BMC Psychiatry

RESEARCH

Open Access

Predicting treatment response in individuals with major depressive disorder using structural MRI-based similarity features

Sutao Song[1*], Songling Wang[1], Jingjing Gao[1], Lingkai Zhu[1], Wenxin Zhang[2], Yan Wang[3], Donglin Wang[4], Danning Zhang[5*] and Kangcheng Wang[2*]

Abstract

Background Major Depressive Disorder (MDD) is a prevalent mental health condition with significant societal impact. Structural magnetic resonance imaging (sMRI) and machine learning have shown promise in psychiatry, offering insights into brain abnormalities in MDD. However, predicting treatment response remains challenging. This study leverages inter-brain similarity from sMRI as a novel feature to enhance prediction accuracy and explore disease mechanisms. The method's generalizability across adult and adolescent cohorts is also evaluated.

Methods The study included 172 participants. Based on remission status, 39 participants from the Hangzhou Dataset and 34 from the Jinan Dataset were selected for further analysis. Three methods were used to extract brain similarity features, followed by a statistical test for feature selection. Six machine learning classifiers were employed to predict treatment response, and their generalizability was tested using the Jinan Dataset. Group analyses between remission and non-remission groups were conducted to identify brain regions associated with treatment response.

Results Brain similarity features outperformed traditional metrics in predicting treatment outcomes, with the highest accuracy achieved by the model using these features. Between-group analyses revealed that the remission group had lower gray matter volume and density in the right precentral gyrus, but higher white matter volume (WMV). In the Jinan Dataset, significant differences were observed in the right cerebellum and fusiform gyrus, with higher WMV and density in the remission group.

Conclusions This study demonstrates that brain similarity features combined with machine learning can predict treatment response in MDD with moderate success across age groups. These findings emphasize the importance of considering age-related differences in treatment planning to personalize care.

Trial registration Clinical trial number: not applicable.

Keywords MDD, Structural MRI, Treatment outcome, Brain connectivity, Machine learning

*Correspondence:
Sutao Song
sutao.song@sdnu.edu.cn
Danning Zhang
haohaizi21c@sina.com
Kangcheng Wang
wangkangcheng@sdnu.edu.cn
[1]School of Information Science and Engineering, Shandong Normal University, Jinan 250358, China
[2]School of Psychology, Shandong Normal University, Jinan 250358, China
[3]Department of Psychiatry, Affiliated Hospital of Southwest Medical University, Luzhou, China
[4]Institutes of Psychological Sciences, College of Education, Hangzhou Normal University, Hangzhou 311121, China
[5]Shandong Mental Health Center, Shandong University, Jinan, Shandong, China

Introduction

Major Depressive Disorder (MDD) has a serious impact on the physical and mental health of patients. As one of the most prevalent mental disorders, it has become an important public health challenge and one of the leading causes of disability [1, 2]. A number of different approaches can be used for the treatment of MDD, such as pharmacotherapy, psychotherapy, and electroconvulsive therapy (ECT) [3]. The first-line treatment drug for MDD, selective serotonin reuptake inhibitors (SSRI), is only effective in 30–35% of patients [4, 5]. For patients who do not achieve remission, multiple medications or other treatments usually need to be tried. One study showed that two-thirds of patients need to try multiple antidepressants and one-third of patients remain ineffective even after trying multiple medications [6]. For patients, this not only adds to their psychological and emotional burden, but also increases the cost of treatment and even the side effects associated with more medications. There have been many studies that have made attempts to predict treatment response in MDD [7], and being able to predict treatment response prior to starting therapy could reduce the costs associated with trial-and-error approaches, providing significant benefits to patients.

Structural magnetic resonance imaging (sMRI) is a non-invasive neuroimaging technique. Several sMRI studies have found differences in brain structure between individuals with MDD and healthy individuals. For example, patients with MDD may exhibit reduced gray matter volume (GMV) in brain regions such as the frontal lobe and limbic structures compared to healthy individuals [7, 8]. In addition to gray matter changes, patients experiencing a first episode of illness were found to have reduced white matter volume (WMV) in the right superior frontal gyrus and around the anterior portion of the corona radiata. In contrast, those with recurrent episodes showed an increase in WMV around the parietal lobe and superior temporal gyrus [9]. In addition, medication or physical therapy can have an effect on brain structure. One study found a general increase in cortical thickness after treatment, and another study noted that ECT significantly increased cortical thickness in the medial temporal lobe network, including the hippocampus and its adjacent cortex [10]. Individual differences in structural indicators exhibited by patients often lead directly to significant differences in treatment outcomes, i.e., whether or not symptomatic relief is achieved. For example, one study found that patients with thicker right caudal anterior cingulate cortex at baseline showed greater symptom improvement during follow-up. At the same time, studies have also found subtle temporal differences in brain structure between remitters and non-remitters. Specifically, hippocampal volume and frontal cortex thickness showed an increasing trend in remitters, while the opposite was true for non-remitters, who showed a decreasing trend in the volume and thickness of these regions [11]. More importantly, individual differences in structural indices—such as hippocampal volume, white matter hyperintensities, GMV, and cortical thickness—often have a significant impact on treatment outcomes and can directly determine whether a patient achieves symptomatic relief. While some studies have suggested that certain structural metrics may help predict treatment effects [12–15], these studies also report that specific indicators, such as white matter hyperintensities and hippocampal volume, do not show a direct correlation with antidepressant treatment outcomes [16]. This suggests that more in-depth studies are needed to fully understand the relationship between brain structure and treatment outcome.

Machine learning methods can more sensitively capture the differences in neurological indicators between those who respond effectively to treatment and those who do not, providing a powerful tool to unlock the selection of precise treatment options for individuals with MDD. Using machine learning methods, the researchers successfully predicted the treatment response of patients with MDD to medication and ECT. For instance, researchers were able to predict the effectiveness of medication by evaluating structural indicators, including the volumes of the left middle frontal gyrus and the right angular gyrus in patients [17]. Similarly, by analyzing sMRI data prior to ECT, some studies have been able to predict changes in depressive symptoms and whether patients experience remission after undergoing ECT [18, 19]. However, there are still some challenges, such as feature selection based only on raw features, which does not fully consider the importance of structural similarity of brain regions. This study aims to address these gaps by using inter-brain similarity extracted from sMRI data as a novel feature for predicting treatment outcomes in MDD. Brain region similarity can help identify brain regions with similar structures, thus providing important information about the connections between these regions. Changes in structural similarity may be associated with neurological diseases or pathological states, providing important clues for studying the mechanisms of these diseases. In addition, existing studies have rarely been validated in independent samples, so it is crucial to address the generalizability of the models across different datasets and to improve the prediction accuracy. A key contribution of this study is the use of inter-brain similarity to capture structural changes in the brain that may be relevant to treatment response, thus improving the prediction model's accuracy. To this end, there is a need to further enhance dataset diversity, optimize feature selection and model training strategies, and enhance independent validation and cross-dataset validation. These

measures will help to improve the reliability and accuracy of machine learning for MDD treatment outcome prediction.

Through sMRI and machine learning techniques, we aim to accurately predict treatment response in patients with MDD. The contributions of this study include the introduction of inter-brain similarity as a key feature, the application of various machine learning methods to predict treatment response, and the exploration of age-related differences in treatment outcomes across adolescent and adult cohorts. In this study, we propose to use the inter-brain similarity of sMRI as a core feature extraction method to extract the corresponding features from the brain GMV, gray matter density (GMD), WMV, and white matter density (WMD) feature space, which can be used to predict treatment response in MDD. We will extensively test the current mainstream machine learning methods and clarify the impact of various machine learning methods on model performance. More importantly, we plan to conduct pre-intervention structural image data collection at two independent sites for adolescents and adults with MDD, respectively. This initiative aims to comprehensively assess the generalizability of our research methodology to this population of adolescents and to conduct controlled analyses with the adult dataset. The innovative aspect of the study is that for the first time, we used inter-brain similarity of sMRI as a key feature for machine learning and conducted a whole-brain analysis of the four features of sMRI. By conducting a comparative study between these two different age groups, we not only expect to validate the reliability of the model in an adult population, but also its generalization and applicability in a population of adolescents with MDD. This study's findings will provide new insights into personalized treatment approaches and contribute to optimizing MDD treatment strategies across diverse populations. This study design not only offers innovative perspectives and support for personalized and precise MDD treatment but also establishes a crucial scientific foundation for the early identification and intervention of adolescent MDD.

Materials and methods

Introduction to the data set

In this study, data on MDD from mental health hospitals in two Chinese provinces were used. One of the datasets, Dataset I, was from one of our previous studies [20]. It contained data from Hangzhou City, China, and is referred to in this study as the Hangzhou Dataset. It included a sample of adult patients. Dataset II was from Jinan City, China, and is referred to in this study as the Jinan Dataset. It included a sample consisting of adolescents. Dataset II was used to assess the stability of the performance of structural indicators of MDD treatment effects across age groups.

Subjects

For the Hangzhou Dataset, a total of 118 subjects were recruited, including 74 patients with MDD (MDD group) and 44 healthy volunteers (HC group). The patients were from the Department of Psychiatry of the Seventh People's Hospital of Hangzhou and the Department of Psychiatry of the Second People's Hospital of Hangzhou. MDD was diagnosed by a senior psychiatrist following the Brief International Neuropsychiatric Interview (MINI), and the severity was assessed by the 24-item version of the Hamilton Depression Scale (HAMD-24). Patient enrollment criteria and the recruitment process were described in detail in [20]. The MDD group received 3 months of SSRI (e.g., fluoxetine, paroxetine, sertraline, etc.) treatment. Eventually, a total of 74 participants received two sMRI scans before and after treatment, including 39 patients with MDD (28 females and 11 males, mean age 27.31 ± 8.36 years) and 35 patients with HC (24 females and 11 males, mean age 28 ± 11.15 years). The study was approved by the Ethics Committee of the Institute of Psychological Science, Hangzhou Normal University. All participants obtained written informed consent before participating in the study procedures.

For Jinan Dataset, a total of 54 adolescent patients with MDD (MDD group) were recruited. Patients were recruited from the Mental Health Center of Shandong Province. MDD was diagnosed by two clinical psychiatrists as first-episode depression according to the Diagnostic and Statistical Manual of Mental Disorders-5 (DSM-V) criteria, with severity rated on the Children's Depression Inventory (CDI). Exclusion criteria for all participants included: (1) contraindications to MRI scanning (e.g., metal implants or claustrophobia); and (2) no history of other major psychiatric or neurological disorders, head injuries, or substance abuse. The MDD group received two months of SSRI treatment. Fifty-four (44 females, 10 males, mean age 15.78 ± 1.79 years) adolescents with MDD underwent two pre- and post-treatment sMRI Scanning. The study was approved by the Ethics Committee of the School of Psychology, Shandong Normal University. Written informed consent was obtained from all participants and their parents before they participated in the study procedures.

Image acquisition and pre-processing

For the Hangzhou Dataset, a GE 3T scanner (MR750, GE Medical Systems, Milwaukee, WI) was used to complete the structural scanning at the Center for Cognition and Brain Disorders of Hangzhou Normal University. The scanning parameters were as follows: TR = 9 ms, TE = 3.664 ms, FOV = 240×240 mm^2, matrix = 300×300,

Table 1 Demographic and clinical characteristics of subjects

	Hangzhou Dataset	Jinan Dataset
Gender (male/female)	28/11	26/8
Age	27.31(8.36)	15.39(1.26)
HAMD-24(pre/post)	27.87(8.02)/12.74(5.66)	/
CDI(pre/post)	/	25.61(7.92)/18.18(7.83)

HAMD-24: the 24-item version of the Hamilton Depression Scale, CDI: Children's Depression Inventory

flip angle = 13°, slice thickness = 0.8 mm, and acquisition time = 13 min 37 s.

For the Jinan Dataset, a cranial 3D-T1 structural image scan was completed at Qilu Hospital of Shandong University using a Siemens 3.0T Verio magnetic resonance instrument. A magnetized prepared rapid gradient echo (MPRAGE) scanning sequence was used. The scanning parameters were as follows: sagittal orientation, number of slices = 208, repetition time (TR) = 2400 ms, echo time (TE) = 2.19 ms, slice thickness = 1 mm, flip angle = 8°, field of view (FOV) = 224 × 224 mm^2, matrix size = 256 × 256, and total scan time = 11 min 18 s.

In this study, we focused on four structural indicators, including GMV, GMD, WMV, and WMD. The sMRI data were preprocessed using DPABI_V3.1 (Data Processing and Analysis Toolbox for Brain Imaging) [21]. The data were normalized to the Montreal Neurological Institute (MNI) standard space using the "Dartel + segment" method. During processing, the images were smoothed using a Gaussian kernel with a half-maximum (FWHM) of 8 mm.

Criteria for division into remission and non-remission groups

For the Hangzhou Dataset, based on the criteria proposed by Jiang et al. [19], we classified patients with ≥ 50% change in HAMD-24 scores before and after treatment and a posttest score of ≤ 10 into the remission group, whereas patients who did not meet this criterion were categorized into the non-remission group. There were 19 patients in the remission group and 20 in the non-remission group.

For the Jinan Dataset, since the patients were adolescents, a reduction of ≥ 30% in the total score of the scale was defined as a treatment remission according to the criteria of Emslie et al. [22]. In order to ensure a balanced dataset, we ultimately chose 14 remitters and 20 non-remitters from the 54 patients. Table 1 presents the demographic and clinical characteristics of subjects, including gender, age, and pre- and post-test scores for both groups.

Method

Our study follows a complete set of machine learning processes, as shown in Fig. 1. First is the data preprocessing stage, which has been described in detail in the previous section. Next is the data reading and processing, in which we normalized the image data, followed by resizing the image and dividing it into more smaller regions for subsequent feature extraction. This was followed by the feature extraction phase in which we defined three different similarity matrices to describe the relationships between the data. This was closely followed by feature selection, where we applied statistical testing methods to filter out features that contribute significantly to the performance of the classifier. Finally, we performed model training and evaluation, choosing appropriate performance metrics to assess the model's performance. This systematic process ensures the rigor and credibility of our study.

Reading and processing of data

The sMRI images were processed using commonly used libraries in python. After acquiring the image data, the minimum and maximum values of the image data were calculated, and the data were normalized so that the values ranged between 0 and 1. Next the normalized images were resampled and new pixel values were calculated using cubic spline interpolation method and finally the images were resized to a dimension of 40 × 40 × 40. Next, the resized images were split into four regions along each of the x-axis, y-axis, and z-axis, resulting in a total of 64 small regions. Since the image was resized to 40 × 40 × 40, the size of each small region on each axis was 40/4 = 10. The final extracted small regions were cubic in shape,

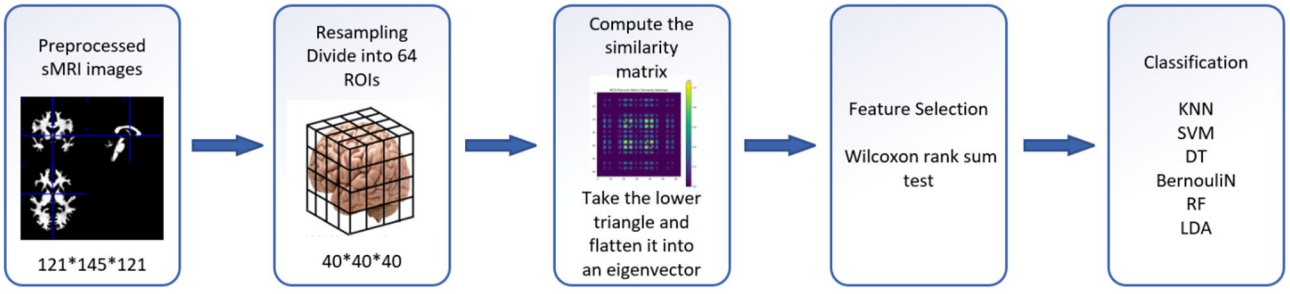

Fig. 1 Schematic diagram of machine learning workflow based on structural similarity features

with each side measuring 10 units in length. A total of 64 small regions covered the entire $40 \times 40 \times 40$ image space.

Feature extraction

First, for each small region, a matrix representing the degree of similarity between the small regions was constructed by calculating their eigenvectors. The similarity matrix was a two-dimensional array in which each element represented the degree of similarity between two small regions. For brain sMRI data, the data distribution was usually unknown because brain structures might have large individual variability among different individuals. The pixel values in an MRI image, after preprocessing such as segmentation, reflected the volume of the tissue or the density of the tissue, and these factors may varied greatly between individuals. Therefore, even for the same type of sMRI image, the data distribution might be different due to individual differences. In this case, we chose to use Pearson's correlation, Spearman's correlation, and Minimum Covariance Determinant (MCD), respectively, as similarity measures to gauge the correlation between two small regions. This step involved calculating the correlation between small regions, thus reflecting their degree of similarity in the feature space.

In this way, the above three methods were used to compute the similarity between each pair of small regions, and for each pair of small regions (i, j), the similarity of their eigenvectors was computed and filled into the corresponding positions of the similarity matrix to obtain a similarity matrix. Since the similarity matrix was symmetric, here, only the lower triangles were filled and the value of the upper triangles in the similarity matrix was set to zero. Through the above steps, the similarity matrix was populated, where each element represented the degree of similarity between the corresponding small regions. Finally, the populated similarity matrix was spread into a one-dimensional vector that served as the final feature representation for each sample.

Feature selection

In the feature selection section, we first executed the Wilcoxon rank sum test for two independent samples. For each feature, comparisons were made between the categories 'remission group' and 'non-remission group', respectively. Specifically, for each feature, we first extracted samples from the training set for the categories 'remission group' and 'non-remission group', and then performed the rank sum test on the samples from these two categories separately. We calculated p-values for each feature to measure the significance difference of that feature between the two categories ('remission group' and 'non-remission group'). We then selected the top 70 features that were significantly different based on the size of the p-value as the final set of features. This was done to exclude features that contribute less to the performance of the classifier, thus improving the generalization ability and performance of the model.

Model training and evaluation

To ensure the diversity and comparability of the models, we selected six classical models, namely K-Nearest Neighbor (KNN), Support Vector Machine (SVM), Decision Tree (DT), Bernoulli Naive Bayes (BernoulliNB), Random Forest (RF), and Linear Discriminant Analysis (LDA). These classifiers were chosen based on their well-documented effectiveness in handling medical imaging tasks and their ability to provide insights into various aspects of classification performance. The KNN algorithm is simple and intuitive, is insensitive to outliers, but computationally expensive when dealing with large datasets. The SVM performs well in high-dimensional spaces, can handle nonlinear data, and has strong generalization capabilities. DT is an interpretable classifier that handles nonlinear relationships. BernoulliNB, though simple, is effective for binary classification tasks and has been successfully applied in medical domains. RF improves accuracy and robustness by aggregating predictions from multiple decision trees. LDA is particularly useful for dimensionality reduction and feature extraction, making it suitable for the high-dimensional datasets that are typically encountered in neuroimaging studies.

Regarding similarity metrics, we selected Pearson, Spearman, and MCD due to their complementary strengths in capturing different aspects of the data. Pearson correlation is widely used and effective for linear relationships, while Spearman correlation can capture monotonic but not necessarily linear relationships, which can be important in brain imaging data where relationships may not always be linear. MCD is a robust estimator that handles outliers and non-normal distributions effectively, making it suitable for neuroimaging data with potential outlier or noise issues. These metrics were chosen to assess inter-brain similarity from multiple perspectives, thereby enhancing the robustness and interpretability of the results.

Ten-fold cross-validation was used for model evaluation. The evaluation metrics included Accuracy, Precision, Recall and F1-score and Area under the ROC curve (AUC).

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP} \tag{1}$$

$$Sensitivity/\mathrm{Re}\,call = \frac{TP}{TP + FN} \tag{2}$$

$$\mathrm{Pr}\,ecision = \frac{TP}{TP + FP} \tag{3}$$

Table 2 Group analysis results of two datasets

Dataset	Data Type	Cluster location	Cluster size	MNI			T value
				x	y	z	
Hangzhou Dataset	GMV	Precentral_R	74	34.5	1.5	37.5	-6.40
	GMD	Precentral_R	49	34.5	1.5	37.5	-6.24
		Cingulum_Ant_L	16	-1.5	25.5	30	-5.81
	WMV	Precentral_R	44	34.5	1.5	37.5	7.14
Jinan Dataset	GMD	Cerebelum_7b_R	19	12	-75	-42	-5.42
	WMV	Cerebelum_7b_R	13	13.5	-73.5	-40.5	5.06
	WMD	Cerebelum_7b_R	19	12	-75	-42	5.43
		Fusiform_L	11	-37.5	-78	-12	4.48

GMV: Gray matter volume, GMD: Gray matter density, WMV: White matter volume, WMD: White matter density, R: right, L: left

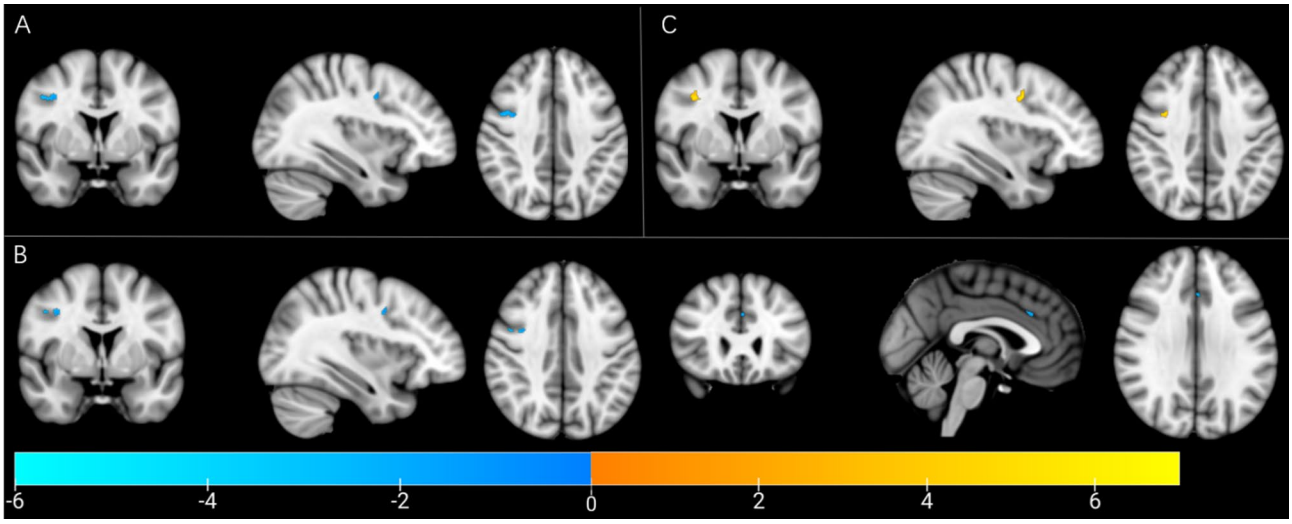

Fig. 2 Visualization of differences between responding and non-responding groups of Hangzhou Dataset in (**A**) gray matter volume (GMV), (**B**) gray matter density (GMD), and (**C**) white matter volume (WMV)

$$F1 = \frac{2 \times (\mathrm{Pr}\,ecision \times \mathrm{Re}\,call)}{\mathrm{Pr}\,ecision + \mathrm{Re}\,call} \quad (4)$$

In addition, we performed model testing using images after preprocess as features to compare the impact of feature selection schemes on classifier performance.

Replicability studies

In order to fully validate the generalizability of our proposed methodology to the adolescent population, we specifically selected the Jinan Dataset, a dataset that focuses on adolescents, and carried out a data process that is fully consistent with the main study. To ensure clarity, the Jinan Dataset was utilized as an external validation set to evaluate the generalizability of the proposed methodology, thereby providing further validation of its applicability to different populations.

Intergroup analysis

A two-sample t-test was performed on the structural characteristics (GMV, GMD, WMV, WMD) of the remission and non-remission groups using the Statistical Analysis function in the DPABI software [21]. During this analysis, age and gender of the subjects were corrected as covariates. The results of the group analysis were visualized using the Viewer function of the software. In the visualization process, we used the GRF correction method by setting the p-value of voxel to 0.001 and the p-value of cluster to 0.05, and chose the two-tailed test mode.

Results

Group analysis results

For the Hangzhou Dataset, the results of group analysis were shown in Table 2; Fig. 2. The results showed that the remission group had lower grey matter volume and density on the right side of the precentral gyrus (t = -6.40/-6.24, $p < 0.001$), while the WMV was significantly higher (t = 7.14, $p < 0.001$), compared to the non-remission group. In addition, the left grey matter density in the anterior cingulate and paracingulate gyrus was also significantly lower in the remission group (t = -5.81, $p < 0.001$).

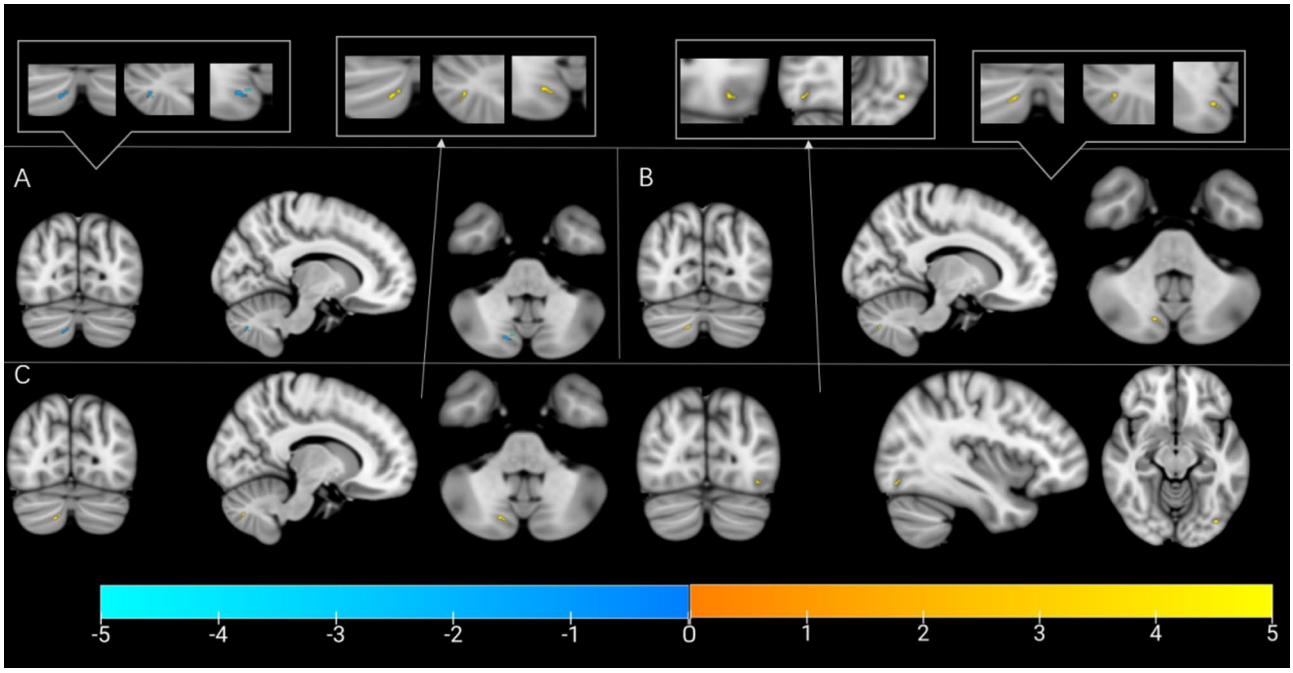

Fig. 3 Visualization of differences between responding and non-responding groups of Jinan Dataset in (**A**) gray matter density (GMD), (**B**) white matter volume (WMV) and (**C**) white matter density(WMD)

Table 3 Hangzhou dataset prediction results based on original image features (%)

	Classifier	Accuracy	Precision	Recall	F1 Score
GMV	KNN	52.50	75.00	25.00	31.67
	SVM	48.33	93.33	10.00	5.00
	DT	61.67	68.33	75.00	69.67
	BernoulliNB	48.33	90.00	0	0
	RF	60.83	61.67	55.00	54.67
	LDA	46.67	53.33	45.00	42.33
GMD	KNN	51.67	78.33	20.00	22.33
	SVM	60.00	36.67	45.00	39.67
	DT	60.00	36.67	45.00	39.67
	BernoulliNB	48.33	90.00	0	0
	RF	59.17	56.67	45.00	43.00
	LDA	56.67	68.33	45.00	47.33
WMV	KNN	61.67	76.67	45.00	40.00
	SVM	48.33	88.33	20.00	11.67
	DT	50.83	45.00	50.00	43.33
	BernoulliNB	48.33	90.00	0	0
	RF	65.83	71.67	55.00	54.67
	LDA	49.17	58.33	50.00	42.67
WMD	KNN	65.00	71.67	65.00	65.00
	SVM	48.33	90.00	0	0
	DT	48.33	46.67	50.00	46.33
	BernoulliNB	51.67	10.00	0	0
	RF	**75.00**	**55.00**	**50.00**	**47.00**
	LDA	60.83	83.33	35.00	40.67

Bolded text indicates the optimal result predicted by the model

The results and visualisation of the Jinan Dataset were presented in Table 2; Fig. 3, showing the differences in neuroimaging features between the remission and non-remission groups. We observed a significant decrease in the right cerebellar grey matter density in the remission group compared to the non-remission group ($t = -5.42$, $p < 0.001$), whereas the right cerebellar WMV and density increased significantly ($t = 5.06/5.43$, $p < 0.001$), and we observed a similarly significant increase in the WMD of the left fusiform gyrus in the remission group ($t = 4.48$, $p < 0.001$).

It is worth noting that we did not find significant group differences in other brain regions, which may be due to individual differences in neuroanatomy, limitations in sample size, or the specific parameters used in the statistical analysis.

Prediction results

Prediction results based on original image features

The accuracy, precision, recall, and F1 scores of all classifiers were at or slightly above the random level when raw features (including GMV, GMD, and WMV) were used. For example, when GMV was used, the accuracy of the classifiers ranged from 46.67 to 61.67%, while when GMD was used, the accuracy ranged from 48.33 to 60.00%. However, only when WMD was used as features, the RF classifier achieved an accuracy of 75%, which was higher than the random level, but its precision, recall, and F1 score were only 55%, 50%, and 47%, respectively (Table 3).

The lower performance of other classifiers with GMV, GMD, and WMV could be attributed to the relatively low discriminatory power of these features for treatment response prediction. This might reflect the need for more robust feature extraction methods or improved classifiers.

Prediction results based on brain similarity features

The prediction results of features extracted based on Pearson, Spearman, and MCD as similarity matrix calculation methods were shown in Table 4. Compared to the original structural image features, the metrics showed different degrees of improvement in the classification tasks of GMV, GMD, WMV, and WMD. WMV demonstrated the most significant performance in terms of classification accuracy, with an accuracy of 77.5%, followed by WMD, with an accuracy of 74.17%. These results were derived from the evaluations performed with the MCD matrix, LDA classifier, and Spearman matrix, RF classifier. The results of the analysis of the ROC curves, as shown in Fig. 4, also showed that the performance of these two features was good with AUC values of 0.88 and 0.76 respectively.

While the performance with similarity-based features was improved, some combinations still performed poorly. This may be due to the complexity of combining neuroimaging data with similarity matrices or the classifiers not being optimal for all feature types.

Results of the model's generalizability assessment in adolescent populations

The model generalizability study conducted for Jinan Dataset demonstrated that the grey matter density presented relatively good performance through the DT classifier in the original image classification task, as shown in Table 5, with an accuracy of 61.67% but a precision of up to 90.00%, with a recall and F1 score of 0.5500% and 0.5167, respectively. Other feature and model combinations also showed different levels of performance, with the accuracy, precision, recall and F1 score of each classifier varying under GMV, GMD, WMV and WMD features, further highlighting the importance of feature selection and classifier. Some feature-model combinations performed poorly, likely due to the complexity of the data and feature interactions. For example, the low performance in certain classifiers could be a result of the inherent difficulty in predicting treatment response using such features.

With the introduction of our method, all the metrics are significantly improved than when using the original images as features (Table 6), especially the WMD extracted through the MCD matrix, which reaches

Table 4 Prediction results (%) of Hangzhou dataset applying three similarity methods and six classifiers

		Accuracy/Precision/Recall/F1 Score		
		Pearson	Spearman	MCD
GMV	KNN	48.33/38.33/45.00/41.00	62.50/51.67/45.00/46.33	54.17/46.67/30.00/34.67
	SVM	38.33/31.67/45.00/36.33	45.83/33.33/45.00/37.67	71.67/71.67/65.00/64.67
	DT	46.67/38.33/30.00/32.33	51.67/46.67/40.00/39.67	56.67/55.00/50.00/48.67
	BernoulliNB	59.17/46.67/55.00/48.33	59.17/46.67/55.00/48.33	65.83/66.67/65.00/63.33
	RF	54.17/40.00/45.00/40.33	55.83/46.67/55.00/48.00	61.67/61.67/45.00/48.33
	LDA	36.67/30.00/40.00/36.67	36.67/30.00/40.00/33.67	61.67/55.00/60.00/55.33
GMD	KNN	54.17/40.00/40.00/37.00	43.33/20.00/25.00/21.67	66.67/66.67/55.00/56.33
	SVM	58.33/55.00/60.00/53.33	45.83/41.67/45.00/39.67	67.50/63.33/80.00/69.33
	DT	59.17/51.67/65.00/55.00	56.67/55.00/70.00/58.67	49.17/36.67/45.00/45.00
	BernoulliNB	54.17/45.00/50.00/43.67	54.17/45.00/50.00/43.67	62.50/63.33/65.00/62.33
	RF	50.83/46.67/50.00/44.67	43.33/35.00/35.00/31.67	62.50/61.67/60.00/58.00
	LDA	48.33/53.33/45.00/46.00	48.33/53.33/45.00/46.00	66.67/70.00/70.00/64.00
WMV	KNN	64.17/50.00/75.00/59.33	38.33/28.33/40.00/33.00	75.00/70.00/65.00/65.33
	SVM	49.17/38.33/55.00/44.67	61.67/48.33/65.00/55.00	65.00/56.67/65.00/60.00
	DT	64.17/68.33/75.00/66.33	61.67/63.33/70.00/62.67	65.83/56.67/65.00/58.67
	BernoulliNB	64.17/56.67/65.00/58.00	64.17/56.67/65.00/58.00	75.00/73.33/70.00/69.33
	RF	64.17/63.33/65.00/61.00	64.17/63.33/65.00/59.67	72.50/65.00/65.00/68.33
	LDA	56.67/50.00/55.00/49.00	56.67/50.00/55.00/49.00	**77.50/75.00/80.00/77.00**
WMD	KNN	61.67/58.33/50.00/50.67	49.17/51.67/50.00/48.00	67.50/65.00/70.00/65.67
	SVM	65.83/60.00/75.00/65.67	74.17/66.67/65.00/33.33	70.00/73.00/65.00/66.00
	DT	65.83/66.67/55.00/56.33	60.83/60.00/55.00/53.67	68.33/68.33/65.00/64.33
	BernoulliNB	60.83/60.00/55.00/55.00	60.83/60.00/55.00/55.00	70.00/68.33/65.00/64.33
	RF	71.67/61.67/65.00/61.67	74.17/71.67/70.00/68.33	69.17/70.00/70.00/65.67
	LDA	60.83/58.33/70.00/61.33	60.83/58.33/70.00/61.33	68.33/63.33/70.00/63.33

Bolded text indicates the optimal result predicted by the model

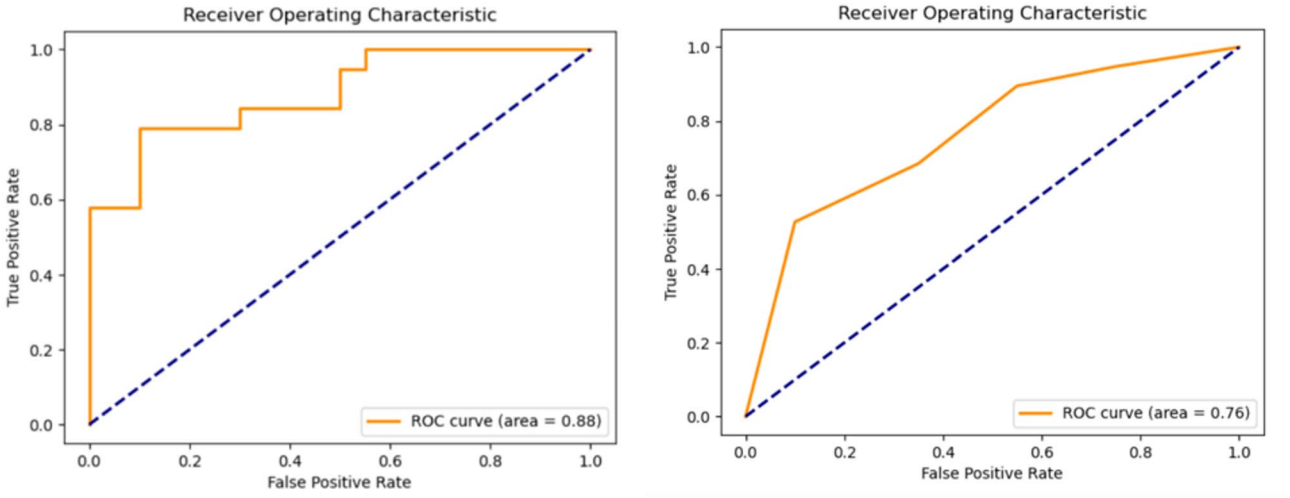

Fig. 4 ROC curves for classification based on similarity methods (left: white matter volume, MCD matrix, LDA classifier; right: white matter density, Spearman matrix, RF classifier)

Table 5 Jinan dataset prediction results based on original image features (%)

	Classifier	Accuracy	Precision	Recall	F1 Score
GMV	KNN	48.33	60.00	40.00	36.67
	SVM	36.67	63.33	60.00	33.33
	DT	55.00	63.33	10.00	05.00
	BernoulliNB	36.67	73.33	40.00	20.00
	RF	38.33	48.33	40.00	24.67
	LDA	53.33	70.00	45.00	40.00
GMD	KNN	36.67	50.00	15.00	16.67
	SVM	26.67	53.33	40.00	20.00
	DT	**61.67**	**90.00**	**55.00**	**51.67**
	BernoulliNB	31.67	63.33	40.00	20.00
	RF	28.33	61.67	75.00	61.67
	LDA	31.67	58.33	50.00	26.67
WMV	KNN	50.00	66.67	40.00	31.33
	SVM	36.67	63.33	60.00	33.33
	DT	30.00	48.33	20.00	11.67
	BernoulliNB	36.67	63.33	60.00	33.33
	RF	41.67	51.67	20.00	14.67
	LDA	45.00	55.00	55.00	41.67
WMD	KNN	41.67	51.67	60.00	50.00
	SVM	36.67	68.33	50.00	26.67
	DT	35.00	58.33	40.00	36.67
	BernoulliNB	31.67	58.33	50.00	26.67
	RF	41.67	53.33	35.00	25.00
	LDA	36.67	63.33	60.00	33.33

Bolded text indicates the optimal result predicted by the model

71.63% accuracy, 63.33% precision, 0.7000 recall, and 0.6333 F1 scores after applying the BernoulliNB classifier. These results signify that our method has good generalisability on different datasets, which provides a solid foundation and important support for further exploration of neuroimaging in the field of MDD treatment effect prediction.

Discussion

In this study, we introduced inter-brain similarity features extracted from structural MRI (sMRI) and combined them with machine learning techniques to predict treatment response in patients with MDD. Our results indicated that these inter-brain similarity features lead to more accurate predictions than raw image features, especially when GMV and WMD were used as predictors. Additionally, the method chosen for constructing the inter-brain similarity matrix significantly impacted the prediction outcomes, with MCD yielding the best results. To validate the robustness of these findings, we also tested the method using an adolescent dataset, further supporting its applicability.

Our findings were consistent with the growing body of literature highlighting the importance of brain structural features in predicting treatment response in MDD [12–15]. This study offered a significant contribution by integrating inter-brain similarity features from sMRI into the prediction of treatment response in MDD. These features captured subtle structural differences between brain regions, which could improve the accuracy of predictive models. Previous studies highlighted structural network disruptions in the brains of patients with MDD, particularly in regions such as the amygdala and medial prefrontal cortex, which played critical roles in depressive pathology [23]. Further studies had identified structural covariance patterns that differentiate patients with MDD from healthy controls and suggested that structural connectivity may be crucial in predicting treatment response [24]. Our findings aligned with these results, providing additional evidence that brain region similarity features held promise in predicting treatment outcomes in MDD.

For the extraction of brain region similarity features, we compared three different computational methods:

Table 6 Prediction results (%) of Jinan dataset applying three similar methods and six classifiers

		Accuracy/Precision/Recall/F1 Score		
		Pearson	Spearman	MCD
GMV	KNN	40.00/33.33/40.00/33.33	55.00/36.67/45.00/39.67	50.00/36.67/50.00/41.33
	SVM	38.33/31.67/35.00/31.33	31.67/33.33/40.00/33.33	46.67/35.00/55.00/41.33
	DT	45.00/40.00/45.00/40.00	45.00/43.33/55.00/45.00	55.00/51.67/45.00/46.33
	BernoulliNB	30.00/28.33/40.00/30.00	30.00/28.33/40.00/30.00	53.33/33.33/50.00/39.33
	RF	36.67/38.33/45.00/36.67	36.67/38.33/45.00/36.67	46.67/31.67/45.00/36.33
	LDA	41.67/30.00/30.00/28.33	41.67/30.00/30.00/28.33	53.33/45.00/70.00/53.00
GMD	KNN	46.67/46.67/50.00/41.67	38.33/20.00/25.00/19.67	43.33/28.33/30.00/26.67
	SVM	56.67/46.67/70.00/52.67	53.33/50.00/70.00/54.67	36.67/28.33/40.00/31.67
	DT	43.33/28.33/40.00/29.67	56.67/40.00/45.00/38.00	68.33/51.67/55.00/51.33
	BernoulliNB	51.67/50.00/60.00/51.33	51.67/50.00/60.00/50.00	46.67/40.00/45.00/39.67
	RF	50.00/40.00/60.00/44.33	53.33/43.33/60.00/46.33	65.00/60.00/60.00/58.33
	LDA	55.00/53.33/55.00/48.33	55.00/53.33/55.00/48.33	43.33/26.67/35.00/28.00
WMV	KNN	43.33/26.67/30.00/28.00	56.67/50.00/75.00/56.00	51.67/58.33/60.00/55.00
	SVM	60.00/50.00/55.00/51.67	38.33/30.00/40.00/31.33	45.00/46.67/65.00/50.00
	DT	51.67/53.33/50.00/46.67	55.00/46.67/55.00/46.33	56.67/38.33/40.00/37.67
	BernoulliNB	36.67/31.67/35.00/31.33	36.67/31.67/35.00/31.33	55.00/33.33/45.00/36.67
	RF	36.67/33.33/40.00/33.33	43.33/35.00/40.00/33.00	51.67/50.00/55.00/48.33
	LDA	40.00/28.33/50.00/34.67	40.00/28.33/50.00/34.67	46.67/33.33/45.00/35.00
WMD	KNN	53.33/55.00/75.00/59.67	48.33/41.63/50.00/44.67	60.00/53.33/65.00/55.00
	SVM	65.00/60.00/60.00/59.67	66.67/66.67/65.00/64.67	66.67/63.33/65.00/60.00
	DT	61.67/46.67/40.00/41.33	56.67/51.67/50.00/49.67	61.67/56.67/70.00/57.67
	BernoulliNB	60.00/56.67/70.00/61.33	60.00/56.67/70.00/61.33	**71.63/63.33/70.00/63.33**
	RF	53.33/50.00/75.00/58.00	56.67/51.67/60.00/53.00	58.33/58.33/60.00/53.33
	LDA	68.33/63.33/65.00/60.00	68.33/63.33/65.00/60.00	66.67/65.00/70.00/64.67

Bolded text indicates the optimal result predicted by the model

Pearson's correlation coefficient, Spearman's correlation coefficient, and MCD. Our results showed that MCD was the most effective method, particularly for analyzing high-dimensional and noisy data. However, MCD may underperformed in scenarios with extreme heterogeneity in datasets, where its robustness was compromised. This limitation should have been considered when applying MCD to datasets with significant variation across subjects. MCD, a robust covariance estimation technique, proved superior by accurately identifying structural changes associated with neurological disorders and improving the generalization of the model across different datasets [25–26]. This robustness made MCD an ideal tool for brain region similarity analysis, enabling it to capture the true structural characteristics of brain regions while effectively handling outliers.

In the group analysis, the Hangzhou Dataset revealed significant structural differences between patients in remission and those not in remission. Specifically, the remission group exhibited lower GMV and density in the right precentral gyrus but greater WMV. These findings aligned with previous studies that reported structural differences in the precentral gyrus and other brain regions related to MDD [7]. Additionally, we observed differences in the anterior cingulate gyrus and orbitofrontal cortex, consistent with other research suggesting these regions were critical in depressive pathology [8]. Furthermore, the Jinan Dataset confirmed these findings, with significant changes in GMV and WMV in the remission group, particularly in regions like the cerebellum and fusiform gyrus. These results further reinforced the potential of brain structure as a predictor of treatment response in MDD.

However, while these structural differences offered insight into potential treatment-related biomarkers, their clinical application remained to be fully explored. Tailoring treatment approaches based on these structural features could have potentially led to more pearsonalized care, though further research was required to determine how these features correlate with treatment efficacy in real-world settings.

Our study also demonstrated the generalizability of this method across different datasets and age groups. When comparing the predictive performance between adult and adolescent cohorts, we observed that WMV had the best predictive performance in the adult group (Hangzhou Dataset). In contrast, WMD was the strongest predictor in the adolescent group. This variability highlighted the importance of considering age-related differences in brain structure when using these features for prediction. Previous studies have suggested that MDD manifests differently in adolescents compared to adults, with distinct

pathophysiological mechanisms at play [27]. Structural differences in brain development between these age groups could explain why different brain features were predictive in each group. For example, studies have shown that the cerebral cortex exhibits significant developmental differences between adolescents and adults with MDD, which might impact treatment response [28]. This variability underscored the importance of designing prediction models that accounted for age-related structural differences in MDD.

Despite these promising results, several challenges remained. First, the sample size in our analysis was relatively small, which was a critical limitation. According to the PROBAST guideline, the ratio of sample size to predictors should be 20 or greater to ensure reliable model performance and generalizability. In our case, the small sample size could have led to underrepresentation of certain subpopulations, which might have skewed the results. This limitation emphasizes the need for larger, more diverse cohorts to enhance the robustness of our findings and ensure they are applicable to broader populations. Second, our dataset predominantly represents a Chinese population, which might limit the generalizability of the model to other populations with different genetic, cultural, and environmental backgrounds. The applicability of the model to diverse populations, including those from different countries and ethnic groups, needs further validation. To mitigate these challenges, future research could focus on increasing the sample size, diversifying the datasets, and addressing potential biases in model development. This will help refine the predictive capabilities of the model and improve its clinical utility across various populations. Additionally, incorporating more robust preprocessing and feature selection methods could help mitigate dataset biases and improve the generalizability of the model across different population groups. Furthermore, while the study acknowledges the longitudinal nature of depression, we did not explore the temporal evolution of brain structure in relation to treatment response. Future work could consider how brain structural changes over time may influence treatment outcomes, as this could provide more accurate predictions in clinical settings.

Conclusion

Our study is the first to use the inter-brain similarity of sMRI of the brain as a feature for machine learning to predict the treatment response of MDD. By integrating the neuroimaging features extracted by the similarity method and combining them with the machine learning technique, we succeeded in predicting the treatment response of patients with MDD, and the method not only possesses high accuracy but also shows good generalisability, which provides a new perspective to support individualized and precise MDD treatment strategies. The finding that brain structure and density features of depressed individuals exhibit significant variability across age highlights the need to adequately account for age differences in feature categorisation tasks. These results enrich our understanding of the relationship between neuroimaging features and treatment outcomes in MDD and provide an important reference for future clinical practice and research.

Abbreviations

MRI	Magnetic resonance imaging
MDD	Major depressive disorder
HC	Healthy controls
GMD	Gray matter density
GMV	Gray matter volume
WMD	White matter density
WMV	White matter volume
SVM	Support vector machine
sMRI	Structural magnetic resonance imaging
KNN	K-Nearest Neighbor
DT	Decision Tree
BernoulliNB	Bernoulli Naive Bayes
RF	Random Forest
LDA	and Linear Discriminant Analysis
L	Left
R	Right
MCD	Minimum Covariance Determinant
ECT	electroconvulsive therapy
MNI	Montreal Neurological Institute
HAMD-24	The 24-item version of the Hamilton Depression Scale
CDI	Children's Depression Inventory
DSM-V	Diagnostic and Statistical Manual of Mental Disorders-5
AUC	Area under the ROC curve
SSRI	Selective Serotonin Reuptake Inhibitors

Acknowledgements

We thank all participants for their willingness to participate in the study.

Author contributions

SS and SW conceived the work, SS, SW, KW, DW，ZD and WZ conducted the analyses. SW drafted the original manuscript. SS, KW, DW, JG, LZ and WZ revised the manuscript. SS, DW, KW and YW contributed to the acquisition of data. All authors approved the submission of the manuscript.

Funding

This research was supported by the National Natural Science Foundation of China (32000760), China Postdoctoral Science Foundation Funded Project (2023T160397), and the Youth Innovation Team in Universities of Shandong Province (2022KJ252). This research was also supported by the Natural Science Foundation of Shandong Province, China (ZR2021MF043), the Humanities and Social Sciences Foundation for the Youth Scholars of the Ministry of Education of China (18YJCZH149). This research was also approved by the Health Commission of Shandong Province, China (202103091036), and was financially supported by the Shandong Mental Health Center of Shandong University. Financially, the funding sources had no involvement in study design; in the collection, analysis and interpretation of data; in the writing of the report; in preparation of the article; nor in the decision to submit the article for publication.

Data availability

The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

For Hangzhou Dataset, this study was approved by the ethics committee of the Institutes of Psychological Sciences, Hangzhou Normal University. All methods were performed in accordance with the relevant guidelines and regulations. All patients' legally authorized representatives and controls provided written informed consent before participation in the study procedures. For Jinan Dataset, the study was approved by the Ethics Committee of the School of Psychology, Shandong Normal University. Written informed consent was obtained from all participants and their parents before participating in the study procedures. All methods were performed in accordance with the Declaration of Helsinki.

Consent for publication

not applicable.

Conflict of interest

The authors declare no conflicts of interest.

Received: 4 January 2025 / Accepted: 7 May 2025
Published online: 26 May 2025

References

1. Ferrari AJ, Charlson FJ, Norman RE, Patten SB, Freedman G, Murray CJ, Vos T, Whiteford HA. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. PLoS Med. 2013;10(11):e1001547. https://doi.org/10.1371/journal.pmed.1001547
2. Bartlett EA, DeLorenzo C, Sharma P, Yang J, Zhang M, Petkova E, Weissman M, Parsey RV. Pretreatment and early-treatment cortical thickness is associated with SSRI treatment response in major depressive disorder. Neuropsychopharmacology. 2018;43(11):2221–30. https://doi.org/10.1038/s41386-018-0122-9
3. Al-Harbi KS. Treatment-resistant depression: therapeutic trends, challenges, and future directions. Patient Prefer Adherence. 2012;6:369–88. https://doi.org/10.2147/ppa.S29716
4. Papakostas GI, Fava M, Thase ME. Treatment of SSRI-resistant depression: a meta-analysis comparing within- versus across-class switches. Biol Psychiatry. 2008;63(7):699–704. https://doi.org/10.1016/j.biopsych.2007.08.010
5. Thase ME, Entsuah AR, Rudolph RL. Remission rates during treatment with Venlafaxine or selective serotonin reuptake inhibitors. Br J Psychiatry. 2001;178:234–41. https://doi.org/10.1192/bjp.178.3.234
6. McGrath PJ, Stewart JW, Fava M, Trivedi MH, Wisniewski SR, Nierenberg AA, Thase ME, Rush AJ. Tranylcypromine versus Venlafaxine plus Mirtazapine following three failed antidepressant medication trials for depression: a STAR*D report. Am J Psychiatry. 2006;163(9):1531–41. https://doi.org/10.1176/ajp.2006.163.9.1531. quiz 666.
7. Lorenzetti V, Allen NB, Fornito A, Yücel M. Structural brain abnormalities in major depressive disorder: a selective review of recent MRI studies. J Affect Disord. 2009;117(1–2):1–17. https://doi.org/10.1016/j.jad.2008.11.021
8. Lai CH, Hsu YY, Wu YT. First episode drug-naïve major depressive disorder with panic disorder: Gray matter deficits in limbic and default network structures. Eur Neuropsychopharmacol. 2010;20(10):676–82. https://doi.org/10.1016/j.euroneuro.2010.06.002
9. Carceller-Sindreu M, Serra-Blasco M, de Diego-Adeliño J, Vives-Gilabert Y, Vicent-Gil M, Via E, Puigdemont D, Portella MJ. Altered white matter volumes in first-episode depression: evidence from cross-sectional and longitudinal voxel-based analyses. J Affect Disord. 2019;245:971–77. https://doi.org/10.1016/j.jad.2018.11.085
10. Wolf RC, Nolte HM, Hirjak D, Hofer S, Seidl U, Depping MS, Stieltjes B, Thomann PA. Structural network changes in patients with major depression and schizophrenia treated with electroconvulsive therapy. Eur Neuropsychopharmacol. 2016;26(9):1465–74. https://doi.org/10.1016/j.euroneuro.2016.06.008
11. Phillips JL, Batten LA, Tremblay P, Aldosary F, Blier PA, Prospective. Longitudinal study of the effect of remission on cortical thickness and hippocampal volume in patients with Treatment-Resistant depression. Int J Neuropsychopharmacol. 2015;18(8). https://doi.org/10.1093/ijnp/pyv037
12. Gerlach AR, Karim HT, Peciña M, Ajilore O, Taylor WD, Butters MA, Andreescu C. MRI predictors of pharmacotherapy response in major depressive disorder. Neuroimage Clin. 2022;36:103157. https://doi.org/10.1016/j.nicl.2022.103157
13. Liu X, Hou Z, Yin Y, Xie C, Zhang H, Zhang H, Zhang Z, Yuan Y. Decreased cortical thickness of left premotor cortex as a treatment predictor in major depressive disorder. Brain Imaging Behav. 2021;15(3):1420–26. https://doi.org/10.1007/s11682-020-00341-3
14. Paolini M, Harrington Y, Colombo F, Bettonagli V, Poletti S, Carminati M, Colombo C, Benedetti F, Zanardi R. Hippocampal and parahippocampal volume and function predict antidepressant response in patients with major depression: A multimodal neuroimaging study. J Psychopharmacol. 2023;37(11):1070–81. https://doi.org/10.1177/02698811231190859
15. Zarate-Garza PP, Ortega-Balderas JA, de la Ontiveros-Sanchez JA, Lugo-Guillen RA, Marfil-Rivera A, Quiroga-Garza A, Guzman-Lopez S, Elizondo-Omaña RE. Hippocampal volume as treatment predictor in antidepressant Naïve patients with major depressive disorder. J Psychiatr Res. 2021;140:323–28. https://doi.org/10.1016/j.jpsychires.2021.06.008
16. Motter JN, Lee S, Sneed JR, Doraiswamy PM, Pelton GH, Petrella JR, Devanand DP. Cortical thickness predicts remission of depression with antidepressants in patients with late-life depression and cognitive impairment. J Affect Disord. 2021;295:438–45. https://doi.org/10.1016/j.jad.2021.08.062
17. Korgaonkar MS, Rekshan W, Gordon E, Rush AJ, Williams LM, Blasey C, Grieve SM. Magnetic resonance imaging measures of brain structure to predict antidepressant treatment outcome in major depressive disorder. EBioMedicine. 2015;2(1):37–45. https://doi.org/10.1016/j.ebiom.2014.12.002
18. Gärtner M, Ghisu E, Herrera-Melendez AL, Koslowski M, Aust S, Asbach P, Otte C, Bajbouj M. Using routine MRI data of depressed patients to predict individual responses to electroconvulsive therapy. Exp Neurol. 2021;335:113505. https://doi.org/10.1016/j.expneurol.2020.113505
19. Jiang R, Abbott CC, Jiang T, Du Y, Espinoza R, Narr KL, Wade B, Calhoun VD. SMRI biomarkers predict electroconvulsive treatment outcomes: accuracy with independent data sets. Neuropsychopharmacology. 2018;43(5):1078–87. https://doi.org/10.1038/npp.2017.165
20. Li H, Song S, Wang D, Tan Z, Lian Z, Wang Y, Zhou X, Pan C. Individualized diagnosis of major depressive disorder via multivariate pattern analysis of thalamic sMRI features. BMC Psychiatry. 2021;21(1):415. https://doi.org/10.1186/s12888-021-03414-9
21. Yan CG, Wang XD, Zuo XN, Zang YF. DPABI: data processing & analysis for (resting-state). Brain Imaging Neuroinformatics. 2016;14(3):339–51. https://doi.org/10.1007/s12021-016-9299-4
22. Emslie GJ, Heiligenstein JH, Wagner KD, Hoog SL, Ernest DE, Brown E, Nilsson M, Jacobson JG. Fluoxetine for acute treatment of depression in children and adolescents: a placebo-controlled, randomized clinical trial. J Am Acad Child Adolesc Psychiatry. 2002;41(10):1205–15. https://doi.org/10.1097/00004583-200210000-00010
23. Korgaonkar MS, Fornito A, Williams LM, Grieve SM. Abnormal structural networks characterize major depressive disorder: a connectome analysis. Biol Psychiatry. 2014;76(7):567–74. https://doi.org/10.1016/j.biopsych.2014.02.018
24. Yang X, Kumar P, Nickerson LD, Du Y, Wang M, Chen Y, Li T, Pizzagalli DA, Ma X. Identifying subgroups of major depressive disorder using brain structural covariance networks and mapping of associated clinical and cognitive variables. Biol Psychiatry Glob Open Sci. 2021;1(2):135–45. https://doi.org/10.1016/j.bpsgos.2021.04.006
25. Hubert M, Debruyne MJWC. Minimum Covariance Determinant. 2010;2(1):36–43. https://doi.org/10.1002/wics.61
26. Mousavian M, Chen J, Traylor Z, Greening SJJIIS. Depression detection from sMRI and rs-fMRI images using machine learning. 2021;57:395–418. https://doi.org/10.1007/s10844-021-00653-w
27. Rice F, Riglin L, Lomax T, Souter E, Potter R, Smith DJ, Thapar AK, Thapar A. Adolescent and adult differences in major depression symptom profiles. J Affect Disord. 2019;243:175–81. https://doi.org/10.1016/j.jad.2018.09.015
28. Schmaal L, Hibar DP, Sämann PG, Hall GB, Baune BT, Jahanshad N, Cheung JW, Veltman DJ. Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the ENIGMA major depressive disorder working group. Mol Psychiatry. 2017;22(6):900–09. https://doi.org/10.1038/mp.2016.60

Publisher's note