain: a network analysis of the default mode hypothesis. Proc Natl Acad Sci USA 100:253–258.

Guy W (1976): Clinical Global Impressions: In ECDEU Assessment Manual for Psychopharmacology. Revised DHEW Pub. (ADM). Rockville, MD: National Institute for Mental Health.

Guyon I, Elisseeff A (2003): An introduction to variable and feature selection. J Mach Learn Res 3:1157–1182.

Hamilton M (1959): The assessment of anxiety states by rating. Br J Med Psychol 32:50–55.

Hamilton M (1960): A rating scale for depression. J Neurol Neurosurg Psychiatry 23:56–62.

Hardoon DR, Mourão-Mirand J, Brammer M, Shawe-Taylor J (2007): Unsupervised analysis of fMRI data using kernel canonical correlation. Neuroimage 37:1250–1259.

Horn DI, Yu C, Steiner J, Buchmann J, Kaufmann J, Osoba A, Eckert U, Zierhut KC, Schiltz K, He H, Biswal B, Bogerts B, Walter M. (2010): Glutamatergic and resting-state functional connectivity correlates of severity in major depression—The role of pregenual anterior cingulate cortex and anterior insula. Front Syst Neurosci 4:1–10.

Jain AK, Murty MN, Flynn PJ (1999): Data clustering: A review. ACM Comput Survey 31:264–323.

Johansen-Berg H, Gutman DA, Behrens TEJ, Matthews PM, Rushworth MFS, Katz E, Lozano AM, Mayberg HS (2008): Anatomical connectivity of the subgenual cingulate region targeted with deep brain stimulation for treatment-resistant depression. Cereb Cortex 18:1374–1383.

Kawasaki Y, Suzuki M, Kherif F, Takahashi T, Zhou S-Y, Nakamura K, Matsui M, Sumiyoshi T, Seto H, Kurachi M (2007): Multivariate voxel-based morphometry successfully differentiates schizophrenia patients from healthy controls. Neuroimage 34:235–242.

Kelley WM, Macrae CN, Wyland CL, Caglar S, Inati S, Heatherton TF (2002): Finding the self? An event-related fMRI study. J Cogn Neurosci 14:785–794.

Kennedy SH, Evans KR, Krüger S, Mayberg HS, Meyer JH, McCann S, Arifuzzman AI, Houle S, Vaccarino FJ (2001): Changes in regional brain glucose metabolism measured with positron emission tomography after paroxetine treatment of major depressi. Am J Psychiatry 158:899–905.

Koutsouleris N, Meisenzahl EM, Davatzikos C, Bottlender R, Frodl T, Scheuerecker J, Schmitt G, Zetzsche T, Decker P, Reiser M, Möller H-J, Gaser C. (2009): Use of neuroanatomical pattern classification to identify subjects in at-risk mental states of psychosis and predict disease transition. Arch Gen Psychiatry 66:700–712.

Kriegeskorte N, Goebel R, Bandettini P (2006): Information-based functional brain mapping. Proc Natl Acad Sci USA 103:3863–3868.

Kross E, Davidson M, Weber J, Ochsner K (2009): Coping with emotions past: The neural bases of regulating affect associated with negative autobiographical memories. Biol Psychiatry 65:361–366.

LeDoux JE (2000): Emotion circuits in the brain. Annu Rev Neurosci 23:155–184.

Li Y-F, Tsang IW, Kwok JT, Zhou Z-H (2009): Tighter and convex maximum margin clustering. AISTATS'09. Clearwater Beach, FL: MIT Press, p W&CP 5.

Luxburg UV (2007): A tutorial on spectral clustering. Stat Comput 17:395–416.

Madhulatha TS (2012): An overview on clustering methods. IOSR J Eng 2:719–725.

Margulies DS, Kelly AMC, Uddin LQ, Biswal BB, Castellanos FX, Milham MP (2007): Mapping the functional connectivity of anterior cingulate cortex. Neuroimage 37:579–588.

Mayberg HS (1997): Limbic-cortical dysregulation: A proposed model of depression. J Neuropsychiatry Clin Neurosci 9(3):471–481.

Mayberg HS (2003): Modulating dysfunctional limbic-cortical circuits in depression: towards development of brain-based algorithms for diagnosis and optimised treatment. Br Med Bull 65:193–207.

Mayberg HS, Brannan SK, Tekell JL, Silva JA, Mahurin RK, McGinnis S, Jerabek PA (2000): Regional metabolic effects of fluoxetine in major depression: Serial changes and relationship to clinical response. Biol Psychiatry 48:830–843.

Mayberg HS, Liotti M, Brannan SK, McGinnis S, Mahurin RK, Jerabek PA, Silva JA, Janet L. Tekell, Martin CC, Lancaster JL, Fox PT. (1999): Reciprocal limbic-cortical function and negative mood: Converging PET findings in depression and normal sadness. Am J Psychiatry 156:675–682.

Mayberg HS, Lozano AM, V V, McNeely HE, D S, C H, Schwalb JM, Kennedy SH (2005): Deep brain stimulation for treatment-resistant depression. Neuron 45:651–660.

McCormick LM, Ziebell S, Nopoulos P, Cassell M, Andreasen NC, Brumm M (2006): Anterior cingulate cortex: An MRI-based parcellation method. Neuroimage 32:1167–1175.

Mourão-Miranda J, Hardoon DR, Hahn T, Marquand AF, Williams SCR, Shawe-Taylor J, Brammer M (2011): Patient classification as an outlier detection problem: An application of the One-Class Support Vector Machine. Neuroimage 58:793–804.

Mwangi B, Ebmeier KP, Matthews K, Steele JD (2012): Multi-centre diagnostic classification of individual structural neuroimaging scans from patients with major depressive disorder. Brain 135:1508–1521.

Oquendo MA, Baca-Garcia E, Artés-Rodríguez A, Perez-Cruz F, Galfalvy HC, Blasco-Fontecilla H, Madigan D, Duan N (2012): Machine learning and data mining: strategies for hypothesis generation. Mol Psychiatry 17:956–959.

Orrù G, Pettersson-Yeo W, Marquand AF, Sartori G, Mechelli A (2012): Using Support Vector Machine to identify imaging biomarkers of neurological and psychiatric disease: A critical review. Neurosci Biobehav Rev 36:1140–1152.

Osuch EA, Ketter TA, Kimbrell TA, George MS, Benson BE, Willis MW, Herscovitch P, Post RM (2000): Regional cerebral

metabolism associated with anxiety symptoms in affective disorder patients. Biol Psychiatry 48:1020–1023.

Pereira F, Mitchell T, Botvinick M (2009): Machine learning classifiers and fMRI: A tutorial overview. NeuroImage 45:S199–S209.

Radua J, Phillips ML, Russell T, Lawrence N, Marshall N, Kalidindi S, El-Hage W, McDonald C (2010): Neural response to specific components of fearful faces in healthy and schizophrenic adults. Neuroimage 49:936–946.

Raichle ME (2001): A default mode of brain function. Proc Natl Acad Sci USA 98:676–682.

Sheline YI, Barch DM, Price JL, Rundle MM, Vaishnavi SN, Snyder AZ, Mintun MA, Wang S, Coalson RS, Raichle ME (2009): The default mode network and self-referential processes in depression. Proc Natl Acad Sci USA 106:1942–1947.

Sheline YI, Price JL, Yan ZZ, Mintun MA (2010): Resting-state functional MRI in depression unmasks increased connectivity between networks via the dorsal nexus. Proc Natl Acad Sci USA 107:11020–11025.

Shen H, Wang L, Liu Y, Hu D (2010a): Discriminative analysis of resting-state functional connectivity patterns of schizophrenia using low dimensional embedding of fMRI. Neuroimage 49:3110–3121.

Shen X, Papademetris X, Constable RT (2010b): Graph-theory based parcellation of functional subunits in the brain from restingstate fMRI data. Neuroimage 50:1027–1035.

Shi J, Malik J (2000): Normalized cuts and image segmentation. IEEE Trans Pattern Anal Mach Intell 22:888–905.

Sue S (1998): In search of cultural competence in psychotherapy and counseling. Am Psychol 53:440–448.

Tanti A, Belzung C (2010): Open questions in current models of antidepressant action. Br J Pharmacol 159:1187–1200.

Uddin LQ, Menon V, Young CB, Ryali S, Chen T, Khouzam A, Minshew NJ, Hardan AY (2012): Multivariate searchlight classification of structural magnetic resonance imaging in children and adolescents with autism. Biol Psychiatry 70:833–841.

Van Dijk KRA, Sabuncu MR, Buckner RL (2012): The influence of head motion on intrinsic functional connectivity MRI. Neuroimage 59:431–438.

Vapnik V (1995): The Natures of Statistical Learning Theory. New York: Springer-Verlag.

Vogt BA, Nimchinsky EA, Vogt LJ, Hof PR (1995): Human cingulate cortex: Surface features, flat maps, and cytoarchitecture. J Comp Neurol 359:490–506.

Walter M, Henning A, Grimm S, Schulte RF, Beck J, Dydak, U, Schnepf B, Boeker H, Boesiger P, Northoff G (2009): The relationship between aberrant neuronal activation in the pregenual anterior cingulate, altered glutamatergic metabolism, and anhedonia in major depression. Arch Gen Psychiatry 66:478–486.

Wang F, Zhao B, Zhang C (2010): Linear time maximum margin clustering. IEEE Trans Neural Netw 21:319–332.

Xu L, Neufeld J, Larson B, Schuurmans D (2005): Maximum Margin Clustering. Advances in Neural Information Processing Systems. Cambridge, MA. p 1537–1544.

Xu L, Schuurmans D (2005): Unsupervised and semi-supervised multi-class support vector machines. In: The 20th National Conference on Artificial Intelligence. Pittsburgh, PA: AAAI Press. p 904–910.

Zar JH (1996): Biostatistical Analysis. Upper Saddle River, New Jersey: Prentice Hall.

Zeng L-L, Shen H, Liu L, Wang L, Li B, Fang P, Zhou Z, Li Y, Hu D (2012): Identifying major depression using whole-brain functional connectivity: A multivariate pattern analysis. Brain 135:1498–1507.

Zhang S, Li C-SR (2012): Functional connectivity mapping of the human precuneus by resting state fMRI. Neuroimage 59:3548–3562.

Zhu X, Wang X, Xiao J, Liao J, Zhong M, Wang W, Yao S (2012): Evidence of a dissociation pattern in resting-state default mode network connectivity in first-episode, treatment-naive major depression patients. Biol Psychiatry:611–617.