Support Vector Machine Classification of Obsessive-Compulsive Disorder Based on Whole-Brain Volumetry and Diffusion Tensor Imaging

Cong Zhou 1,2†, Yuqi Cheng 1†, Liangliang Ping 1,2, Jian Xu 3, Zonglin Shen 1, Linling Jiang 1, Li Shi 3, Shuran Yang 1, Yi Lu 4 and Xiufeng Xu 1*

1 Department of Psychiatry, First Affiliated Hospital of Kunming Medical University, Kunming, China, 2 Postgraduate College, Kunming Medical University, Kunming, China, 3 Department of Internal Medicine, First Affiliated Hospital of Kunming Medical University, Kunming, China, 4 Department of Medical Imaging, First Affiliated Hospital of Kunming Medical University, Kunming, China

Magnetic resonance imaging (MRI) methods have been used to detect cerebral anatomical distinction between obsessive-compulsive disorder (OCD) patients and healthy controls (HC). Machine learning approach allows for the possibility of discriminating patients on the individual level. However, few studies have used this automatic technique based on multiple modalities to identify potential biomarkers of OCD. High-resolution structural MRI and diffusion tensor imaging (DTI) data were acquired from 48 OCD patients and 45 well-matched HC. Gray matter volume (GMV), white matter volume (WMV), fractional anisotropy (FA), and mean diffusivity (MD) were extracted as four features were examined using support vector machine (SVM). Ten brain regions of each feature contributed most to the classification were also estimated. Using different algorithms, the classifier achieved accuracies of 72.08, 61.29, 80.65, and 77.42% for GMV, WMV, FA, and MD, respectively. The most discriminative gray matter regions that contributed to the classification were mainly distributed in the orbitofronto-striatal “affective” circuit, the dorsolateral, prefronto-striatal “executive” circuit and the cerebellum. For WMV feature and the two feature sets of DTI, the shared regions contributed the most to the discrimination mainly included the uncinate fasciculus, the cingulum in the hippocampus, corticospinal tract, as well as cerebellar peduncle. Based on whole-brain volumetry and DTI images, SVM algorithm revealed high accuracies for distinguishing OCD patients from healthy subjects at the individual level. Computer-aided method is capable of providing accurate diagnostic information and might provide a new perspective for clinical diagnosis of OCD.

Keywords: obsessive-compulsive disorder, support vector machine, structural magnetic resonance imaging, brain volumetry, diffusion tensor imaging
INTRODUCTION

The core characteristics of obsessive-compulsive disorder (OCD) are intrusive recurrent thoughts and/or repetitive behaviors. Abnormalities of gray matter (GM) and white matter (WM) microstructures are considered to be related with its neural pathogenesis. Nowadays magnetic resonance imaging (MRI) approaches provide a perspective to investigate the neuropathological changes of OCD and allow researchers to identify better biological markers of this disease. Voxel-based morphometry (VBM) analysis allows the investigation of gray matter volume (GMV) and white matter volume (WMV) in the whole brain (1). Diffusion tensor imaging (DTI), on the other hand, is available to measure the in vivo water molecule diffusion within the WM fibers, which renders more exquisite details on microstructure changes in WM (2). Fractional anisotropy (FA) and mean diffusivity (MD) are the two most widely used diffusion indices to investigate the pathology of OCD (3, 4).

To date numerous neuroimaging studies have used the between—group comparison—types of analyses to investigate subtle differences between OCD patients and healthy controls (HC). Reported volumetric abnormalities lied in multiple neural structures. Volume reduction in the medial orbitofrontal, anterior cingulate and temporolimbic cortices, and tissue expansion in the striatum and thalamus was among the most widely accepted pathological model of OCD which assumes brain abnormalities in the “affective circuit” (1, 5, 6). Additionally, volume changes of the cortical source of the dorsolateral (DL) prefronto-striatal “executive” circuit (dorsomedial, DL, ventrolateral and frontopolar prefrontal cortices), and of reciprocally connected regions (temporo-parieto-occipital associative areas) are consistently described in OCD patients (6). Reported WM integrality abnormalities also involved extensive brain areas, such as the corpus callosum (CC), cingulum bundles, corticospinal tract, superior longitudinal fasciculus (SLF), uncinate fasciculus (UNC), and cerebellum (1, 4). However, the results obtained by various studies were substantially heterogeneous. Neuropsychiatric disorders are usually characterized by indistinct structure abnormalities rather than a significant abnormal region (7). The group-level analysis usually requires large samples and needs to be corrected by strict multiple comparisons, and the alterations are often too small to detect and lack a reliable differentiation between patients and control subjects (8). Furthermore, group-based methods are neither helpful to diagnose patients at the individual level nor to infer specific clinical outcomes for an individual patient (9).

Advances in neuroimaging have resulted in the use of automated techniques for multivariate pattern analysis. The MRI-related machine learning technique offers a systematic approach in developing sophisticated, automatic, and objective classification frameworks for analyzing high-dimensional data and provides promise for improving the sensitivity and/or specificity of detection and diagnosis of disease (10–12). The analysis based on multivariate pattern is more sensitive to identify subtle differences in the brain structure than group-level statistics. Machine learning-based pattern recognition techniques hold the potential to detect biomarkers on the basis of neuroimaging data and make it possible to combine complementary information from different sources efficiently (13). The techniques also enable people to predict the future course of the disease and the response to treatment at the individual patient level (14). Thus, it holds high clinical values. The support vector machine (SVM) pattern recognition has been demonstrated to be useful for MRI classification. The SVM algorithm establishes model by discriminating the different categories (such as patients and controls) and further applying new data to test its generalizability (15, 16). So far, SVM classifications have been successfully applied to a range of MRI modalities aiming to automate the diagnosis of neuropsychiatric disorders in numerous studies (17–29), and the high accuracies were extremely impressive.

To our best knowledge, few studies have used the SVM approach based on MRI data to perform diagnostic or predictive investigations in OCD patients. A recent review of OCD multivariate pattern analysis based on neuroimaging data found that it is able to classify OCD diagnosis with accuracies ranging from 66 up to 100% (30). Among these studies, Li et al. (16) found FA value showed promise for discriminating OCD from healthy controls (HC) using SVM. The study provided the first evidence of a quite high identification rate presented as a sensitivity of 86%, a specificity of 82% and a significant accuracy of 84%. Moreover, they found the white matter regions which contributed the most to such discrimination mainly included bilateral prefrontal and temporal regions, inferior fronto-occipital fasciculus, superior fronto-parietal fasciculus, splenium of corpus callosum, and left middle cingulum bundle. Hu et al. (31) used structural neuroanatomy of GM and WM volume and reported these structural images could accurately discriminate between patients with OCD and HC. They found the classification accuracies for SVM using GM and WM anatomy were all above 75%, and three main distributed networks including the fronto-striatal circuit, the tempo-parieto-occipital junction and the cerebellum provided high discriminative power. Several other OCD studies conducted different feature selection algorithms and revealed a more comprehensive characterization of the disorder, thus yielding a superior identification of OCD patients based on their brain anatomy (32, 33). Besides, in a psychiatric symptoms predicting study, Hoexter et al. (34) employed support vector regression (SVR) in 37 treatment-naïve adult patients to evaluate whether gray matter volumes of the cortical–subcortical loops contain informative value to predict OCD symptom severity. They found the left medial orbitofrontal cortex and the left putamen may identify neurobiological markers to predict OCD symptom severity based on individual structural MRI datasets.

Multimodal neuroimaging features may reflect different aspects of brain tissue and may be a supplement to comprehend the pathological mechanism of OCD (28, 35). Our previous work has found that OCD patients exhibited symptom-related reduced right posterior cingulate cortex cortical thickness, and disrupted WM integrity in the genu and body of corpus callosum (36). Biomarkers from multiple modalities may provide complementary information for the diagnosis of OCD. However, existing MRI-based SVM analysis studies focused on only a single
modality of MRI. To further comprehensively and systematically evaluate the SVM approach applied in distinguishing patients with OCD from HC, and also to provide classification result of multiple indices and estimate which type of data is more valuable for the detection of OCD, we performed an automated classification with structural-MRI-derived GMV and WMV, along with DTI-derived FA and MD values in a large sample of OCD patients and matched HC. Ten brain regions that weighted the most in the identification of the four features were also calculated. To the best of our knowledge, this is the first study using SVM for classification of OCD patients based on whole-brain volumetry and tractography indices. We expected our study would be informative for the early detection of patients with OCD in future clinical applications.

MATERIALS AND METHODS

Subjects
This study was approved by the Ethics Review Board of Kunming Medical University. The written informed consent of each participant was obtained before the study.

In this study, 48 OCD patients were recruited from the psychiatry department of the First Affiliated Hospital of Kunming Medical University. One experienced clinical psychiatrist made the OCD diagnosis according to the Diagnostic and Statistical Manual of Mental Disorders-Fourth Edition (DSM-IV) criteria. Medication status of OCD sample were listed in the Supplementary Material (see, Table S1). Forty-five well-matched healthy control volunteers were recruited by advertisement. All the participants were right-handed Han Chinese individuals aged from 18 to 55 years. The exclusion criteria for participants involved in both groups were as below: (1) having a previous history of other psychiatric or neurological illness or serious physical disease, (2) present or previous history of substance abuse, (3) physical limitations to undergo an MRI scan, (4) pregnant women, (5) right handedness. Besides, all patients’ obsessive-compulsive symptoms were not caused by another mental disorder or physical disease. Demographic data, including age, sex, duration of illness, and clinical symptom ratings were obtained. The Yale-Brown Obsessive-Compulsive Scale (Y-BOCS) was used to evaluate obsessive-compulsive symptoms. In consideration of the results might be influenced by comorbidity of depression and anxiety, the patients with elevated depression (The Hamilton Depression Rating Scale score, HDRS >17) or anxiety (Hamilton Anxiety Scale score, HAMA >14) symptoms were excluded.

Demographic and clinical characteristics differences between two groups were analyzed using IBM SPSS Statistics (version 21.0; IBM, Armonk, NY, USA). Age difference was assessed by the Independent samples t-test. Chi-Square test was performed to compare gender difference.

Image Acquisition
MRI scanning was carried out by a skilled radiological technician at a Philips Achieva 3.0T scanner (Philips Healthcare, Best, the Netherlands). First, acquisitions included a conventional normal T1- and T2-weighted sequences to rule out obvious structural abnormalities such as cerebrovascular diseases.

3D T1-weighted volumetric structural MRI scan sequence was acquired using a fast spoiled gradient recalled acquisition (FSPGR) with the following parameters: TR/TE = 7.38/3.4 ms, matrix size = 256 × 256, FOV = 250 × 250 mm, number of slices: 230, flip angle = 90°, scan time = 6 min 53 s.

DTI images were acquired using an echo-planar imaging (EPI) sequence in 50 axial planes and was collected along 33 independent orientations through the whole brain using the following parameters: TR/TE = 6.800 ms/80 ms, slice thickness = 3 mm, FOV = 230 mm² × 230 mm², matrix size = 116 × 112, voxel size = 1.98 mm × 2.05 mm × 3 mm, b value = 1000 s/mm², flip angle = 90°.

Image Preprocessing
Structural MRI images were preprocessed using the VBM8 toolbox (http://dbm.neuro.uwi-jena.de/vbm) with the Diffeomorphic Anatomical Registration using the Exponentiated Liealgebra (DARTEL) toolbox (37) implemented in the statistic parametric mapping software package (SPM8, http://www.fil.ion.ucl.ac.uk/spm) running on Matlab 2012a (MathWorks, Natick, MA, USA). This procedure comprises creating a study-specific template and segmenting each individual image using the template aiming to maximize accuracy and sensitivity (29). Then the GM, WM and cerebrospinal fluid (CSF) were automatically segmented. After Jacobian modulation, the GM images and WM images were smoothed with 8-mm full width at half maximum (FWHM) Gaussian kernel for further SVM analysis.

The diffusion MRI data were processed by the FMRIB Software Library (FSL, Version 5.0; available from http://fsl.fmrib.ox.ac.uk/fsl). Tract-based spatial statistics (TBSS) was used to perform voxel-wise processing of diffusivity measures. First, the skull was stripped using Brain Extraction Tool (BET) of FSL. Then the head motion and eddy current distortion correction were conducted using the b = 0 volume as a reference. FMRIB’s Diffusion Toolbox (FDT) was used to fit the tensor model and to compute the FA and MD images using the standard FSL protocol similarly with the previous DTI studies (38–40). Then, the FA and MD of each subject were aligned to the FMR158-FA template (http://fsl.fmrib.ox.ac.uk/fsl/fslwiki/FMRIB58_FA) by using the non-linear image registration tool (FNIRT). Subsequently, the mean FA and MD images in each WM tract were calculated for each subject. Then a mean FA skeleton was generated as the mean FA image was created by averaging all aligned FA maps, which represents the centers of all fiber tracts common to all subjects. A threshold of FA ≥0.2 was set to include the major white matter pathways. The MD maps were acquired using the same non-linear transformations as the FA maps.

Support Vector Machine Analysis
SVM was applied by using the Pattern Recognition for Neuroimaging Toolbox (PRoNTo) (41) (http://www.mlnl.cs.ucl.ac.uk/pronto) to estimate potential WM areas contribute most in calcifying OCD. Briefly, the main steps of the SVM method include: (i) feature extraction and feature selection, (ii) selecting discriminative regions, (iii) training the SVM classifier model
using the training data, and (iv) evaluating the performance of the SVM model using the evaluation data (9, 38).

In neuroimaging studies, the number of features is much more than the number of subjects, which induces the “curse of dimensionality” in machine learning studies (42, 43). Feature extraction allows the original data to transform into analyzable input data of SVM. In the present study, each 3D image was transformed into a column vector of features and each value corresponded to a single corresponding voxel intensity. Thus, this feature vector encoded the pattern of GMV, WMV, FA, or MD values. By comparison, feature selection involves the selection of a subset of features which facilitates learning (15, 44). In this study, feature selection consisted of identifying brain regions that are expected to differ between groups (15, 44). Above procedures were automatically processed in PRONTo’s “Prepare feature set” programs.

A leave-one-out cross-validation method was carried out to perform SVM classifier validation, where the feature selection was performed each time on the training partition of the data to avoid circularity effects. In this study, it involved the exclusion of a single subject from each group and training the classifier using the remaining subjects. Then the excluded subject pair was used to test the ability of the classifier to classify new cases reliably. Above procedures were repeated for each subject pair so that it could obtain a relatively unbiased estimate of generalizability (44). Above procedures were automatically processed in PRONTo’s “Specify model” programs. And the whole process had been described in previous studies detailedly (24).

As for performance evaluation, once the SVM algorithm has been established, it is used to predict a new and previously unseen subject to and decide which group it belongs (44). A 1,000-times non-parametric permutation test (24, 28, 45, 46) was used to obtain a corrected p-value to determine the statistical significance of the accuracy, sensitivity and specificity. In detail, accuracy is the proportion of subjects correctly classified into the patient or control group. Sensitivity and specificity represent the proportion of subjects classified correctly. Besides, receiver operating characteristic (ROC) analysis and AUC (the area under the ROC curve) were used to evaluate the performance of the classifiers. AUC represents the classification power of a classifier, and a larger AUC indicates better classification ability (28, 47). This involves the repetitions of the classification procedure with training group labels by multiple random distribution of the computer in the aim of generating a null distribution of accuracies (44). We also performed a support vector regression of the magnitude of differences and illness duration.

For each model, the PRONTo allows it possible to calculate images representing the weights per voxel and also images summarizing the weights per regions of interest as defined by an atlas (41). The region contributions can be ranked in descending order, yielding a sorted list of regions according to their contribution to the classification model. To investigate classification power of specific locations in the brain, we computed vector weights and listed 10 brain regions that have the highest value of the discrimination map of each feature. For the GMV feature, the most commonly used AAL atlas (48) was selected, which contains 116 cortical and subcortical anatomical structures. For the WMV and DTI feature sets, the weights were calculated based on the ICBM-DTI-81 white-matter atlas, containing 48 WM fiber tract labels (49).

### RESULT

#### Demographic and Clinical Characteristics

The demographic and clinical characteristics of the subjects were presented in Table 1. No significant difference between groups was found in gender and age. The 48 OCD patients had an average duration of illness of 45.42 ± 41.02 months, the total Y-BOCS score was 25.50 ± 3.56, the Y-BOCS obsession score was 12.90 ± 2.40, the Y-BOCS compulsion score was 12.58 ± 3.07. The total HDRS and HAMA scores were 8.10 ± 3.71 and 9.29 ± 2.89, respectively (Table 1).

| Demographics and clinical characteristics of the sample. | OCD patients (n = 48) | HC (n = 45) | t/F/χ² | p-Value |
|----------------------------------------------------------|-----------------------|------------|--------|--------|
| Age, years                                               | 32.29 ± 12.62         | 30.62 ± 9.02 | 4.733  | 0.464  |
| Gender (male/female)                                     | 27/21                 | 24/21      | 0.080  | 0.778  |
| Duration (month)                                         | 45.42 ± 41.02         | –          | –      | –      |
| Y-BOCS total score                                       | 25.50 ± 3.56          | –          | –      | –      |
| Y-BOCS obsession score                                   | 12.90 ± 2.40          | –          | –      | –      |
| Y-BOCS compulsion score                                  | 12.58 ± 3.07          | –          | –      | –      |
| HDRS score                                               | 8.10 ± 3.71           | –          | –      | –      |
| HAMA score                                               | 9.29 ± 2.89           | –          | –      | –      |

OCD, obsessive-compulsive disorder; HC, Healthy Controls; Y-BOCS, Yale-Brown Obsessive-Compulsive Scale; HDRS, Hamilton Depression Rating Scale; HAMA, Hamilton Anxiety Scale.

#### Classifier Evaluation

In the classification of the two groups, for the whole brain GMV, the accuracy was 72.04% (permutation p = 0.001) with a sensitivity of 70.83% and a specificity of 73.33%. For the whole brain WMV, the accuracy was 61.29% (permutation p = 0.040) with a sensitivity of 64.58% and a specificity of 57.78%. Results for the FA and MD values were more promising. A classification accuracy of 80.65 to 77.42% were achieved for the two feature sets of DTI (permutation p = 0.001), respectively. Sensitivity of the two parameters was 81.25 and 75%, respectively, and specificity of both FA and MD values were 80% (Table 2). No significant results were found between the four features and duration (see Table 3).

Receiver Operating Characteristic (ROC) curve and area under the ROC (AUC) were also achieved (Figure 1, Table 2).
Regions Contributed Most for Classification

For GMV feature, the most informative regions for classification between OCD patients and HC included right anterior cingulate gyrus (ACG), right angular gyrus, right inferior parietal, bilateral paracentral lobule, left inferior frontal gyrus, and bilateral cerebelum regions. For the WMV feature and two feature sets of DTI, regions contributed the most to the discrimination were relatively consistent, mainly included the UNC, the cingulum in the hippocampus, corticospinal tract, as well as cerebellar peduncle. Additionally, right external capsule, left fornix and stria terminalis, left anterior corona radiate, bilateral cerebral peduncle, pontine crossing tract and bilateral cerebral peduncle are among the informative regions for classification (Table 3). (Detailed results could be obtained in Table S2).

DISCUSSION

To the best of our knowledge, the present study is the first to simultaneously detect whole-brain volumetric and tractography

(Continued)
The current results showed that it is achievable to classify SVM studies on OCD were mainly based on single modal. Nowadays clarify the neuropathological mechanism of OCD. Nowadays, these provided various perspectives for the investigation of pathological characteristics, as well as the detection of biomarkers for OCD diagnosis. GM regions with high discriminative power are consistent with the previous structural MRI multivariate pattern analysis (31) and numerous VBM studies of OCD (1, 6), mainly distributed in the orbitofrontal region “affective” circuit, the DL prefrontal “executive” circuit and the cerebellum (6). The cingulate cortex plays a key role in the “affective” circuit and ACG is considered to be involved in executive control, particularly of emotion-related processes (6, 53). Abnormal regional ACG volume has also been reported in previous VBM studies (54–57). The ACG volume deficit might mediate the cognitive symptoms (e.g., cognitive behavioral inflexibility) which were often observed in OCD patients (6). Previous functional MRI study in our research group also found that OCD patients exhibited decreased regional activity in the posterior cingulate cortex and increased activity in the ACG (58). Thus, characteristic morphometry changes of ACG might be a crucial biomarker of OCD. The right angular gyrus and right inferior parietal gyrus also showed a relatively high weight in the classification. The angular gyrus lies in the superior portion of the right lateral parietal cortex (59). Neuroimaging studies have reported abnormalities in the right parietal lobe, particularly in the angular gyrus (53, 60). Parietal lobe is important in a variety of executive tasks involving attention, spatial perception and working memory functions (53). Considering that impairment in some of these functions are consistently reported in OCD, such as attentional shifting, it is conceivable that parietal lobe dysfunction, particularly within the parietal cortex, could contribute to the cognitive deficits evident in OCD (53, 61). In addition, the parietal lobe might interact with the frontal subcortical circuitry of OCD through direct anatomical connections between associative parietal areas and some of the key regions implicated in the disorder such as the lateral orbitofrontal cortex, the striatum and the mediodorsal thalamic nucleus (6). The paracentral lobule is divided into anterior and posterior to the central sulcus. The anterior portion of the paracentral lobule is part of the frontal lobe and is often referred to as the supplementary motor area; the posterior portion is considered part of the parietal lobe (62, 63). The function of the parietal lobule has been proved to be primarily associated with skill learning, attention, and working memory (63). Therefore, in addition to the ACG, the frontal lobe and parietal lobe, paracentral lobule may also play key roles in the pathology of OCD and can provide much reference for classification of the disease.

Myelin is closely related with the complex connectivity of the human brain. Notably, in this study, three distinct

| Modality | Hemisphere | Brain regions | ROI index | Discriminative weight (%) | Cluster Size |
|----------|------------|---------------|-----------|----------------------------|--------------|
| L        | Inferior cerebellar peduncle | 12 | 4.991 | 125 |
| R        | Corticospinal tract | 7 | 4.763 | 168 |
| R        | Cingulum (hippocampus) | 37 | 3.709 | 172 |
| R        | Cerebral peduncle | 15 | 3.459 | 301 |
| L        | Pontine crossing tract | 2 | 3.328 | 198 |
| L        | Cerebral peduncle | 16 | 3.276 | 312 |
| R        | Superior cerebellar peduncle | 13 | 3.268 | 128 |
| L        | Cingulum (hippocampus) | 38 | 3.182 | 157 |

L, left; R, right. GMV feature was estimated using AAL atlas. WMV, FA, and MD features were estimated using ICBM-DTI-81 white-matter atlas.

abnormalities of OCD patients using SVM. Compared with the mass univariate analysis, SVM takes into account the inter-regional correlations, and provides numerical indicators for group membership without multiple comparison biases (44). We made this study aiming to develop an objective method facilitate the accuracy of the diagnosis and assist humans with clinical judgment. The primary finding of this preliminary study is that it is feasible to discriminate OCD patients from HC with high accuracies by using a neuroimaging-based computer-aided method.

According to the recent review of multivariate pattern analysis techniques on neuroimaging data of OCD (30), studies were able to classify OCD diagnosis with accuracies ranging from 66 up to 100%. Methodological limitations such as heterogeneity of sample characteristics, differences in data processing, acquisition or statistical analyses might induce these inconsistencies. In our study, the SVM analysis demonstrated better performance by using diffusion parameters than volumetric features for distinguishing OCD patients from HC. Among the four feature sets, FA came out to be the most significant one with a classification accuracy of 80.65%. This is close to the SVM classification on DTI-derived FA features performed by Li et al. (16). As FA and MD reflect white matter microstructure characteristics, our finding indicated that OCD patients might exhibited more detectable alterations in white matter integrity. Our previous multimodal group-level study detected OCD patients exhibited reduced cortical thickness of the right posterior cingulate cortex and altered WM integrity in the genu and body of corpus callosum. Other multimodal imaging analyses reported OCD patients displayed alterations in brain structures and functions, involving complex brain networks (50–52). Multivariate methods provide a broader perspective to clarify the neuropathological mechanism of OCD. Nowadays SVM studies on OCD were mainly based on single modal. The current results showed that it is achievable to classify OCD patients and healthy people with multiple indices using an automatic machine learning way. The SVM algorithm calculates a region’s discriminative power depends not only on between-group differences in its absolute values, but also on any between-group differences in its structural correlations with other regions. Therefore, this method is particularly suitable to explore mental disorders such as OCD in which abnormalities are distributed across the whole brain (16).
WM classification models (WMV, FA, and MD) revealed a coherent set of discriminative features, which were primarily associated with processes of decision making or behavior control. UNC and cingulum are within the association fiber tracts (49). The UNC connects the frontal lobe (orbital cortex) and the anterior temporal lobe, while cingulum especially the hippocampal regions carries information from the cingulate gyrus to the hippocampus (49). The corticospinal tract is an important part of the motor system, and its integrity change has been widely detected in various studies (4). Disrupted white matter connectivity in these regions might underlie the neural mechanisms of decline in memory and behavior control in OCD patients. Since both the WM volumetric feature and the tractography indices indicated high discriminative power in the UNC, the cingulum in the hippocampus and corticospinal tract, we hypothesized that the above three regions (the UNC, the cingulum in the hippocampus and corticospinal tract) hold the most discriminative WM connections of OCD and are likely to be the specific biological markers of OCD. Additionally, other discovered informative white matter regions of the right external capsule, left fornix, and stria terminalis are parts of the association fibers (cortex-cortex connections), and the left anterior corona radiate and bilateral cerebral peduncle are defined as projection fibers (cortex-spinal cord, and cortex-thalamus connections) (49, 64). Damages of these WM tracts might interfere with the connectivity between brain regions and disrupt the brain networks involved in mood and cognition. These regions might provide informative value for detection of OCD as well.

It is particularly noteworthy that cerebellar regions showed a fairly high discriminative weight in all four features. This finding is in line with a previous structural study using the multivariate pattern analysis in OCD (31). Existing structural and functional neuroimaging studies have observed abnormal structures and functions in the cerebellum of OCD patients (65) and enriched our understanding of the great role of cerebellum in OCD illness models. Increasing evidence has demonstrated that in addition to its well-known role in motor control, the cerebellum also plays roles in cognitive and emotional regulatory processes (31, 66). The cerebellum is structurally and functionally connected to the parallel cortico-striato-thalamo-cortical circuits, as well as the limbic-cortical network (31, 67), which forms a feedback information flow that allows the cerebellum to involve in advanced neural activities. Moreover, abnormal cerebellar functions are suggested to be related to inflexible movements and repetitive behaviors in OCD (68). Hence, the cerebellum deserves more attention in future pathological model of OCD, and our results provide further evidence for neural implications of this region in OCD.

There are some limitations in this study to be addressed. First, this study aims not only to use multi-modal features for OCD classification, but also estimates which type of data is more valuable for the detection of OCD. So, this study did not investigate combined imaging modalities for classification. Multi-kernel learning based on multi-modal features with large sample might provide a better perspective to classify OCD from healthy controls in future studies. Second, our patients group consisted of both medicated and drug-naïve subjects. Medication status of the OCD group were various and complicated. Correlation between medication status and brain images was not analyzed separately. Therefore, we did not clarify how medication status would affect the brain. Last but not least, we innovatively estimate the 10 most discriminative brain regions to exhibit a preliminary result of the multiple feature sets. Some unlisted regions may also contain valuable discriminative information and these regions should receive full considerations in future research.

**CONCLUSION**

Based on whole-brain volumetry and DTI images, SVM algorithm revealed high accuracies for distinguishing OCD patients from healthy controls at the level of the individual. Feature sets of DTI seem to offer better predictive value than volumetric features. Though machine learning method sole is not enough to diagnose the OCD patient than the clinical symptom, we believe neuroimaging-based machine learning techniques may suggest neurobiological markers for automatic detection of OCD patients and develop an objective method to facilitate the accuracy of the diagnosis. We hope computer-aided method based on imaging biomarker could assist clinicians with the diagnosis of mental disorders in the future.

**ETHICS STATEMENT**

This study was approved by the Ethics Review Board of Kunming Medical University (ClinicalTrials.gov: NCT01298622). All of the subjects involved in the study gave written informed consent in accordance with the Declaration of Helsinki.

**AUTHOR CONTRIBUTIONS**

XX designed the study and revised the manuscript. CZ, YC, LP, JX, ZS, LS, and SY carried out experiments and undertook the statistical analysis. LJ managed the diagnosis and clinical measurement. CZ wrote the first draft of manuscript. YL were responsible for acquiring MRI image data and assisted with the analysis. YC and LP contributed to the final manuscript.

**ACKNOWLEDGMENTS**

This study was supported by the National Natural Science Foundation of China (81560233, 81660237, 81460256), National Clinical Research Center on Mental Disorders (2015BAI13B02), Founding of Yunnan Provincial Health Science and Technology Plan (2016NS026), Yunnan Applied Basic Research Projects-Union Foundation [2017FE467(-167)], Yunnan Applied Basic Research Projects-Union Foundation [2017FE468(-176)], Innovative Research Team of Kunming...
Medical University (CXTD201705), and Middle and Young Aged Academic and Technology Leaders Reserve Personnel Foundation of Yunnan Province (2017HB062). We would like to thank Binli Shang for her assistance of the stats.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2018.00524/full#supplementary-material

REFERENCES

1. Friedman I, de Salles Andrade JB, Vigne P, Fontenelle LF. Can neuroimaging provide reliable biomarkers for obsessive-compulsive disorder? a narrative review. Curr Psychiatry Rep. (2016) 18:90. doi: 10.1007/s11920-016-0729-7
2. Le Bihan D, Mangin JF, Poupon C, Clark CA, Pappata S, Molko N, et al. Diffusion tensor imaging: concepts and applications. J Magn Reson Imag. (2001) 13:534–46. doi: 10.1002/jmri.1076
3. Pierpaoli C, Peter J, Peter B, Giovannini DC. Diffusion tensor MR imaging of the human brain. Radiology (1996) 201:637–48. doi: 10.1148/radiol.201.3.8939209
4. Koch K, Reess TJ, Rus OG, Zimmer C, Zaudig M. Diffusion tensor imaging (DTI) studies in patients with obsessive-compulsive disorder (OCD): a review. J Psychiatr Res. (2014) 54:26–35. doi: 10.1016/j.jpsychires.2013.04.006
5. Pujol J, Soriano-Mas C, Alonso P, Menchon JM, Deus J, et al. Mapping structural brain alterations in obsessive-compulsive disorder. Arch Gen Psychiatry (2004) 61:720–30. doi: 10.1001/archpsyc.61.7.7270
6. Piras F, Piras F, Chiapponi C, Girardi P, Callagionere C, Spalletta G. Widespread structural brain changes in OCD: a systematic review of voxel-based morphometry studies. Cortex (2015) 62:89–108. doi: 10.1016/j.cortex.2013.01.016
7. Seeley WW, Crawford RK, Zhou J, Miller BL, Greicius MD. Neurodegenerative diseases target large-scale human brain networks. Neuron (2009) 62:42–52. doi: 10.1016/j.neuron.2009.03.024
8. Kambeitz I, Cabral C, Sacchet MD, Gotlib IH, Zahn R, Serpa MH, et al. Detecting neuroimaging biomarkers for depression: a meta-analysis of multivariate pattern recognition studies. Biol Psychiatry (2017) 82:330–8. doi: 10.1016/j.biopsych.2016.10.028
9. Amarreh I, Meyerand ME, Stafstrom C, Alonso P, Hermann BP, Damasio, et al. Pattern recognition of magnetic resonance imaging-based gray matter volume measurements classifies bipolar disorder and major depressive disorder. J Affect Disord. (2013) 150:457–63. doi: 10.1016/j.jad.2013.08.006
10. Sajda P. Machine learning for detection and diagnosis of disease. Ann Rev Biomed Eng. (2006) 8:537–65. doi: 10.1146/annurev.bioeng.8.061305.095802
11. Rathore S, Habes M, Ifikhar MA, Shacklett A, Davatzikos C. A review on neuroimaging-based classification studies and associated feature extraction methods for Alzheimer’s disease and its prodromal stages. Neuroimage (2017) 155:350–48. doi: 10.1016/j.neuroimage.2017.03.057
12. Volkow ND, Baler RD, Compton WD, Maruyama M. The prevalence of addiction severity and co-occurring disorders: results from the National Epidemiologic Survey on Alcohol and Related Conditions (2001) 37:328–49. doi: 10.1038/nature.2005.8001
13. Koppel S, Abdulkadir A, Jack CR Jr, Joutsaouler M, Mourao-Miranda J, Vemuri P. Diagnostic neuroimaging across diseases. Neuroimage (2012) 61:457–63. doi: 10.1016/j.neuroimage.2011.11.002
14. Noble WS. What is a support vector machine? Nat Biotechnol. (2006) 24:1565–70. doi: 10.1038/nbt1206-1565
15. Li F, Huang X, Tang W, Yang Y, Li B, Kemp GF, et al. Multivariate pattern analysis of DTI reveals differential white matter in individuals with obsessive-compulsive disorder. Hum Brain Mapp (2014) 35:2643–51. doi: 10.1002/hbm.22357
16. Ingallhalikar M, Parker D, Bloy L, Roberts TP, Verma R. Diffusion based abnormality markers of pathology: toward learned diagnostic prediction of ASD. Neuroimage (2011) 57:918–27. doi: 10.1016/j.neuroimage.2011.05.023
17. O’Dwyer L, Lamberton F, Bokde AL, Ewers M, Faluji YO, Tanner C, et al. Using support vector machines with multiple indices of diffusion for automated classification of mild cognitive impairment. PLoS ONE (2012) 7:e32441. doi: 10.1371/journal.pone.0032441
18. Besga A, Termenon M, Grana M, Echeveste J, Perez JM, Gonzalez-Pinto A. Discovering Alzheimer’s disease and bipolar disorder white matter effects building computer aided diagnostic systems on brain diffusion tensor imaging features. Neurosci Lett. (2012) 520:71–6. doi: 10.1016/j.neulet.2012.05.033
19. Dyrba M, Ewers M, Wegryn M, Kilimann I, Plant C, Oswald A, et al. Robust automated detection of microstructural white matter degeneration in Alzheimer’s Disease using machine learning classification of multilcenter DTI Data. PLoS ONE (2013) 8:e64925. doi: 10.1371/journal.pone.0064925.001
20. Folkke NC, Helms G, Schewe S, Pannell P, Bachmann CG, Dechant P, et al. Individual voxel-based subtype prediction can differentiate progressive supranuclear palsy by idiopathic Parkinson syndrome and healthy controls. Hum Brain Mapp. (2011) 32:1905–15. doi: 10.1002/hbm.21161
21. Deng F, Wang Y, Huang H, Niu M, Zhang S, Zhao L, et al. Abnormal segments of right uncinate fasciculus and left anterior thalamic radiation in major and bipolar depression. Prog Neuropsychopharmacol Biol Psychiatry (2018) 81:340–9. doi: 10.1016/j.pnpbp.2017.09.006
22. Bi K, Hua L, Wei M, Qin J, Lu Q, Yao Z. Dynamic functional-structural coupling within acute functional state change phases: evidence from a depression recognition study. J Affect Disord. (2016) 191:145–55. doi: 10.1016/j.jad.2015.11.041
23. Schnyer DM, Clasen PC, Gonzales C, Bevers GS. Evaluating the diagnostic utility of applying a machine learning algorithm to diffusion tensor MRI measures in individuals with major depressive disorder. Psychiatry Res. (2017) 264:1–9. doi: 10.1016/j.psychres.2017.03.003
24. Rozycki M, Satterthwaite TD, Koutsouleris N, Erus G, Doshi J, Wolf DH, et al. Multisite machine learning analysis provides a robust structural imaging signature of schizophrenia detectable across diverse patient populations and within individuals. Schizophr Bull. (2017) 44:1035–44. doi: 10.1093/schbul/sbx137
25. Sacchet MD, Prasad G, Roland-Ross LC, Thompson PM, Gotlib IH. Support vector machine classification of major depressive disorder using diffusion-weighted neuroimaging and graph theory. Front Psychiatry (2015) 6:21. doi: 10.3389/fpsyt.2015.00021
26. Rubin-Falcone H, Zanderigo F, Thapa-Chhetry B, Lan M, Miller JM, Sublette ME, et al. Pattern recognition of magnetic resonance imaging-based gray matter volume measurements classifies bipolar disorder and major depressive disorder. J Affect Disord. (2018) 227:498–505. doi: 10.1016/j.jad.2017.11.043
27. Cui Z, Xia Z, Mu S, Shu H, Gong G. Disrupted white matter connectivity underlying developmental dyslexia: a machine learning approach. Hum Brain Mapp. (2016) 37:1443–58. doi: 10.1002/hbm.23112
28. Krakauer K, Edbrup BH, Glenthoj BY, Raghava JM, Nordholm D, Randers L, et al. Patterns of white matter microstructure in individuals at ultra-high-risk for psychosis: associations to level of functioning and clinical symptoms. Psychol Med. (2017) 47:2689–707. doi: 10.1017/S0033291717001210
29. Bruin W, Denys D, van Wingen G. Diagnostic neuroimaging markers of obsessive-compulsive disorder: initial evidence from structural and functional MRI studies. Prog Neuropsychopharmacol Biol Psychiatry (2018) 81:405–9. doi: 10.1016/j.pnpbp.2017.09.006
30. Hu X, Liu Q, Li B, Tang W, Sun H, Li F, et al. Multivariate pattern analysis of obsessive-compulsive disorder using structural neuroanatomy. Eur Neuropsychopharmacol. (2016) 26:246–54. doi: 10.1016/j.euroneuro.2015.12.014
31. Parrado-Hernandez E, Gomez-Verdejo V, Martinez-Ramon M, Shawe-Taylor J, Alonso P, Pujol J, et al. Discovering brain regions relevant to obsessive-compulsive disorder identification through bagging and transduction. Med Image Anal. (2014) 18:435–48. doi: 10.1016/j.media.2014.01.006
33. Trambaioli LR, Biazoli CE, Jr., Balardín JB, Hoexter MQ, Sato JR. The relevance of feature selection methods to the classification of obsessive-compulsive disorder based on volumetric measures. *J Affect Disord*. (2017) 222:49–56. doi: 10.1016/j.jad.2017.06.061

34. Hoexter MQ, Miguel EC, Dinia JR, Shavitt RG, Busatto GF, Sato JR. Predicting obsessive-compulsive disorder severity combining neuroimaging and machine learning methods. *J Affect Disord*. (2013) 150:1213–6. doi: 10.1016/j.jad.2013.05.041

35. Wee CY, Yap PT, Li W, Denny K, Browndyke JN, Potter GG, et al. Enriched white matter connectivity networks for accurate identification of MCI patients. *Neuroimage* (2011) 58:1841–22. doi: 10.1016/j.neuroimage.2010.10.026

36. Zhou C, Xu J, Ping L, Zhang F, Chen W, Shen Z, et al. Cortical thickness and white matter integrity abnormalities in obsessive-compulsive disorder: a combined multimodal surface-based morphometry and tract-based spatial statistics study. *Depress Anxiety* (2013) 38:742–51. doi: 10.1002/da.22758

37. Ashburner J. A fast diffeomorphic image registration algorithm. *Neuroimage* (2007) 38:95–113. doi: 10.1016/j.neuroimage.2007.07.007

38. Dyrba M, Grothe M, Kirste T, Teipel SJ. Multimodal analysis of functional and structural disconnection in Alzheimer’s disease using multiple kernel SVM. *Hum Brain Mapp.* (2015) 36:2118–31. doi: 10.1002/hbm.22759

39. Inano R, Oishi N, Kunieda T, Arakawa Y, Yamao Y, Shibata S, et al. Voxel-based multivariate clustering imaging by multiparameter diffusion tensor imaging for glia grading. *Neuroimage Clin.* (2014) 5:396–407. doi: 10.1016/j.nicl.2014.08.001

40. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp.* (2002) 17:143–55. doi: 10.1002/hbm.10062

41. Schrouff J, Rosa MJ, Rondina JM, Marquand AF, Chu C, Ashburner J, et al. PRoNTo: pattern recognition for neuroimaging toolbox. *Bioinformatics* (2013) 11:319–37. doi: 10.1093/bioinformatics/bts1201-03-1978-1

42. De Martino F, Gentile F, Esposito F, Balsi M, Di Salle F, Goebel R, et al. Classification of fMRI independent components using IC-fingerprints and support vector machine classifiers. *Neuroimage* (2007) 34:177–94. doi: 10.1016/j.neuroimage.2006.08.041

43. Fort G, Lambert-Lacroix S. Classification using partial least squares. *Comput Stat Data Anal.* (2005) 49:133–47. doi: 10.1016/j.csda.2004.02.012

44. Orru G, Pettersson-Yeo W, Marquand AF, Sartori G, Mechelli A. Using support vector regression to predict neuroimaging features. *Neuroimage* (2010) 46:831–44. doi: 10.1016/j.cortex.2009.11.008

45. Ojala M, Garriga GC. Permutation tests for studying classifier performance. *Neuroimage* (2006) 32:525–49. doi: 10.1016/j.neubiorev.2007.09.005

46. Masumoto R, Ito H, Takahashi H, Ando T, Fujimura Y, Nakayama K, et al. Reduced gray matter volume of dorsal cingulate cortex in patients with obsessive-compulsive disorder: a voxel-based morphometric study. *Psychiatry Clin Neurosci.* (2010) 64:541–7. doi: 10.1111/j.1440-1819.2010.02125.x

47. Fawcett T. An introduction to ROC analysis. *Pattern Recognit. Lett.* (2006) 27:861–74. doi: 10.1016/j.patrec.2005.10.010

48. Tzourio-Mazoyer N, Landeau B, Papathanassiou D, Crivello F, Etard O, Delcroix N, et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* (2002) 19:279–91. doi: 10.1016/S1053-8119(02)00059-1

49. Radua J, Mataix-Cols D. Voxel-wise meta-analysis of grey matter changes in obsessive-compulsive disorder. *Br J Psychiatry* (2009) 195:393–402. doi: 10.1192/bjp.bp.108.055046

50. Bush G, Luu P, Posner MI. Cognitive and emotional influences in anterior cingulate cortex. *Trends Cogn Sci.* (2000) 4:215–22. doi: 10.1016/S1364-6613(00)01483-2

51. Radua J, Huelva OAvd, Surguladze S, Mataix-Cols D. Meta-analytical comparison of voxel-based morphometry studies in obsessive-compulsive disorder vs other anxiety disorders. *Arch Gen Psychiatry* (2010) 67:701–11. doi: 10.1001/archgenpsychiatry.2010.70

52. Moreira PS, Marques P, Soriano-Mas C, Magalhaes R, Sousa N, Soares JM, et al. Regional gray matter abnormalities in obsessive-compulsive disorder: a voxel-based morphometry study. *Brain Imaging Behav.* (2005) 58:479–87. doi: 10.1007/s11682-010-9024-1

53. Zhou et al. SVM Classification of OCD

54. Valente AA Jr., Miguel EC, Castro CC, Amaro E Jr., Duran FL, Buchiguel CA, et al. The neuropsychiatry of obsessive-compulsive disorder: a voxel-based morphometry study. *Brain Imaging Behav.* (2012) 640:99–104. doi: 10.106/j.neulert.2017.01.024

55. Niu Q, Yang L, Song X, Chu C, Liu H, Zhang L, et al. Abnormal resting-state brain activities and functional connectivities of the anterior and the posterior cortices in medication-naive patients with obsessive-compulsive disorder. *PLoS ONE* (2013) 8:e67478. doi: 10.1371/journal.pone.0067478

56. Liu L, Fu X, Wang S, Tang Q, Chen X, Cheng L, et al. Abnormal regional spontaneous neuronal activity associated with symptom severity in treatment-naive patients with obsessive-compulsive disorder revealed by resting-state functional MRI. *Neurosci Lett.* (2017) 640:99–104. doi: 10.106/j.neulert.2017.01.024

57. Culham JC, Kanwisher NG. Neuroimaging of cognitive functions in human parietal cortex. *Curr Opin Neurobiol.* (2011) 21:157–63. doi: 10.1016/j.conb.2011.02.002

58. Wakana S, Jiang H, Nagae-Poetscher LM, Zijl PCMv, Mori S. FiberTract-based atlas of human white matter anatomy. *Radiology* (2004) 230:77–87. doi: 10.1148/radiol.2301021640

59. Eng GK, Sim K, Chen SH. Meta-analytic investigations of structural gray matter, executive domain-related functional activations, and white matter diffusivity in obsessive compulsive disorder: an integrative review. *Neurosci Biobehav Rev.* (2015) 52:233–57. doi: 10.1016/j.neubiorev.2015.03.002

60. Ramnani N. Frontal lobe and posterior parietal contributions to the cortico-cerebellar system. *Cerebellum* (2012) 11:366–83. doi: 10.10371/021311-01-0272-3

61. Schmahmann JD, Weilburg JB, Sherman JC. The neuropsychiatry of the cerebellum - insights from the clinic. *Cerebellum* (2007) 6:254–67. doi: 10.1080/14734220701490995

62. Stoodley CJ, Schmahmann JD. Evidence for topographic organization in the cerebellum of motor control versus cognitive and affective processing. *Cortex* (2010) 46:831–44. doi: 10.1016/j.cortex.2009.11.008

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Zhou, Cheng, Ping, Xu, Shen, Jiang, Shi, Yang, Lu and Xu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.