frontiers | Research Topics

Clinical application of machine learning methods in psychiatric disorders

Edited by
Xiaozheng Liu, Zhi Xu, Weikai Li and Zhen Zhou

Published in
Frontiers in Psychiatry

ISSN 1664-8714
ISBN 978-2-8325-2698-9
DOI 10.3389/978-2-8325-2698-9

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Clinical application of machine learning methods in psychiatric disorders

Topic editors

Xiaozheng Liu — Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, China

Zhi Xu — Southeast University, China

Weikai Li — Chongqing Jiaotong University, China

Zhen Zhou — University of Pennsylvania, United States

Topic coordinator

Xize Jia — Hangzhou Normal University, China

Citation

Liu, X., Xu, Z., Li, W., Zhou, Z., eds. (2023). *Clinical application of machine learning methods in psychiatric disorders*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-2698-9

Table of contents

frontiers | Frontiers in Psychiatry

TYPE Editorial
PUBLISHED 31 May 2023
DOI 10.3389/fpsyt.2023.1209615

OPEN ACCESS

EDITED AND REVIEWED BY
Francisco Rodrigues,
University of São Paulo, Brazil

*CORRESPONDENCE
Xiaozheng Liu
✉ lxz_2088@hotmail.com

RECEIVED 21 April 2023
ACCEPTED 12 May 2023
PUBLISHED 31 May 2023

CITATION
Liu X, Xu Z, Li W, Zhou Z and Jia X (2023) Editorial: Clinical application of machine learning methods in psychiatric disorders. *Front. Psychiatry* 14:1209615. doi: 10.3389/fpsyt.2023.1209615

Editorial: Clinical application of machine learning methods in psychiatric disorders

Xiaozheng Liu[1]*, Zhi Xu[2], Weikai Li[3], Zhen Zhou[4] and Xize Jia[5]

[1]Department of Radiology, Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, Wenzhou, China, [2]Department of Psychosomatics and Psychiatry, Zhongda Hospital, School of Medicine, Southeast University, Nanjing, China, [3]School of Mathematics and Statistics, Chongqing Jiaotong University, Chongqing, China, [4]Department of Radiology, Center for Biomedical Image Computing and Analytics, University of Pennsylvania, Philadelphia, PA, United States, [5]School of Psychology, Zhejiang Normal University, Hangzhou, China

KEYWORDS

machine learning, psychiatric disorders, clinical application, natural language processing, functional magnetic resonance imaging (fMRI)

Editorial on the Research Topic
Clinical application of machine learning methods in psychiatric disorders

Introduction

With the increasing prevalence of psychiatric disorders, there has been a significant increase in the demand for effective psychiatric disorder treatments in recent years. However, the complexity of neuronal degeneration and the heterogeneity of patients make early diagnosis and treatment of these disorders difficult. To meet these challenges, scientists, clinicians, and patients can benefit from the application of machine learning theories and algorithms. Machine learning, which includes methods for feature extraction, selection, and classification, has demonstrated significant benefits in the pathological analysis of psychiatric disorders (1). These methods can learn features from brain neuroimaging data and adapt to data variation, thereby improving the reliability, performance, and accuracy of disease-specific diagnostic systems. Furthermore, machine learning can accurately assess the conditions of patients. With the use of cutting-edge machine learning algorithms, clinical diagnosis, and clinical interventions, along with clinical neuroimaging data of the brain, this Research Topic aims to incorporate theoretical and technological innovations and assess the performance of machine learning in clinical studies on psychiatric disorders.

Machine learning methods in psychiatric disorders

Deep learning-based natural language processing techniques were applied to assess depressive symptoms in clinical interviews. The F1 score (a measure of model performance, harmonic mean of accuracy, and recall) was 0.719 when classifying the four-level severity of depression, and 0.890 when identifying the presence of depressive symptoms (Li et al.). Multidimensional speech feature diagnosis and evaluation system (MSFDA) combining multidimensional speech features and deep learning in the auxiliary diagnosis of major depressive disorder in children and adolescents. The sensitivity (92.73% vs. 76.36%)

and specificity (90.91% vs. 85.45%) of the MSFDA system were significantly higher than those of HAMD-24. The area under the curve of the MSFDA system is also higher than that of HAMD-24. The difference between the two groups was statistically significant ($p < 0.05$), and the diagnostic accuracy was higher in both groups. In addition, the diagnostic efficiency of the MSFDA system is higher than that of HAMD-24 in terms of Youden index, diagnostic accuracy, likelihood ratio, diagnostic odds ratio and predictive value (Luo et al.).

The combination of artificial intelligence and imaging data to guide clinical diagnosis and intervention is a major direction for future clinical research. Venkatapathy et al. present an integrated model for resting-state functional MRI data analysis and graph convolution networks based on graph theory. For the classification of patients with major depressive symptoms and healthy controls, the ensemble models achieved 71.18% upsampling accuracy and 70.24% downsampling accuracy. When comparing patients with first-episode major depressive symptoms to those with recurrent major depressive symptoms, the accuracy of upsampling was 77.78% and the accuracy of downsampling was 71.96% (Venkatapathy et al.). A multimodal MRI image-based imagingomics study predicted prognostic outcome of stroke with an accuracy of 0.831, sensitivity of 0.739, specificity of 0.902, F1 score of 0.788 and area under the curve of 0.902 (Yu et al.). Resting-state-based amplitude of low-frequency fluctuation and support vector machine models help distinguish schizophrenia patients from healthy controls (Gao et al.).

Peripheral blood is easy to obtain and less invasive, making it an ideal specimen for clinical trials. Lymphocyte subpopulation-based features help distinguish bipolar depression from major depressive disorder with an accuracy of >90% (Su et al.). The accuracy of the peripheral non-enzymatic antioxidant combined with the xGboost model for differentiating bipolar disorder from major depressive disorder was 0.849 and for differentiating bipolar disorder with depressive episodes from major depressive disorder was 0.899 (Gong et al.).

Conclusion

Each paper focuses on a different but equally important aspect of clinical application of machine learning methods in psychiatric disorders. We believe that by presenting and highlighting the latest novel and emergent machine learning technologies, implementations, and applications relating to psychiatric disorders, we will raise awareness in the scientific community. Finally, we'd like to thank all of the authors who contributed to this Research Topic with their research. We would also like to thank the many experts in the field who participated in the review process and offered helpful suggestions to the authors to improve the articles' contents and presentations.

Author contributions

XL were major contributors in writing the manuscript. All authors read and approved the final manuscript.

Conflict of interest

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.

Publisher's note

References

1. Shehab M, Abualigah L, Shambour Q, Abu-Hashem MA, Shambour MKY, Alsalibi AI, et al. Machine learning in medical applications: a review of state-of-the-art methods. *Comput Biol Med.* (2022) 145:105458. doi: 10.1016/j.compbiomed.2022.105458

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 25 August 2022
DOI 10.3389/fpsyt.2022.949512

OPEN ACCESS

EDITED BY
Yu-Feng Zang,
Hangzhou Normal University, China

REVIEWED BY
Ying Wang,
First Affiliated Hospital of Jinan University, China
Zhikun Zhang,
Guangxi Medical University, China

*CORRESPONDENCE
Gang Wang
532614206@qq.com
Yi Li
psyyee@gmail.com
Gaohua Wang
wanggaohua64@126.com

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Neuroimaging and Stimulation, a section of the journal Frontiers in Psychiatry

RECEIVED 23 May 2022
ACCEPTED 19 July 2022
PUBLISHED 25 August 2022

CITATION
Gao Y, Tong X, Hu J, Huang H, Guo T, Wang G, Li Y and Wang G (2022) Decreased resting-state neural signal in the left angular gyrus as a potential neuroimaging biomarker of schizophrenia: An amplitude of low-frequency fluctuation and support vector machine analysis. *Front. Psychiatry* 13:949512. doi: 10.3389/fpsyt.2022.949512

Decreased resting-state neural signal in the left angular gyrus as a potential neuroimaging biomarker of schizophrenia: An amplitude of low-frequency fluctuation and support vector machine analysis

Yujun Gao[1†], Xin Tong[2,3†], Jianxiu Hu[3], Hanjun Huang[3], Tian Guo[3], Gang Wang[3*], Yi Li[3*] and Gaohua Wang[1*]

[1]Department of Psychiatry, Renmin Hospital of Wuhan University, Wuhan, China, [2]School of Mental Health and Psychological Science, Anhui Medical University, Heifei, China, [3]Wuhan Mental Health Center, Wuhan, China

Objective: Schizophrenia (SCH) is primarily diagnosed based on specific clinical symptoms, with the lack of any objective SCH-related biomarkers often resulting in patient misdiagnosis and the underdiagnosis of this condition. This study was developed to assess the utility of amplitude of low-frequency fluctuation (ALFF) values analyzed *via* support vector machine (SVM) methods as a means of diagnosing SCH.

Methods: In total, 131 SCH patients and 128 age- and gender-matched healthy control (HC) individuals underwent resting-state functional magnetic resonance imaging (rs-fMRI), with the resultant data then being analyzed using ALFF values and SVM methods.

Results: Relative to HC individuals, patients with SCH exhibited ALFF reductions in the left angular gyrus (AG), fusiform gyrus, anterior cingulate cortex (ACC), right cerebellum, bilateral middle temporal gyrus (MTG), and precuneus (PCu) regions. No SCH patient brain regions exhibited significant increases in ALFF relative to HC individuals. SVM results indicated that reductions in ALFF values in the bilateral PCu can be used to effectively differentiate between SCH patients and HCs with respective accuracy, sensitivity, and specificity values of 73.36, 91.60, and 54.69%.

Conclusion: These data indicate that SCH patients may exhibit characteristic reductions in regional brain activity, with decreased ALFF values of the bilateral PCu potentially offering value as a candidate biomarker capable of distinguishing between SCH patients and HCs.

KEYWORDS

amplitude of low-frequency fluctuation, schizophrenia, resting-state fMRI, support vector machine, imaging biomarker

Introduction

Schizophrenia (SCH) is a psychological disorder characterized by progressive changes in brain function that result in symptoms including reduced social function, decreased motivation and emotion, hallucinations, and delusions (1). Currently, SCH diagnoses are primarily made based on the symptoms and clinical signs that patients exhibit (2). As other psychiatric conditions including bipolar disorder and major depressive disorder (MDD) can exhibit symptoms similar to those of SCH, this can often result in patient misdiagnosis, underscoring the need for the establishment of objective biomarkers capable of aiding in SCH diagnostic efforts.

Resting-state functional magnetic resonance imaging (rs-fMRI) has emerged as a powerful tool that offers a potential means of identifying novel, sensitive biomarker signatures associated with specific brain disorders (3–6). A growing body of evidence suggests that abnormal changes in patients with SCH are primarily found in the striatum (7), temporal lobe (8), default-mode network (DMN) (9), and frontoparietal network (10), although the specific nature of these changes has varied across studies, with some reporting functional signal increases (11–13), decreases (14), or both (15). These discrepant findings may be attributable to the differences in sample size, disease course, and the analytical methods employed (16,17). Researchers have focused on these regional brain abnormalities when seeking to define biomarkers of SCH (17,18). Li et al., for example, utilized volumetric decreases in the left insula as a potential diagnostic endophenotype for SCH (19), while others have reported early decreases in global-brain functional connectivity in the bilateral anterior cingulate cortex (ACC) as a promising predictor of SCH patient therapeutic outcomes (20). In a recent article, Li et al. established a novel hypothesis-driven neuroimaging biomarker of SCH through a comparison of these patients and healthy control (HC) individuals, achieving > 80% accuracy (17). However, their developed biomarker necessitated the integration of several functional indicators, making it impractical for routine or urgent clinical use in a diagnostic or therapeutic setting. Ideally, a biomarker of SCH should be readily obtained, non-invasive, and associated with a high degree of diagnostic accuracy, although no such biomarkers have yet been identified despite decades of intensive research (21).

The amplitude of low-frequency fluctuations (ALFF) in the BOLD signal measured during rs-fMRI analyses can provide insight into spontaneous brain functional activity, and it is thus commonly used when evaluating patients with diseases including cervical spondylotic myelopathy (22), SCH (17), and depression (23). ALFF values have been shown to offer great promise as diagnostic biomarkers owing to their high degree of high temporal stability. Using an ALFF approach, SCH showed some abnormalities in spontaneous brain activity in parietal and occipital lobes (24). Furthermore, Kirino et al. (15) determined that SCH patients exhibited changes in spontaneous brain activity in two separate frequency bands. In addition, a large number of previous studies have found that the age of publication, the course of the disease, and the drug have some effects on brain function. Therefore, we selected patients with SCH with an earlier age of onset and a shorter course of disease as the study subjects to reduce confounding factors in this regard.

The application of support vector machine (SVM)-based artificial intelligence methods has been increasingly used to aid in predicting therapeutic outcomes or accurately diagnosing specific conditions (25,26). Multivariate pattern recognition-based SVM methods allow for the detection of patterns within a given dataset, and are well-suited to analyzing high-dimensional data in which there are more features than there are observations, as is common in experimental settings (27). SVM approaches enable optimal hyperplane separation in high-dimensional space, with samples closest to this hyperplane being defined as support vectors. When performing fMRI studies, SVM weights can be overlapped with the original brain space to generate a discriminative map by visually tracing the most important weights to the regions of the brain with the most discriminative value. SVM strategies have been shown to offer great clinical utility in the context of high-dimensional neuroimaging data-based decision-making (28,29). Here, rs-fMRI data from SCH patients were examined and ALFF values from abnormal regions of the brain were extracted and evaluated for their potential utility as neuroimaging biomarkers of SCH through the use of SVM methods. Together, the results of this study have the potential to aid in the more reliable and efficient diagnosis of SCH.

Materials and methods

Subjects

For this analysis, 131 patients diagnosed with SCH and 128 age- and gender-matched HC individuals were consecutively recruited from the inpatients or outpatients of the Wuhan Mental Health Center and Renmin Hospital of Wuhan University. Patients were diagnosed with SCH in accordance with the criteria established in the Diagnostic and Statistical Manual of Mental Disorders—Fourth Edition (DSM-IV). Prior to screening, all study participants were assessed with the Chinese MINI version of the Concise International Neuropsychiatric Interview, and two psychiatrists with different professional titles independently diagnosed all patients. While 5 participants (2 HCs, 3 SCH patients) exhibited excessive head movements during initial imaging, they were rescreened

on the same day such that no data needed to be excluded from this study. All enrolled SCH patients experienced auditory hallucinations, reference delusions, persecutory delusions, and were 18–64 years of age. Participants were excluded from this study if they exhibited a history of substance abuse, electroconvulsive therapy, left-handedness, neurological disease, or severe illness. Prior to scanning, SCH symptoms were assessed based upon the Chinese versions of the Positive and Negative Symptom Scale (PANSS) and the Repeatable Battery for the Assessment of Neuropsychological Status (RBANS) before scanning (30,31). After scanning, patients were successfully followed for 3 months and the diagnosis of SCH was confirmed. HC participants were recruited from universities and the community, were free of any history of severe medical or neuropsychiatric illnesses, and did not exhibit any family history of neuropsychiatric disease among first-degree relatives.

Magnetic resonance imaging scanning procedures

Philips Ingenia 3.0 T scanners in the Mental Health Center of Wuhan Affiliated with Huazhong University of Science and Technology were used to conduct rs-fMRI scanning for all study participants. Participants were directed to close their eyes, remain awake, and avoid thinking about anything in particular to the greatest extent possible. Functional images were captured using a gradient-echo—echo-planer imaging sequence to acquire data with the following settings: TR/TE = 2,000 ms/30 ms, thickness (mm) = 60, 35 slices, 64*64 element matrix, a flip angle of 78, 22.4 cm field of view, 3.5 mm slice thickness, 0.6 mm gap and 1 mm pitch, total scan duration"09:17.7." 3D_T1 scanning parameters: repetition time (TR) = 8.4 ms, echo time (TE) = 3.2 ms, slice thickness = 1 mm, slice spacing = 0 mm, Number of slice = 33, and field of view (FOV) = 256 × 256 cm, total scan duration"04:17.7."

Data processing

The pre-processing of rs-fMRI data was conducted in Matrix Laboratory (MATLAB) using Data Processing Assistant for Resting-State fMRI (DPARSF). The impact of initial signal instability on the resultant analyses was reduced by discarding the first 10 time points. Data were corrected for head movement and slice time. Any participants that exhibited > 2 mm maximum displacement in the x-, y-, or z-axis or > 2° maximum rotation underwent rescanning on the same day until meeting these criteria. Corrected imaging data were subjected to spatial normalization to the T1 imaging and standard Montreal Neurological Institute space. The resultant images were then resampled at $3 \times 3 \times 3$ mm^3, band-pass filtered (0.01–0.08 Hz), and linearly detrended. Spurious covariates were eliminated, including the signal from a region centered in the white matter and the signal from a ventricular seed-based region of interest. The resultant data were then smoothed using a Gaussian kernel of 6 mm full-width at half-maximum.

Amplitude of low-frequency fluctuations analyses

ALFF analyses were proposed by Jia et al. (32), and were conducted in MATLAB with the REST software (33). ALFF values were based on measurements of the rs-fMRI signal for each voxel, and were sensitive to the scale of the raw signal. Time series data for each voxel were subjected to fast Fourier transformation into the frequency domain, with the power spectrum then being calculated and subjected to square root transformation for each voxel. Average square root values were measured as the ALFF across the 0.01–0.08 Hz range for each voxel, with the ALFF then being calculated.

Statistical analyses

Differences in age, years of education, PANSS, and RBANS results were compared between SCH patients and HCs using two-sample *t*-tests, while chi-square tests were used to assess differences in gender distributions between these groups using SPSS 23.0. Age, gender, years of education, and frame-wise displacement were used as covariates. Correlations between abnormal ALFF values and clinical variables were assessed through Spearman's correlation analyses. $P < 0.05$ was the threshold of significance. Differences between groups were identified through a voxel-by-voxel analysis of covariance using individual whole-brain ALFF maps for these two groups. Results were thresholded at $P < 0.01$ and GRF-corrected *via* cluster-extent-based thresholding with a primary threshold of $P < 0.01$ in REST.

Classification and receiver operating characteristic analyses

SVM methods were implemented in MATLAB with a library for support vector machine (LIBSVM) software package. At first, two-sample *t*-tests were conducted to identify significant regions between the patient and control groups, and then SVM was used based on the ALFF values of the identified regions. A grid of parameters was evaluated using LIBSVM, with the accuracies of all parameter settings being acquired after which the highest cross-validation accuracy for these parameters was established (for more detailed procedures, see **Supplementary**

Material). ROC curves were used to analyze the ALFF values in abnormal brain regions.

Results

Participants

In total, this study enrolled 131 patients diagnosed with SCH and 128 age- and gender-matched HCs. Participant clinical and demographic data are summarized in Table 1, revealing no differences among these groups with respect to age, gender, or years of education.

Schizophrenia-related differences in amplitude of low-frequency fluctuations values

Initially, differences in ALFF values were compared between SCH patients and HC individuals using two-sample *t*-tests, revealing significant decreases in these values in the left angular gyrus (AG), ACC, fusiform, right cerebellum, bilateral precuneus (PCu), and middle temporal gyrus (MTG) in SCH patients relative to HCs (Figure 1 and Table 2).

Support vector machine results

Next, abnormal ALFF values in different regions of the brain were used to classify individuals in these two groups. The bilateral PCu and left AG in SCH patients were individually analyzed using the SVM method, revealing that reduced ALFF values in the bilateral PCu could readily differentiate between SCH patients and HCs with good accuracy (73.36%), specificity (54.69%), and sensitivity (91.60%) (Figures 2A, 3). Similarly, decreased ALFF values in the left AG were capable of discriminating between these two groups of patients with satisfactory accuracy (73.36%), specificity (52.34%), and sensitivity (93.89%) (Figures 2B, 3).

Receiver operating characteristic results

Next, ROC curve analyses were employed as a means of comparing the accuracy for analyses of region 6 and region 7, revealing that ALFF values for these two regions could be effectively applied to differentiate between SCH patients and HCs while achieving good sensitivity and specificity. Through this analysis, abnormal ALFF values in the bilateral PCu were found to be superior as a candidate biomarker for distinguishing between SCH patients and HCs (Figures 4, 5 and Table 3).

Correlations between amplitude of low-frequency fluctuations values and clinical variables

Abnormal ALFF values in the identified brain regions were not found to correlate with any clinical variables in this patient cohort.

Discussion

Here, reduced ALFF values were observed in the left AG, ACC, fusiform, right cerebellum, bilateral PCu, and MTG of SCH patients relative to HCs. When ALFF values for these abnormal brain regions were utilized as candidate biomarkers to differentiate between SCH patients and HCs *via* an SVM approach, decreased ALFF values in the bilateral PCu were able to discriminate between these two groups (accuracy: 73.36%, specificity: 54.69%, sensitivity: 91.60%), as were decreased ALFF values in the left AG (accuracy: 73.36%, specificity: 52.34%, sensitivity: 93.89%). Subsequent ROC analyses indicated that abnormal ALFF values in the bilateral PCu may offer value as an SCH-related neuroimaging biomarker, with an AUC of 72.47%.

The PCu is a critical component of the DMN, which corresponds to a series of functionally consistent networked brain regions (the medial prefrontal cortex, posterior cingulate cortex/PCu, medial parietal cortex, lateral parietal cortex, inferior parietal cortex, and cerebellum) that exhibit high activity levels at rest (34), with activity levels being reduced when the brain is engaged in non-specific attention task execution. A growing body of evidence supports a close relationship between the DMN and mental activity (35,36), with abnormal network homogeneity having been reported in the DMN of drug-naïve, first-episode adolescent SCH individuals (37). Additionally, different SCH subtypes have been found to exhibit significant differences in resting-state DMN activity (38). Researchers have also found that there are similar situations in different types of other mental diseases. For example, Chen et al. Found that bipolar depression and MDD have common abnormal brain activity (increased dynamic functional connectivity variability between the left dorsal rostral putamen and the left supplementary motor area, and between the right VRP and the right inferior parietal lobe), Had specific increased dynamic functional connectivity variability between the right dorsal caudal putamen and the left central gyrus compared with MDD and HCs (39). A multimodal meta-analysis of resting-state studies showed that bipolar disorder was characterized by hypo-connectivity within the default network, the affective network, and ventral attention network etc., and decreased gray matter volume in the insula, inferior frontal gyrus, and ACC (40). Recently, increased functional connectivity was observed between the right PCu and cerebellum, temporal lobe, and left superior

TABLE 1 Characteristics of the participants.

Demographic data	Patients (n = 131)	HCs (n = 128)	T (orx^2)	P-value
Gender (male/female)	131 (76/55)	128 (75/53)	0.35	0.85[a]
Age (years)	26.34 ± 10.39	30.33 ± 7.90	3.04	0.08[b]
Education (years)	9.50 ± 2.64	9.30 ± 2.61	0.61	0.34[a]
Onset (years)	19.42 ± 12.21			
Illness course (years)	6.92 ± 4.86			
REBANS	177.61 ± 38.60			
PANSS	89.63 ± 17.00			

[a]The p-value for gender distribution was obtained by chi-square test.
[b]The p-value were obtained by two sample t-tests.
HCs, healthy controls; REBANS, Repeatable Battery for the Assessment of Neuropsychological Status; PANSS, Positive and Negative Symptom Scale.

TABLE 2 Brain regions with abnormal ALFF in schizophrenia patients.

Cluster location	Peak X	(MNI) Y	Z	Cluster size (voxels)	Peak accuracy (%)	T-value
Patients < HCs						
Right cerebellum	18	−81	−27	58	69.11	−7.38
Left MTG	−51	−24	−15	399	70.27	−7.91
Right MTG	60	−39	−3	246	70.27	−8.25
Left fusiform	−30	−33	−21	41	69.88	−7.41
Left ACC	0	33	12	148	70.66	−8.2
Bilateral PCu	−6	−72	18	273	73.36	−8.75
Left AG	−51	−54	24	460	73.36	−8.15

MNI, montreal neurological institute; MTG, middle temporal gyrus; ACC, anterior cingulate cortex; PCu, Precuneus; AG, angular gyrus.

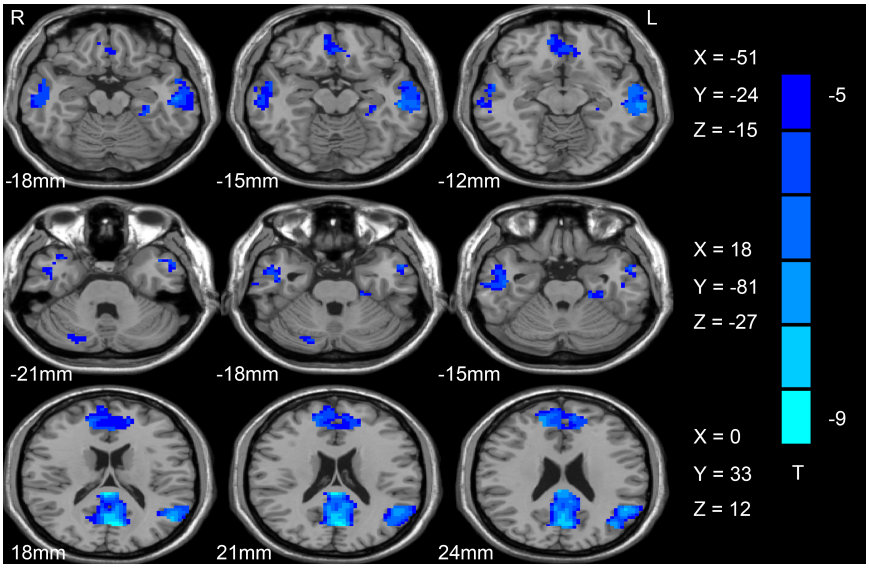


FIGURE 1
Differences in ALFF values between SCH patients and HC individuals. Reduced ALFF values are shown in blue, with the color bar representing t-values in the group analysis. ALFF, amplitude of low-frequency fluctuation; SCH, schizophrenia; HCs, healthy controls.

parietal lobule in non-depressed SCH individuals (41), while another study reported reduced ReHo in the right superior temporal gyrus, left middle frontal gyrus, PCu, and left central anterior gyrus in childhood- and adolescent-onset SCH, with auditory hallucination incidence being associated with abnormal activity in these regions (42). These inconsistent

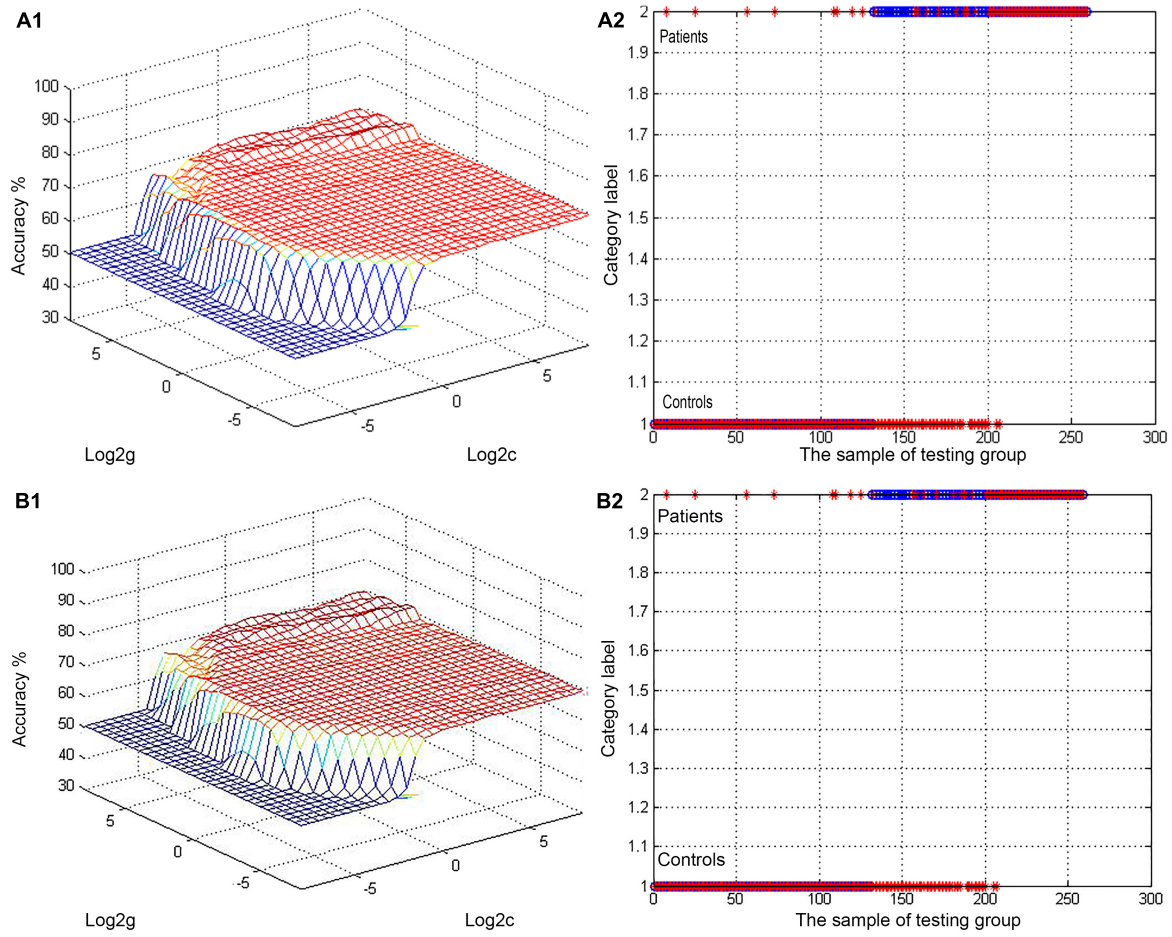


FIGURE 2
(A,B) Visualization of SVM classification based upon reduced ALFF values in the bilateral PCu and left AG as a means of differentiating between SCH patients and HCs. A1/2: 3D visualization of SVM with the most optimal parameters; B1/2: classification map of ALFF values for the bilateral PCu/left AG. SVM, support vector machine; ALFF, amplitude of low-frequency fluctuation; AG, angular gyrus; PCu, precuneus; SCH, schizophrenia; HCs, healthy controls.

results may be attributable to differences in sample size, disease course, or analytical approaches employed among studies. Furthermore, structural MRI findings have suggested that a gray matter volume reduction was evident in the bilateral PCu in SCH patients (43). In one recent meta-analysis, reduced ALFF values within the PCu were found to be reliably associated with memory and theory of mind in both first-episode and chronic SCH patients (44). Moreover, damage to the PCu can impact frontal lobe activity (45). These prior studies suggest that abnormal ALFF values in the PCu, medial superior frontal gyrus, MTG, and illness severity are likely to be correlated with one another. Even so, no such correlative relationships were observed in the present study. Abnormal ALFF values in the DMN may be attributable to the sample size in this study or the specific characteristics of this patient population. Moreover, this analysis revealed that reduced ALFF values in the bilateral PCu may serve as a promising biomarker for differentiating between SCH patients and HCs in an SVM analysis, yielding good accuracy (73.36%), specificity (54.69%), and sensitivity (91.60%).

The AG is located in the posterior of the inferior parietal lobe, and serves as a central hub for different subsystems (46). The AG plays a role in semantic processing, reading, and word comprehension, and is linked to higher memory scores (47). Over the course of SCH progression, patients exhibit increases in brain damage, with chronic SCH patients exhibiting the disruption of normal bilateral AG asymmetry and a reversal of the normal left > right asymmetry relative to HC individuals (48). In this analysis, decreased ALFF values were only observed in the left AG of analyzed SCH patients, potentially owing to the shorter disease course in these individuals. The AG is considered a component of the semantic-lexical network that includes the planum temporal

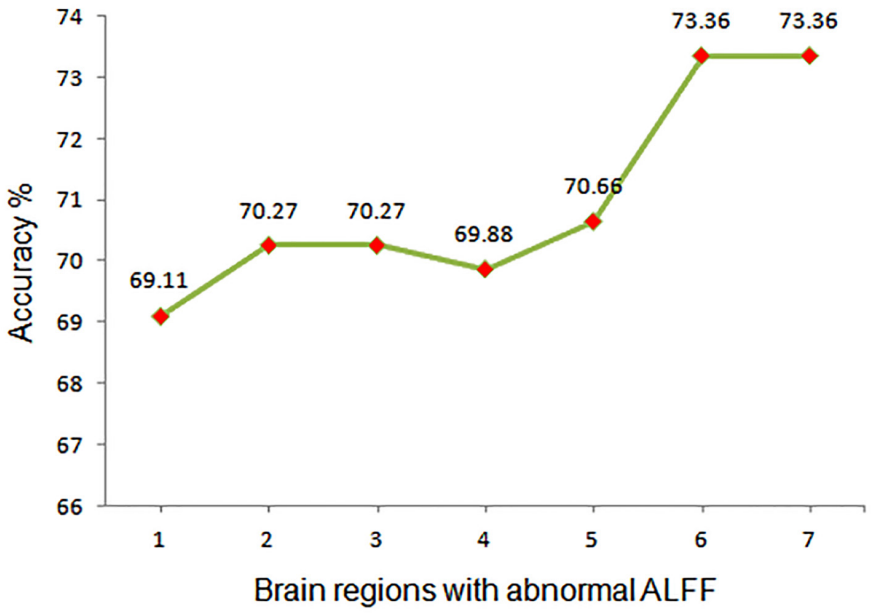


FIGURE 3
Assessment of the accuracy of utilizing abnormal ALFF values in different regions of the brain to differentiate between SCH patients and HCs. ALFF, amplitude of low-frequency fluctuation; 1, right cerebellum; 2, middle temporal gyrus; 3, right middle temporal gyrus; 4, left fusiform; 5, left anterior cingulate cortex; 6, bilateral precuneus; 7, left angular gyrus.

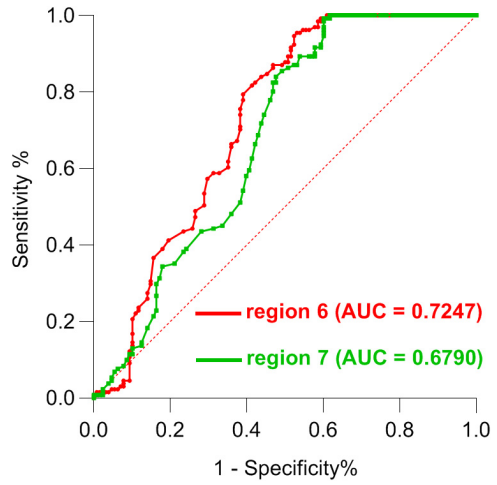


FIGURE 4
ROC curve-based differentiation between SCH patients and HC individuals based upon ALFF values in region 6 and region 7 of the brain. ROC, Receiver operating characteristic; SCH, schizophrenia; HCs, healthy controls; ALFF, amplitude of low-frequency fluctuation; 6, bilateral precuneus; 7, left angular gyrus.

lobe and plays a role in complex thought processes and the perception of auditory hallucinations (49). Previous reports have documented abnormal cortical asymmetry in the superior temporal gyrus in individuals with SCH, particularly in the AG (50). In the present study, SCH patients were generally considered to exhibit memory loss and slower cognition as these traits are associated with reductions in ALFF values within the AG. These abnormal AG findings offer important insight into the neurological basis for the characteristics of thought and language processing in individuals diagnosed with SCH. When ALFF values in the left AG were used as a biomarker to differentiate between SCH patients and HCs *via* an SVM approach, the associated specificity (71.88%), sensitivity (93.89%), and accuracy (52.34%) values suggested that reduced ALFF values in the left AG may represent an effective biomarker of SCH. No correlations, however, were observed between these reduced ALFF values and SCH patient disease severity or course. This may be attributable to the fact that brain function is impacted by a range of factors including compensatory mechanisms, or may suggest that decreased ALFF in left AG is a unique intrinsic characteristic of SCH that does not impact clinical symptoms of this disease.

The MTG plays a critical role in regulating sensory perception and is generally regarded as a central region of the brain necessary for memory and language functions (51,52). Neuroimaging data from SCH patients have consistently revealed a role for this region of the brain in the context of hallucinations and emotional processing. In addition, the MTG serves as a significant node in the network consisting of the frontal lobe, temporal lobe, parietal-occipital lobe, and subcortical structures (53,54). Abnormal MTG activation

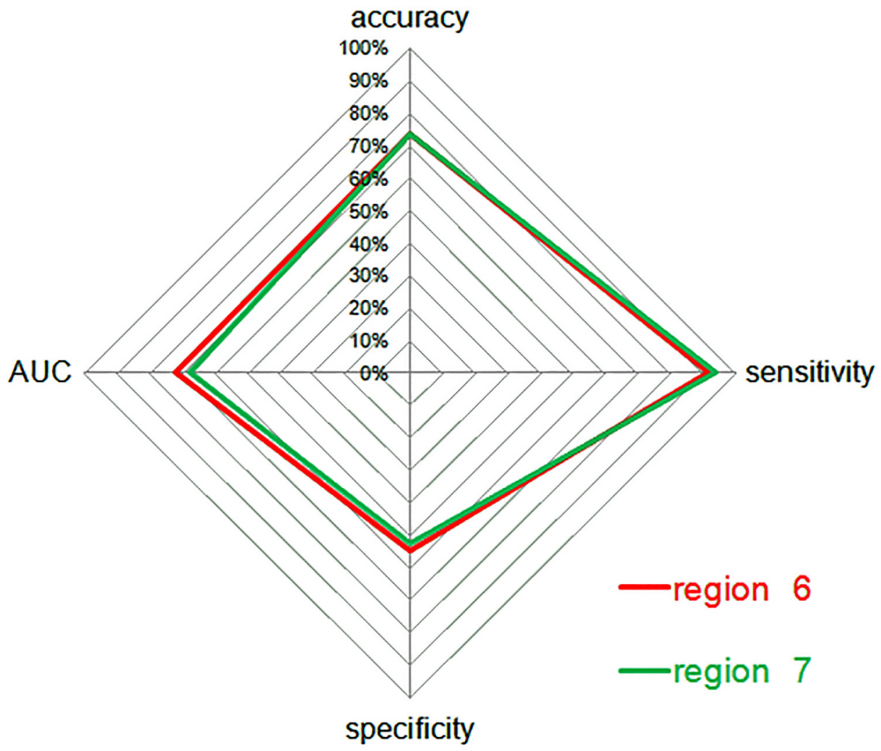


FIGURE 5
A radar plot demonstrating the accuracy, sensitivity, and specificity of classifications for SCH patients and HCs between region 6 and region 7, with corresponding AUC values. AUC, area under the curve; SCH, schizophrenia; HCs, healthy controls; 6, bilateral precuneus; 7, left angular gyrus.

TABLE 3 ROC analysis for differentiating SCH patients from HCs by using ALFF values in region 6 and 7.

Brain regions	Area under the curve	Cut—off point	Sensitivity	Specificity
Region 6	0.7247	1.160	91.60%	54.69%
Region 7	0.6790	1.295	93.89%	52.34%

ROC, receiver operating characteristic; SCH, schizophrenia; HCs, healthy controls; ALFF, amplitude of low-frequency fluctuation; 6, bilateral precuneus; 7, left angular gyrus; AUC, area under the curve.

may thus impair language processing and semantic memory functions. In this analysis, decreased ALFF in the bilateral MTG was observed in patients with SCH, and the resultant abnormalities may impair language and memory functions in these individuals, resulting in symptoms characteristic of this disease.

There are several limitations to this study. For one, some of the included SCH patients had been medicated prior to study initiation, and it is thus not possible to exclude the impact of such treatment on the observed structural and functional alterations in the brains of SCH patients (55). Subsequent studies of SCH patients that have not undergone drug treatment will be necessary to clarify this possibility. Second, this was a single-center study with a relatively small sample size, highlighting a need for additional large-scale multi-center validation of these results. Finally, deep learning is currently the most scientific in the exploration of biomarkers of brain function disorders in psychiatry. However. Our research only focuses on traditional SVM. In the next step, we plan to expand the sample size and collect schizophrenia data from multiple centers to further explore the imaging biomarkers of schizophrenia combined with artificial intelligence technology.

Conclusion

Overall, the results of this analysis suggest that SCH patients exhibit abnormal ALFF values in the left AG, ACC, fusiform, right cerebellum, bilateral PCu, and MTG. In particular, the reduced ALFF values in the left AG may offer value as a candidate biomarker to guide the objective diagnosis of SCH.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author/s.

Ethics statement

The study was approved by the ethics committee of the Wuhan Mental Health Center in accordance with the Declaration of Helsinki, and all subjects signed the written informed consent. The patients/participants provided their written informed consent to participate in this study.

Author contributions

YG and XT conceived the structure of the manuscript and wrote the manuscript. GW, JH, HH, and TG collected and analyzed the data. YL and GW conceived and critically reviewed the manuscript. All authors have read and approved the final manuscript.

Funding

This project received funding from the Wuhan Clinical Study Program (WG17M01).

Acknowledgments

We thank all the participants in this study. We would like to thank all the reviewers who participated in the review and MJEditor (www.mjeditor.com) for its linguistic assistance during the preparation of this manuscript. We also thank all individuals who served as study subjects.

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2022.949512/full#supplementary-material

References

1. Meepring S, Tulyakul P, Sathagathonthun G, Supasri J. A review of factors relating to medication non-adherence in patients with schizophrenia. *Glob J Health Sci.* (2021). 13:1–52.

2. Tandon R, Gaebel W, Barch DM, Bustillo J, Gur RE, Heckers S, et al. Definition and description of schizophrenia in the DSM-5. *Schizophr Res.* (2013) 150:3–10. doi: 10.1016/j.schres.2013.05.028

3. Zhao Z, Wang CH, Ma JD, Shan X, Shi LJ, Wang X, et al. Decreased left amygdala functional connectivity by cognitive-coping therapy in obsessive-compulsive disorder. *Mol Psychiatry.* (2021) 26:6952–62. doi: 10.1038/s41380-021-01131-z

4. Gan C, Wang L, Ji M, Ma K, Sun H, Zhang K, et al. Abnormal interhemispheric resting state functional connectivity in Parkinson's disease patients with impulse control disorders. *NPJ Parkinsons Dis.* (2021) 7:60. doi: 10.1038/s41531-021-00205-7

5. Barraclough M, McKie S, Parker B, Elliott R, Bruce IN. The effects of disease activity on neuronal and behavioural cognitive processes in systemic lupus erythematosus. *Rheumatology.* (2021) 61:195–204. doi: 10.1093/rheumatology/keab256

6. Zhu Q, Huang J, Xu X. Non-negative discriminative brain functional connectivity for identifying schizophrenia on resting-state Fmri. *Biomed Eng.* (2018) 17:32. doi: 10.1186/s12938-018-0464-x

7. Gangadin SS, Cahn W, Scheewe TW, Hulshoff Pol HE, Bossong MG. Reduced resting state functional connectivity in the hippocampus-midbrain-striatum network of schizophrenia patients. *J Psychiatr Res.* (2021) 138:83–8. doi: 10.1016/j.jpsychires.2021.03.041

8. Yang M, Jia X, Zhou H, Ren P, Deng H, Kong Z, et al. Brain dysfunction of methamphetamine-associated psychosis in resting state: approaching schizophrenia and critical role of right superior temporal deficit. *Addict Biol.* (2021) 26:e13044. doi: 10.1111/adb.13044

9. Dauvermann MR, Mothersill D, Rokita KI, King S, Holleran L, Kane R, et al. Changes in default-mode network associated with childhood trauma in schizophrenia. *Schizophr Bull.* (2021) 47:1482–94. doi: 10.1093/schbul/sbab025

10. Sklar AL, Coffman BA, Salisbury DF. Fronto-parietal network function during cued visual search in the first-episode schizophrenia spectrum. *J Psychiatr Res.* (2021) 141:339–45. doi: 10.1016/j.jpsychires.2021.07.014

11. Mothersill O, Tangney N, Morris DW, McCarthy H, Frodl T, Gill M, et al. Further evidence of alerted default network connectivity and association with theory of mind ability in schizophrenia. *Schizophr Res.* (2017) 184:52–8. doi: 10.1016/j.schres.2016.11.043

12. Jimeno N, Gomez-Pilar J, Poza J, Hornero R, Vogeley K, Meisenzahl E, et al. (Attenuated) hallucinations join basic symptoms in a transdiagnostic network cluster analysis. *Schizophr Res.* (2022) 243:43–54. doi: 10.1016/j.schres.2022.02.018

13. Tang S, Wang Y, Liu Y, Chau SW, Chan JW, Chu WC, et al. Large-scale network dysfunction in alpha-synucleinopathy: a meta-analysis of resting-state functional connectivity. *EBioMedicine*. (2022) 77:103915. doi: 10.1016/j.ebiom.2022.103915

14. Su, W, Yuan A, Tang Y, Xu L, Wei Y, Wang Y, et al. Effects of polygenic risk of schizophrenia on interhemispheric callosal white matter integrity and frontotemporal functional connectivity in first-episode schizophrenia. *Psychol Med.* (2022) 1:1–10. doi: 10.1017/S0033291721004840

15. Kirino E, Tanaka S, Fukuta M, Inami R, Inoue R, Aoki S. Functional connectivity of the caudate in schizophrenia evaluated with simultaneous resting-state functional mri and electroencephalography recordings. *Neuropsychobiology.* (2019) 77:165–75. doi: 10.1159/000490429

16. Hyza M, Kuhn M, Ceskova E, Ustohal L, Kasparek T. Hippocampal volume in first-episode schizophrenia and longitudinal course of the illness. *World J Biol Psychiatry*. (2016) 17:429–38. doi: 10.1080/15622975.2016.1199893

17. Li A, Zalesky A, Yue W, Howes O, Yan H, Liu Y, et al. A neuroimaging biomarker for striatal dysfunction in schizophrenia. *Nat Med.* (2020) 26:558–65. doi: 10.1038/s41591-020-0793-8

18. Yao B, Neggers SFW, Kahn RS, Thakkar KN. Altered thalamocortical structural connectivity in persons with schizophrenia and healthy siblings. *Neuroimage Clin*. (2020) 28:102370. doi: 10.1016/j.nicl.2020.102370

19. Li H, Ou Y, Liu F, Su Q, Zhang Z, Chen J, et al. Region-specific insular volumetric decreases in drug-naive, first-episode schizophrenia and their unaffected siblings. *Am J Med Genet B Neuropsychiatr Genet.* (2020) 183:106–12. doi: 10.1002/ajmg.b.32765

20. Li H, Ou Y, Liu F, Chen J, Zhao J, Guo W, et al. Reduced connectivity in anterior cingulate cortex as an early predictor for treatment response in drug-naive, first-episode schizophrenia: a global-brain functional connectivity analysis. *Schizophr Res.* (2020) 215:337–43. doi: 10.1016/j.schres.2019.09.003

21. Falkai PA. Desperate search for biomarkers in schizophrenia. What is going wrong? *World Psychiatry*. (2011) 10:38–9. doi: 10.1002/j.2051-5545.2011.tb00011.x

22. Takenaka S, Kan S, Seymour B, Makino T, Sakai Y, Kushioka J, et al. Resting-state amplitude of low-frequency fluctuation is a potentially useful prognostic functional biomarker in cervical myelopathy. *Clin Orthop Relat Res.* (2020) 478:1667–80. doi: 10.1097/CORR.0000000000001157

23. Liu X, Li L, Li M, Ren Z, Ma P. Characterizing the subtype of anhedonia in major depressive disorder: a symptom-specific multimodal MRI study. *Psychiatry Res Neuroimaging.* (2021) 308:111239. doi: 10.1016/j.pscychresns.2020.111239

24. Yu XM, Qiu LL, Huang HX, Zuo X, Zhou ZH, Wang S, et al. Comparison of resting-state spontaneous brain activity between treatment-naive schizophrenia and obsessive-compulsive disorder. *BMC psychiatry.* (2021) 1:544. doi: 10.1186/s12888-021-03554-y

25. Wu N, Zhang XY, Xia J, Li X, Yang T, Wang JH. Ratiometric 3d DNA machine combined with machine learning algorithm for ultrasensitive and high-precision screening of early urinary diseases. *ACS Nano.* (2021) 15:19522–34. doi: 10.1021/acsnano.1c06429

26. Ramkiran S, Sharma A, Rao NP. Resting-state anticorrelated networks in schizophrenia. *Psychiatry Res Neuroimaging.* (2019) 284:1–8. doi: 10.1016/j.pscychresns.2018.12.013

27. Kharat KD, Pawar VJ, Pardeshi SR editors. "Feature extraction and selection from MRI images for the brain tumor classification,"in *Proceedings of the 2016 International Conference on Communication and Electronics Systems (ICCES)* (Piscataway, NJ: IEEE) (2016).

28. Singh D, Kaur K. Classification of abnormalities in brain MRI Images Using GLCM, PCA and SVM. *Int J Eng Adv Technol.* (2012) 1:243–8. doi: 10.1080/00207454.2021.1901696

29. Yan H, Shan X, Li H, Liu F, Guo W. Abnormal spontaneous neural activity as a potential predictor of early treatment response in patients with obsessive-compulsive disorder. *J Affect Disord.* (2022) 309:27–36. doi: 10.1016/j.jad.2022.04.125

30. Phillips MR, Xiong W, Wang RW, Gao YH, Wang XQ, Zhang NP. Reliability and validity of the chinese versions of the scales for assessment of positive and negative symptoms. *Acta Psychiatr Scand.* (1991) 84:364–70. doi: 10.1111/j.1600-0447.1991.tb03161.x

31. Phillips R, Cheung YB, Collinson SL, Lim ML, Ling A, Feng L, et al. The equivalence and difference between the english and chinese language versions of the repeatable battery for the assessment of neuropsychological status. *Clin Neuropsychol.* (2015) 29(Suppl. 1):1–18. doi: 10.1080/13854046.2015.1034182

32. Jia XZ, Sun JW, Ji GJ, Liao W, Lv YT, Wang J, et al. Percent amplitude of fluctuation: a simple measure for resting-state FMRI signal at single voxel level. *PLoS One.* (2020) 15:e0227021. doi: 10.1371/journal.pone.0227021

33. Kang J, Wang L, Yan C, Wang J, Liang X, He Y. Characterizing dynamic functional connectivity in the resting brain using variable parameter regression and kalman filtering approaches. *Neuroimage.* (2011) 56:1222–34. doi: 10.1016/j.neuroimage.2011.03.033

34. Chang CC, Lin CJ. LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol.* (2007) 2:27. doi: 10.1145/1961189.1961199

35. Chang CC, Lin CJ. LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol.* (2011) 2:27. doi: 10.1145/1961189.1961199.36

36. Raichle ME. The brain's default mode network. *Annu Rev Neurosci.* (2015) 38:433–47. doi: 10.1146/annurev-neuro-071013-014030

37. Hu ML, Zong XF, Mann JJ, Zheng JJ, Liao YH, Li ZC, et al. A review of the functional and anatomical default mode network in schizophrenia. *Neurosci Bull.* (2017) 33:73–84. doi: 10.1007/s12264-016-0090-1

38. Garrity AG, Pearlson GD, McKiernan K, Lloyd D, Kiehl KA, Calhoun VD. Aberrant "default mode" functional connectivity in schizophrenia. *Am J Psychiatry.* (2007) 164:450–7. doi: 10.1176/ajp.2007.164.3.450

39. Chen G, Chen P, Gong J, Jia Y, Zhong S, Chen F, et al. Shared and specific patterns of dynamic functional connectivity variability of striato-cortical circuitry in unmedicated bipolar and major depressive disorders. *Psychol Med.* (2020) 52:747–56. doi: 10.1017/S0033291720002378

40. Gong J, Wang J, Chen P, Qi Z, Luo Z, Wang J, et al. Large-scale network abnormality in bipolar disorder: a multimodal meta-analysis of resting-state functional and structural magnetic resonance imaging studies. *J Affect Disord.* (2021) 292:9–20. doi: 10.1016/j.jad.2021.05.052

41. Zhang S, Yang G, Ou Y, Guo W, Peng Y, Hao K, et al. Abnormal default-mode network homogeneity and its correlations with neurocognitive deficits in drug-naive first-episode adolescent-onset schizophrenia. *Schizophr Res.* (2020) 215:140–7. doi: 10.1016/j.schres.2019.10.056

42. Wang YM, Zou LQ, Xie WL, Yang ZY, Zhu XZ, Cheung EFC, et al. Altered functional connectivity of the default mode network in patients with schizo-obsessive comorbidity: a comparison between schizophrenia and obsessive-compulsive disorder. *Schizophr Bull.* (2019) 45:199–210. doi: 10.1093/schbul/sbx194

43. Li P, Zhou M, Yan W, Du J, Lu S, Xie S, et al. Altered resting-state functional connectivity of the right precuneus and cognition between depressed and non-depressed schizophrenia. *Psychiatry Res Neuroimaging.* (2021) 317:111387. doi: 10.1016/j.pscychresns.2021.111387

44. Li YL, Li YD, Zhang H, Gao ZT, Xia YH, Liang YH, et al. [Relationship between auditory hallucination and regional homogeneity of functional magnetic resonance imaging in first-episode childhood and adolescence-onset schizophrenia]. *Zhonghua Yi Xue Za Zhi.* (2021) 101:1915–20. doi: 10.3760/cma.j.cn112137-20201126-03195

45. Tseng HH, Chiu CD, Chen KC, Lee IH, Chen PS, Yang YK. Absence of negative associations of insular and medial frontal gray matter volume with dissociative symptoms in schizophrenia. *J Psychiatr Res.* (2021) 138:485–91. doi: 10.1016/j.jpsychires.2021.04.017

46. Gong J, Wang J, Luo X, Chen G, Huang H, Huang R, et al. Abnormalities of intrinsic regional brain activity in first-episode and chronic schizophrenia: a meta-analysis of resting-state functional MRI. *J Psychiatry Neurosci.* (2020) 45:55–68. doi: 10.1503/jpn.180245

47. Tanglay O, Young IM, Dadario NB, Briggs RG, Fonseka RD, Dhanaraj V, et al. Anatomy and white-matter connections of the precuneus. *Brain Imaging Behav.* (2022) 16:574–86. doi: 10.1007/s11682-021-00529-1

48. Perret M, Lavallé L, Haesebaert F, Suaud-Chagny MF, Brunelin J, Mondino M. Neuroanatomical correlates of reality monitoring in patients with schizophrenia and auditory hallucinations. *Eur Psychiatry.* (2021) 64:e58. doi: 10.1192/j.eurpsy.2021.2234

49. Rolls ET, Cheng W, Gilson M, Gong W, Deco G, Lo CZ, et al. Beyond the disconnectivity hypothesis of schizophrenia. *Cereb Cortex.* (2020) 30:1213–33. doi: 10.1093/cercor/bhz161

50. Niznikiewicz M, Donnino R, McCarley RW, Nestor PG, Iosifescu DV, O'Donnell B, et al. Abnormal angular gyrus asymmetry in schizophrenia. *Am J Psychiatry.* (2000) 157:428–37. doi: 10.1176/appi.ajp.157.3.428

51. Oertel-Knöchel V, Knöchel C, Matura S, Prvulovic D, Linden DE, van de Ven V. Reduced functional connectivity and asymmetry of the planum temporale in patients with schizophrenia and first-degree relatives. *Schizophr Res.* (2013) 147:331–8. doi: 10.1016/j.schres.2013.04.024

52. Rey R, Suaud-Chagny MF, Bohec AL, Dorey JM, d'Amato T, Tamouza R, et al. Overexpression of complement component c4 in the dorsolateral prefrontal cortex, parietal cortex, superior temporal gyrus and associative striatum of patients with schizophrenia. *Brain Behav Immun.* (2020) 90:216–25. doi: 10.1016/j.bbi.2020.08.019

53. Xu Y, Lin Q, Han Z, He Y, Bi Y. Intrinsic functional network architecture of human semantic processing: modules and hubs. *Neuroimage.* (2016) 132:542–55. doi: 10.1016/j.neuroimage.2016.03.004

54. McDonald CR, Ahmadi ME, Hagler DJ, Tecoma ES, Iragui VJ, Gharapetian L, et al. Diffusion tensor imaging correlates of memory and language impairments in temporal lobe epilepsy. *Neurology.* (2008) 71:1869–76. doi: 10.1212/01.wnl.0000327824.05348.3b

55. Yan M, Chen J, Liu F, Li H, Zhao J, Guo W. Abnormal default mode network homogeneity in major depressive disorder with gastrointestinal symptoms at rest. *Front Aging Neurosci.* (2022) 14:804621. doi: 10.3389/fnagi.2022.804621

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 06 October 2022
DOI 10.3389/fpsyt.2022.1017888

OPEN ACCESS

EDITED BY
Zhi Xu,
Southeast University, China

REVIEWED BY
Chao Huang,
Nantong University, China
Gang Li,
Zhejiang Normal University, China

*CORRESPONDENCE
Zhongxia Shen
snowszx@sina.com
Xiaomei Zhang
zxm11982012@sina.com

SPECIALTY SECTION
This article was submitted to Computational Psychiatry, a section of the journal Frontiers in Psychiatry

RECEIVED 12 August 2022
ACCEPTED 13 September 2022
PUBLISHED 06 October 2022

CITATION
Su L, Shuai Y, Mou S, Shen Y, Shen X, Shen Z and Zhang X (2022) Development and validation of a nomogram based on lymphocyte subsets to distinguish bipolar depression from major depressive disorder.
Front. Psychiatry 13:1017888.
doi: 10.3389/fpsyt.2022.1017888

Development and validation of a nomogram based on lymphocyte subsets to distinguish bipolar depression from major depressive disorder

Liming Su[1], Yibing Shuai[1], Shaoqi Mou[2], Yue Shen[1], Xinhua Shen[1], Zhongxia Shen[1]* and Xiaomei Zhang[1]*

[1]Department of Neurosis and Psychosomatic Diseases, Huzhou Third Municipal Hospital, The Affiliated Hospital of Huzhou University, Huzhou, China, [2]Department of Psychiatry, Wenzhou Medical University, Wenzhou, China

Objective: Bipolar depression (BD) and major depressive disorder (MDD) are both common affective disorders. The common depression episodes make it difficult to distinguish between them, even for experienced clinicians. Failure to properly diagnose them in a timely manner leads to inappropriate treatment strategies. Therefore, it is important to distinguish between BD and MDD. The aim of this study was to develop and validate a nomogram model that distinguishes BD from MDD based on the characteristics of lymphocyte subsets.

Materials and methods: A prospective cross-sectional study was performed. Blood samples were obtained from participants who met the inclusion criteria. The least absolute shrinkage and selection operator (LASSO) regression model was used for factor selection. A differential diagnosis nomogram for BD and MDD was developed using multivariable logistic regression and the area under the curve (AUC) with 95% confidence interval (CI) was calculated, as well as the internal validation using a bootstrap algorithm with 1,000 repetitions. Calibration curve and decision curve analysis (DCA) were used to evaluate the calibration and clinical utility of the nomogram, respectively.

Results: A total of 166 participants who were diagnosed with BD (83 cases) or MDD (83 cases), as well as 101 healthy controls (HCs) between June 2018 and January 2022 were enrolled in this study. $CD19^+$ B cells, $CD3^+$ T cells, $CD3^-CD16/56^+$ NK cells, and total lymphocyte counts were strong predictors of the diagnosis of BD and MDD and were included in the differential diagnosis nomogram. The AUC of the nomogram and internal validation were 0.922 (95%; CI, 0.879–0.965), and 0.911 (95% CI, 0.838–0.844), respectively. The calibration curve used to discriminate BD from MDD showed optimal agreement between the nomogram and the actual diagnosis. The results of DCA showed that the net clinical benefit was significant.

Conclusion: This is an easy-to-use, repeatable, and economical nomogram for differential diagnosis that can help clinicians in the individual diagnosis of BD and MDD patients, reduce the risk of misdiagnosis, facilitate the formulation of appropriate treatment strategies and intervention plans.

KEYWORDS

bipolar depression (BD), major depressive disorder (MDD), lymphocyte subsets, differential diagnosis, nomogram

Introduction

Bipolar disorder (BD) and major depressive disorder (MDD) are two significant spectra of mental disorders characterized by clinical symptoms triggered by dysfunction in the emotional, cognitive and behavioral, and somatic domains (1). It has been reported that the lifetime prevalence of BD and MDD is 5 and 16.2% (2, 3), respectively, and the misdiagnosis of BD as MDD was as high as 40% (4), so the over-diagnosis of MDD is partly caused by the high rate of misdiagnosis of BD; which can be modified by adequate recognition and accuracy in diagnosing MDD from BD (5). In general, patients with BD who have a history of evident mania can be separated from those with MDD (6). However, most patients with BD experience milder or atypical forms of (hypo) mania, which are frequently overlooked by patients and clinicians (7). In addition, Perlis et al. found that it takes an average of 8 years time from the first attack of patients with BD to the proper diagnosis and treatment (8). Unfortunately, the accurate diagnosis of the two diseases can be challenging even for experienced physicians. During depressive episodes, there are overlapping clinical presentations of BD and MDD (9, 10), in which anxiety symptoms are the most commonly comorbidity symptoms (11) and the severity of current anxiety symptoms are strongly associated with the subsequent persistence of depressive symptoms (12). These cause the misdiagnosis of each other. It is noteworthy that the treatment strategies for these two diseases are quite different. In clinical practice, patients with BD have been prescribed with mood stabilizers (such as lithium salts, antiepileptics, and antipsychotics), whereas patients with MDD are prescribed with antidepressants only (13). Additionally, there is considerable debate regarding the effectiveness of antidepressants for BD (14). Given the high rate of misdiagnosis of BD and the differences in treatment strategies between BD and MDD, it is essential to make a correct diagnosis of BD and MDD. Otherwise, it will lead to the use of inappropriate drug treatment, with subsequent adverse outcomes and poor prognosis (e.g., prolonged morbidity, suicide, mania, and disability) (15–17), which severely impairs the quality of life of patients, and further increases the healthcare burden.

Currently, there is no gold standard for diagnosing BD and MDD. The existing diagnostic consensus points out that a comprehensive judgment should be combined with socio-demographics, clinical features, biological markers, brain electrophysiology, neuroimaging examinations, and treatment conditions (9). Nevertheless, there are still many issues to be investigated in the research and practice of BD and MDD identification. In recent years, it has been confirmed that the immune response plays a vital role in the pathogenesis of affective disorders (18, 19), especially lymphocyte subtypes consisting of Breg (B), Treg (T), and natural killer (NK) cells, have been proposed as inflammatory markers and supporting the inflammatory hypothesis underlying the etiopathogenesis of these conditions (19). Regarding the lymphocyte subtypes with the following characteristics and physiological functions: (1) B cells can be developed into plasma cells, which produce the antibodies important against foreign intruders, mainly represented by extracellular bacteria, and participate in the humoral immune response (20), and $CD19^{+}$ is one of the important membrane antigens involved in B cell activation and proliferation, and is a common surface marker for all B cells (21). (2). T cells are antigen specific, generated in the thymus, can be identified based on "cluster of differentiation" (CD) proteins expressed on their surface and/or on the cytokines they produce and are capable of differentiating into T cytotoxic or T helper cells (19, 20). The former present antigens to B cells, while the latter inhibit the function of effector T cells, B cells, and the proliferation of lymphocytes (20). All $CD3^{+}$ T cells express the typical T cell receptor, playing a role in antigen recognition (22). (3) NK cells are predominantly large granular lymphocytes (LGL), which express CD 16 and/or CD 56 surface antigens, have anti-tumor, anti-infection, immunomodulatory, and hematopoietic effects, and are essential lymphocytes in the natural immune system, and their killing effects are spontaneous and generally do not require the presence of antibodies or pre-sensitization (19).

Indeed, lymphocyte subtypes were performed by collecting blood samples and calculating the corresponding parameters, which has the advantage of being low-cost, reproducible, and easily available under simple laboratory conditions. Up to the

present date, several studies have examined the usefulness of lymphocyte subsets as potential biomarkers of neuro-inflammatory response with BD and MDD (23–25). For instance, a study by Bauer et al. showed that when patients with BD or MDD experienced acute psychological stress, the circulation of lymphocytes dysregulated, including the number of lymphatic subsets, and the corresponding functional impairment (1). Breunis, et al. found that patients with symptomatic and remitted BD had higher levels of activated $CD3^+$ T cells compared to healthy controls (HCs) (22) and the same results were obtained in the study of Barbosa et al. (26). However, 21 BD type I patients and 21 age- and sex-matched controls were recruited for a study by Barbosa et al. concluded that BD patients presented reduced proportions of $CD3^+$ T cells (26). Denney et al. observed in 15 MDD patients a decreased number of $CD3^+$ cells relative to HCs (27), but in a study of 20 patients with MDD, a slight increase in $CD3^+$ cells was described, but did not reach a significant difference with HCs (28). Besides, Müller et al., evaluated 37 patients suffering from a subtype of endogenous depression (during the depression and during free intervals), and found significantly higher levels of $CD3^+$ cells compared to HCs (29). Wu et al. reported that the distribution of T cells in BD patients was significantly different from that in MDD patients and patients with BD had significantly lower levels of cytotoxic T cells than patients with MDD (19, 25). The evidence from studies related to T-cell subsets in BD and MDD has not been clearly conclusive, indicating a gap in the field that deserves to be filled. Additionally, several studies have pointed to abnormal numbers and activity of B and NK cells in affective disorders. Mays et al. investigated whether the number or percentage of B cells reflected depressive states and noticed that depressed patients had significantly higher numbers of $CD19^+$ B cells compared to HCs (30). A recent study that compared BD patients with HCs on distribution of B lymphocyte subsets yielded a similar result: patients with BD who were in remission and depressive episodes had a higher percentage of $CD19^+$ B cells than HCs (31). The last but not least, Benschop et al. showed that acute psychological stress induces a transient increase in lymphocyte numbers, with NK cell counts being the most prominent (32); on the contrary, Patas et al. discovered that decreased NK cell counts were observed in patients with MDD (33, 34). Studies on exploration of NK cell levels in BD are lacking, and the differences in level of NK and B cell between BD and MDD are unclear. In light of the close relationship of immune response between BD and MDD, B and NK cells may play different roles, though the specific mechanism of this association remains to be fully determined. Taken together, these findings suggest that additional more investigation into differences in lymphocyte subsets between BD and MDD is necessary to reveal the relevant pathophysiological mechanisms and assist in clinical diagnosis.

Typically, because of the interplay of variables among lymphocyte subsets, focusing on the significance of a single factor may overlook the highly correlated factors from a clinical standpoint. Henceforth, to avoid these deficiencies, a clinical decision-making approach currently widely used in affective disorders – based on machine learning techniques, may help deal with complex factors, and show high predictive power (35, 36). Regretfully, it is rarely utilized to differentiate between BD and MDD disorders. Of these computational methods or models, the least absolute shrinkage and selection operator (LASSO) is one of the most widely utilized algorithms, which can comprehensively consider the synergistic effect among multiple influencing factors (37, 38), and select the most effective predicting factors from the available data set (35). Furthermore, a nomogram is an individualized and evidence-based predictive model that provides accurate information for decision-making objectively (39). As a result, applying machine learning techniques based on previous studies can provide a trustful methodological reference for this study to determine the predictors of BD and MDD.

The objectives of this study were as follows: (1) to explore the diagnostic value of lymphocyte-based subsets in distinguishing BD from MDD; (2) to develop and validate a differential diagnosis model to differentiate between BD and MDD with a base-machine learning technique, and was presented as a nomogram to improve the accuracy and usefulness of the predictive model. In all, in order to provide a simple and effective diagnostic tool for early differentiation of these two disorders and provide evidence support for follow-up clinical intervention.

Materials and methods

Study design

From June 2018 to January 2022, a prospective cross-sectional study was performed at Huzhou Third Municipal Hospital. This protocol was approved by the ethical review board of this institution and all participants provided signed informed consent before enrollment. In this study, assuming the nomogram could distinguish BD from MDD with a sensitivity and specificity of 90%, at least 158 participants (40).

Inclusion criteria were as follows: (1) age 18–65 years; (2) diagnosed with BD or MDD according to diagnostic and statistical manual of mental Disorders-IV (DSM-IV) criteria; (3) 17-item Hamilton rating scale for depression ($HAM\text{-}D_{17}$) score $\geq$14. Exclusion criteria were as follows: (1) with immune system diseases (e.g., rheumatoid arthritis, systemic lupus erythematosus); (2) used anti-inflammatory or immunosuppressive drugs in the past year; (3) infectious diseases in the past 2 weeks; (4) comorbidity of other mental or nervous system diseases; (5) any clinically serious physical diseases (including chronic diseases such as hypertension, hyperlipidemia, and diabetes); (6) history of pregnancy and

lactation; (7) blood dyscrasia, hepatic or renal failure, or obesity (body mass index (BMI) > 30 kg/m^2); and (8) alcohol, drugs or other psychotropic substance use disorder.

Finally, 166 participants diagnosed with BD (n = 83) or MDD (n = 83) met the inclusion criteria. According to the match ratio of 1:1:1, 83 cases of HCs were required, and 101 cases of HCs were enrolled. This prediction model study was carried out in accordance with the Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD) checklist (41) (shown in **Supplementary material 1**).

Information collection

Totally 29 factors, included clinical features [age, gender, height, weight, and BMI, onset age, illness duration, current episode duration, smoking, drinking, education level, employment status, marital status, family history of affective disorders, Hamilton rating scale for anxiety (HAMA) score, HAM-D_{17} score], and the characteristic of lymphocyte subsets [$CD19^+$ B cell, $CD19^+$ B cell (%), $CD3^+$ total T cell, $CD3^+$ total T cell (%), $CD3^+/CD4^+$ T-helper cell, $CD3^+/CD4^+$ T-helper cell (%), $CD3^+/CD8^+$ T-cytotoxic cell, $CD3^+/CD8^+$ T-cytotoxic cell (%), $CD3^-/CD16^+56^+$ NK cell, $CD3^-/CD16^+56^+$ NK cell (%), $CD4^+CD8^+$ (ratio), total lymphocyte count, and DN T (%)].

Instruments for evaluating clinical symptoms

All participants underwent the structural clinical interview for DSM-IV disorders (SCID) by experienced psychiatrists and were finally diagnosed with BD or MDD according to DSM-IV. The Chinese version of the HAM-D_{17} (42) and HAMA (43) scales were used to determine the severity of depressed mood and anxiety symptoms in participants with BD and MDD, respectively. These were the most frequently used assessment tools in clinical practice.

The procedure for collecting blood samples

Peripheral blood samples were collected after a 12 h overnight fast; approximately 2 mL of blood was drawn and placed in a blood sample tube containing EDTA anticoagulant. Blood samples were processed within 30 min of collection using BriCyte E6 flow cytometry (Mindray, Shenzhen, Guangdong, China) and performed in the Hangzhou Dian laboratory by technicians who were blinded to this study.

The procedure for detecting lymphocyte subsets

According to the sample preparation instructions provided by the manufacturer, two flows of sample tubes should be taken, and they should be sequentially numbered with the test number; The procedure for using quality control materials (Beckman Coulte, Inc., Brea, CA, USA) is the same as the procedure for using standard sample tests. First, 10 μl of CD3/CD8/CD45/CD4 and $CD3/CD16^+CD56$/CD45/CD19 antibodies (Mindray, Shenzhen) were injected into the numbered flow sampling tubes, respectively. After inverting and mixing the blood samples 6–8 times, 50 μl of blood samples were injected into the corresponding numbered tubes by the trans-addition technique, mixed well, and left at room temperature in the dark for 15 min. Then, 450 μl of red blood cell lysates (1:9 dilution with deionized water) were injected into each tube, mixed well, and left again at room temperature in the dark for 15 min. Finally, the tests were performed on a calibrated instrument.

Statistical analysis

Statistical Package for the Social Sciences (SPSS) statistics software for windows (version 26.0) was used to make a descriptive statistical analysis for illustrating the clinical information and the characteristics of lymphocyte subsets in the HCs, BD, and MDD groups. Fisher's exact test and Chi-square test were used to analyze proportional differences for categorical variables [n (%)], which were presented as frequency and percentage. Before applying statistical analysis to the continuous variables, the Shapiro–Wilk's test was used to check the data normality; Continuous variables with a normal distribution were expressed as mean and standard deviation [Mean (S.D.)], and the univariate ANOVA was used for comparison between multiple groups and Bonferroni correction was used for pairwise comparison between groups and an independent t-test was performed to assess differences between BD and MDD groups in univariate analysis; Non-normal distribution continuous data was represented as median and interquartile range [Median (IQR)], the Kruskal–Wallis H test was used for comparison between multiple groups, and Nemenyi correction was used for pairwise comparison between groups, the Mann–Whitney U test was used to evaluate differences between BD and MDD groups in univariate analysis.

Rest analyses were conducted using R statistical software for windows (version 4.0.4). The LASSO regression algorithm with the "glmnet" package (44) was used to screen for factors that were significantly associated with BD versus MDD status and eliminate multicollinearity between factors. The performance of this classification algorithm is mainly determined by the

parameter (lambda). As a result, 10-fold cross-validation (45) was adopted to select lambda with the minimum criteria. The optimal factors determined by LASSO analysis were further examined by univariate and multivariate logistic regression analysis, with the results demonstrated as odds ratio (OR) with 95% confidence interval (CI) and *P*-value. Subsequently, factors with statistically significant univariate (*P*-value < 0.05) were included multivariable logistic regression model after adjusting for age, sex, and BMI, and the backward stepwise selection was performed, with improvements in the goodness of fit measured by a decrease in the Akaike information criterion (AIC) (46). Following that, a differential diagnosis model was constructed based on the results of multivariate analysis.

Additionally, other R software packages that were used were (1) the "pROC" package for the receiver operating characteristic (ROC) curve graphics, which represented the area under the curve (AUC) and Harrell's concordance index (C-index), as well as computed the optimal cut-off value; (2) the "rms" package was utilized to generate a developed nomogram and calibration curve in order to visualize and calibrate the differential diagnosis model; (3) the "generalhoslem" package was used to perform the Hosmer–Lemeshow test (47) for detected goodness-of-fit of the nomogram, the result was supported by a *P*-value > 0.05; and (4) the "rmda" package was applied to perform the decision curve analysis (DCA) to judge the clinical utility of the nomogram based on the net benefit and threshold probability (48). All these packages are available on the website.[1] All *P*-value < 0.05 (two-sided) were considered statistically significant.

Ultimately, the accuracy of the nomogram was evaluated by internal validation through a bootstrap algorithm with 1,000 repetitions (49).

Results

Participants information

Gender, age, and lymphocyte subsets characteristics of the HCs (101 cases), BD (83 cases), and MDD (83 cases) groups were shown in **Supplementary Table 1**. Excluded sex-age differences and lymphocyte subsets were statistically different among the multiple comparisons. Further pairwise comparisons revealed that some lymphocyte subsets had statistically different comparisons of HCs and MDD group, HCs and BD group, and MDD and BD group. This vital evidence suggests that lymphocyte subsets may be effective biomarkers in distinguishing BD from MDD based on neuro-inflammation involved in the pathogenesis of BD and MDD. Based on the above, to assess the levels of MDD and BD lymphocyte subtypes in more detail, the potential factors for identifying MDD and BD were further explored in conjunction with clinical features and were shown in **Supplementary Table 2**. A total of 166 participants with BD and MDD were ultimately analysis, with more females than males in both groups. Participants with BD were younger in median onset age (25.74 years BD versus 33.39 years MDD), longer illness duration (136 months BD versus 36 months MDD), and higher median HAMD score (25 points BD versus 24 points MDD) compared to participants with MDD, and all three factors were statistically significant. In addition, there were nine factors of lymphocyte subsets that were significantly different between the MDD and BD groups, included $CD19^+$ B cell counts, $CD3^+$ T cell counts, $CD3^+/CD4^+$ T-helper cell counts, $CD3^+/CD4^+$ T-helper cell (%), $CD3^+/CD8^+$ T cell counts, $CD3^+/CD8^+$ T cell (%), $CD3^-CD16/56^+$ NK cell counts, $CD4^+CD8^+$ (ratio), and total lymphocyte counts. There are no factors with missing data.

Feature selection

A total of twelve factors with statistically significant differences in **Supplementary Table 2** were used as the original model and selected by LASSO regression. And based on the principle of the most regularized and stabilized model, the optimal parameter (lambda) through 10-fold cross-validation with minimum criteria was chosen. Then, the twelve factors were later reduced to eleven factors, including onset age, illness duration, HAMD score, $CD19^+$ B cell counts, $CD3^+$ T cell counts, $CD3^+/CD4^+$ T cell counts, $CD3^+/CD8^+$ T cell counts, $CD3^+/CD4^+$ T cell (%), $CD3^-CD16/56^+$ NK cell counts, $CD4^+CD8^+$ (ratio), and total lymphocyte counts, all with non-zero coefficients for the above. A cross-validated plot and a coefficient plot regarding the LASSO regression model are detailed in **Supplementary Figures 1A,B**.

Development of differential diagnosis model

The univariate and multivariate logistic regression was used to analyze further the eleven factors assessed by the LASSO regression model, with the findings reported in **Supplementary Table 3**. The final analysis revealed four independent predictors that were strongly associated with the diagnosis of BD and MDD, including $CD19^+$ B cell counts (OR, 1.109; 95% CI, 1.068–1.152), $CD3^+$ T cell counts (OR, 1.107; 95% CI, 1.065–1.150), $CD3^-CD16/56^+$ NK cell counts (OR, 1.120; 95% CI, 1.074–1.168), total lymphocyte counts (OR, 0.904; 95% CI, 0.871–0.939), and shown in a forest plot (**Supplementary Figure 2**). The AIC for the multivariable model was 117.06. Based on this, a differential diagnosis model with those mentioned above independent four predictors (*P*-value < 0.05) were created

1 https://cran.r-project.org/web/packages

and displayed as a ROC curve (**Supplementary Figure 3**) and nomogram (**Supplementary Figure 4**). The nomogram had an excellent discriminating ability with a high AUC of 0.922, the optimum cut-off value of this nomogram was 0.602, the sensitivity of 0.904, and specificity of 0.843. The nomogram shows the risk probability of each predictor from the multivariable model on the differential diagnosis of BD and MDD at the endpoint.

Performance and clinical utility of differential diagnosis nomogram

The curve shape of the calibration plot indicated that the differential diagnosis nomogram was well-calibrated (**Supplementary Figure 5A**), with a non-significant Hosmer–Lemeshow test result confirming the goodness-of-fit of the model ($\chi^2 = 9.471$, P-value = 0.395). The AUC of the nomogram was 0.922 (95% CI, 0.879–0.965) and was confirmed as 0.911 (95% CI, 0.838–0.844) after internal validation, indicating that the model had an excellent discriminatory ability. The apparent performance of the differential diagnosis nomogram demonstrated a good prediction ability for BD and MDD. The DCA for the differential diagnosis nomogram was shown in **Supplementary Figure 5B**. The DCA result demonstrated that the threshold probability of more than 0.05, with a substantial benefit for distinguishing between the BD and MDD groups.

Discussion

An overview of the findings

In the current study, given that blood sample collection is an easy-to-acquire technique, we used a flow cytometric analysis technique to quantify factors of lymphocyte subtypes from participants with BD and MDD as well as compare them with HCs. Finally, it was concluded that lymphocyte subtypes play an important role in neural pathways during depressive episodes in participants with MDD and BD as an indicator of status. Based on this, we applied the machine learning technique to identify four independent predictors in lymphocyte subtypes strongly associated with BD and MDD diagnosis. Simultaneously, a differential diagnosis nomogram had a discriminatory ability (C-index, 0.922) was developed and validated. According to our knowledge, this is the first study to provide valuable evidence that $CD19^+$ B cell counts, $CD3^+$ T cell counts, $CD3^-CD16/56^+$ NK cell counts, and total lymphocyte counts were diagnostic markers for individualized prediction of BD and MDD, which paves the way for the subsequent selection of appropriate clinical treatment strategies (e.g., drug selection) to guarantee a favorable prognosis and avoid unwanted outcomes.

Meaningful clinical features linked with the diagnosis of BD and MDD

In the univariate model of our study, onset age, illness duration, and HAMD score were associated with the diagnosis of BD and MDD, consistent with previous studies (50–52). Even though these factors were not included in the final nomogram, their clinical significance cannot be disregarded. It can be noted that the median onset age of the BD group was 7.65 years earlier than the MDD group in our cohort. This was closed to the findings of Tondo et al., who studied the onset age of BD and MDD from a cohort combined with 3,014 mood disorder patients, and found that the median onset age for the two disorders was 23 and 32 years, respectively, a difference of about 9 years. (53). It is widely believed that onset age helps differentiate BD from MDD, the typical onset of BD takes place from late adolescence to early adulthood (54). At the same time, MDD is more inclined to later ages onset (55). On the whole, BD patients are of earlier onset age, longer illness duration (56), with more severe clinical presentation and poorer outcomes (such as more episodes of depression, greater severity of depression) (51, 56).

Intriguing lymphocyte subsets are strongly associated with the diagnosis of BD and MDD

Based on multivariate logistic regression, the nomogram highlighted the four predictors used to differentiate BD from MDD based on lymphocyte subtypes: higher $CD19^+$ B cell counts, $CD3^+$ T cell counts, $CD3^-CD16/56^+$ NK cell counts, and lower total lymphocyte counts. Our nomogram illustrates how these various predictors can be combined to help predict the likelihood of BD. For example, for a participant with $CD19^+$ B cell counts, $CD3^+$ T cell counts, $CD3^-CD16/56^+$ NK cell counts, and total lymphocyte counts of 424.8, 2,056.6, 233.4, and 2,706 (cells/1 μl), respectively, the risk of being diagnosed with BD would be approximately 92.9%.

For elevated $CD19^+$ and $CD3^+$ cell counts in BD patients than MDD patients of our nomogram, and the results of the pairwise comparison further showed that $CD19^+$ B cells were higher in both BD and MDD than in HCs; this result is consistent with previous studies exploring the comparison of $CD19^+$ levels in BD or MDD with HCs (30, 31). Interestingly, our study compared the difference in $CD19^+$ levels between BD and MDD, with MDD being lower than BD. Moreover, patients with MDD had decreased $CD3^+$ T cells, whereas patients with BD had greater $CD3^+$ cell levels than HCs, and BD than MDD. As stated by Karlijn et al., the absence of T cells in MDD and a normal or overactive T cell system may be characteristic of BD (19). Therefore, the $CD3^+$ T cells and $CD19^+$ B cells become diagnosis indicators of BD and MDD, which relevant

T lymphocyte paradigm may help explain the finding. There is evidence that T cells activation has indeed been highlighted as a possible trait in BD patients (57), and is a response to the chronic low inflammatory state caused by an excess of immune modulation which inducing relevant symptoms (58), particularly the elevation level of activated $CD3^+$ T cells (22). Later findings have shown that an increased proportion of activated T cells and a trend in T helper 2 (Th2) activation in BD patients (22, 26). Moreover, there is a cognate interaction between B cells and T cells (59), B cells do not exhibit regulatory functions (60) due to the lack of major histocompatibility complex class II and B7 (61, 62), however, the activation of Th2 from T cell creates the conditions for activating B cells and the production of autoantibodies, further increased $CD19^+$ levels (22). In this line, it shows that the Th2 cell play a crucial role as a hub. Leday et al. suggested that MDD is associated with the downregulation of genes related to T lymphocyte function and adaptive immunity (63). On the one hand, it may be due to the lack of Th2 expressed by T cells, therefore, the inability to activate B cells, resulting in mildly reduced or near-normal $CD19^+$ levels. On the other hand, it is the lack of T lymphocytes and the inability to induce f interleukin (IL)-10 in the meninges and prefrontal cortex (PFC); hence, a genetic predisposition to produce less IL-10 is associated with a higher risk for depressive symptoms (64).

Compared to MDD, NK levels were higher in BD, but both were below normal levels compared to HCs, all of which are supported by the results of previous studies. Studies revealed that patients with BD exhibit a decrease in NK cells when they experience mania as compared to HCs. (65), whereas in outpatients with stable BD, there was no change in the number of NK cells compared to HCs (22). Furthermore, MDD may be associated with reduced natural killer cell activity (NKCA) and NK cell counts, especially early-onset-age MDD (66). The reason for this discrepancy between BD and MDD suggests that an interaction between affective disorders and various parameters of the immune system may be associated with altered immunity (65, 66). It is proposed that chronic pro-inflammatory processes in individuals with MDD may directly affect NK cells, suppressing NK cell function and numbers (67). We all know that when depression occurs in patients with BD, it is accompanied by the onset of mania. Tsai et al. demonstrated the presence of cell-mediated activation when BD patients experience mania. They found that, compared to HCs, BD patients with the manic had a proliferation of lymphocytes responding to phytohemagglutinin (PHA), and soluble interleukin-2R receptor (sIL-2R) was released from activated T cells into the blood, leading to a significant increase in plasma sIL-2R levels, thereby reducing cellular immune function (68).

Interestingly, some studies thought lymphopenia is a common observation in patients with BD and MDD (25, 64). Nevertheless, not absolute, Denis et al. presented a study on the alterations in neutrophils and lymphocytes in the blood of mood disorder patients, and discovered that lymphopenia occurs solely in MDD and not in BD (69), which is the same as our conclusion. We found significantly fewer total lymphocyte counts in MDD compared with HCs, but no difference was found in BD; and total lymphocyte counts in BD were more significant than in MDD, which is consisted with the study of Abeer et al. (65). In addition, combining the results of multivariate logistic regression further revealed that total lymphocyte count was negatively associated with the occurrence of BD and, conversely, positively associated with the occurrence of MDD. As shown by the results of our nomogram, individuals with lower total lymphocyte counts were more likely to be diagnosed with BD than MDD. As of now, the mechanisms of total lymphocyte count in mood disorders remain incompletely understood, and there is contradictory information concerning lymphocyte subsets and total lymphocyte counts in patients with BD and MDD (70). Insufficient numbers of circulating T helper cells and their subtypes in MDD and BD by previous scholars can lead to increased T cell apoptosis (19), which may further cause a decrease in total lymphocyte counts. Neuroinflammation mechanisms in patients with BD and MDD have been extensively described in terms of the levels of cytokines, HPA axis, neurotransmitters, and neurotrophic factors (71). Therefore, apoptosis may be associated with neuroinflammatory mechanisms, the findings of previous studies may provide some evidence to explain the difference between MDD and BD, which can be include as the following aspects.

(i) Patients with BD produce higher levels of pro-inflammatory cytokines during acute exacerbations due to the inhibition of T helper cell proliferation and function by tumor necrosis factor alpha (TNF-α) (72, 73). However, there is also a trend for elevated pro-inflammatory cytokine levels in MDD patients (74), which has been suggested that elevated pro-inflammatory cytokines could be responsible for the difference in T helper cell apoptosis in MDD and BD patients.

(ii) Differences in relevant pharmacological treatment strategies between MDD and BD. In particular, lithium or valproic acid do affect on cell apoptosis/proliferation (75, 76), and therefore it can be assumed that this class of drugs produces the effect. However, Ezequiel et al. found that when patients received any therapeutic dose of an emotional drug (lithium or valproic acid) or not, the proportion of lymphocytes exhibiting apoptosis was significantly increased in BD patients compared to HCs, which correlated with the high expression of Bax protein in patient cells (77). Similarly, Bei et al. found similar results in BD patients where increased lymphocyte apoptosis was caused by increased cytochrome C content in the cytoplasm and translocation of Bax protein (78). It can be seen that the difference between BD and MDD is not only limited to drug differences, but is more caused by the own endogenous substances of patients.

(iii) It may associate with tryptophan depletion. Tryptophan is a proliferation stimulator for T cells, and depletion of tryptophan has been shown to lead to decreased T cell proliferation and apoptosis of T cells (79). Activation of indoleamine-2,3-dioxygenase (IDO) causes the enhanced synthesis of tryptophan in MDD, leading to excessive degradation of tryptophan (80). However, low tryptophan levels are also present in BD patients (81). Therefore, tryptophan depletion needs to be explored in depth.

(iv) The abnormal state of microglia (microglia activation) and other immunocompetent cells in the brain is the driving force behind the phenomenon (82). Stertz et al. postulated that BD patients undergoing an acute emotional episode induces neuronal damage, which results in the discharge of damage-associated molecules that actuate microglia (83). Activated microglia can then generate pro-inflammatory cytokines and neurotrophic factors. These molecules cause alteration in the synaptic environment *via* synaptic pruning as an adaption attempt to the damage caused by acute episode. Then, after multiple repeated episodes, the overproduction of pro-inflammatory cytokines outstrips the normal down-regulatory ability to down-regulate them, causing microglia to remain constantly activated and inducing apoptosis failure to develop adequate (adaptive) stress-related responses (84, 85). Again, this phenomenon has been found to be present in patients with MDD. Studies have confirmed the ability of Th1 and Th17 cells to promote neuroinflammation and the activation of microglia and astrocytes (86, 87). In addition, high levels of T helper 1 (Th1) and T helper 17 (Th17) are often considered to be markers of inflammation, while high levels of T regulatory, and Th2 cells play an anti-inflammatory role. And Th17, T regulatory, and Th2 cells are all trending up in BD patients and conversely, down in MDD patients. In other words, it seems that both pro-inflammatory and anti-inflammatory forces are activated in BD, but both are suppressed in MDD, both of which contribute to a greater degree of apoptosis in BD than in MDD (19, 88–90).

Satisfactory performance for differential diagnosis nomogram

This study is based on a novel and widely used methodologies in mental disorders-machine learning techniques, including the LASSO analysis method and nomogram has the ability to superimpose the relative risk of important features and improve its predictive power (91), and eliminate interactions between features (92). Nomogram as a distinguishable predictive tool, not only correctly determines disease conditions based on the level of connection between features and diseases, but also delivers targeted risk assessments for the corresponding disease groups by building risk thresholds for diagnostic decisions (93, 94). Based on the selected features, we constructed a differential diagnosis nomogram that can serve as a feasible scoring system. We obtained a good verification result through the bootstrapping method (C-index, 0.911) and the calibration curve and DCA curve also exhibited favorable agreement and applicability, respectively. In view of our findings, all these predictors could be low-economical and easily obtained in the blood collection, which is of great significance for patients with mood disorders. Additionally, clinical staff can use this individualized diagnostic model to diagnose the probability of BD and MDD, which aligns with the current concept of precision medicine (95).

Strength

One of the strengths of this study is that as a prospective study, it facilitated us to design stringent diagnostic, inclusion, and exclusion criteria for an accurate distinction between BD and MDD to guarantee the authenticity and reliability of the data source. Another strength is our study based on machine learning techniques to explore $CD19^+$ B cell counts, $CD3^+$ T cell counts, $CD3^-CD16/56^+$ NK cell counts, and total lymphocyte counts as new predictors to differentiate BD and MDD.

Limitation and future directions

Although we present a helpful tool for distinguishing between BD and MDD, there are not, however, devoid of limitations. Firstly, the recruitment of participants for our study was performed with a small sample of participants who had BD and MDD at a single center. Although the internal validation of nomograms has good performance, it is still debatable whether nomograms have extrapolation power, and further expansion of sample size or replication in combination with multicenter is needed. Secondly, on the one hand, our nomogram was developed with the clinical features and the characteristic of lymphocyte subsets. The clinical features should be further refined, such as obesogenic diet and medication use; in addition to lymphocyte subsets, immunophenotyping and plasma cytokine should be added. On the other hand, incorporating the computed tomography (CT), electroencephalographic (EEG), and magnetic resonance imaging (MRI) information; combining the above additional variables to continue the prospective data collection effort could optimize our current nomogram. Thirdly, we did not collect blood samples after follow-up, so in future studies, we will build a follow-up database to obtain more dynamic changes in lymphocyte subtype levels and further investigate the dynamic role of lymphocyte subtypes in the neuroinflammatory pathways of BD and MDD. Fourth, from Müller et al., there is heterogeneity in patients with BD and MDD, leading to conflicting neuro-immunological findings in patients with

affective disorders, especially related to different depression subtypes (29). The participants recruited in this study included both first-episode and relapsed BD and MDD participants and did not distinguish between BD and MDD subtypes to explore differences in lymphocyte subtypes, which could lead to selection bias and may affect the study results, which requires refinement of the study protocol in future studies by designing strict inclusion criteria further explore the differences between first-episode BD and MDD patients, relapsed BD and MDD patients, BD subtypes and MDD subtypes patients between the lymphocyte subtype levels, and to verify whether nomograms have similar diagnostic power when explored in the above stratification. Finally, all participants are likely to be treated with mood stabilizing medications or antidepressants, which inevitably impact the regulation of immune function. Therefore, the next step should be to collect detailed information about medication use of participants and explore changes in lymphocyte subtypes under drug stratification.

Conclusion

Our findings suggest that lymphocyte subsets ($CD19^+$ B cell count, $CD3^+$ T cell count, $CD3^-CD16/56^+$ NK cell count, and total lymphocyte count) play an important role in the pathogenesis of BD and MDD as an indicator of biological status, especially during depressive episodes. As blood samples are a simple and economical test that is routinely performed in patients with BD and MDD, investment in this line of research may lead to the discovery of a firmer link between biological parameters and psychopathological indicators of BD and MDD. These complex biomarkers of affective disorders could in turn facilitate the identification of new therapeutic strategies.

Furthermore, the differential diagnosis nomogram developed in our study based on machine learning techniques is well calibrated and discriminatory and could be an easy-to-use, repeatable, and economical diagnostic tool to help clinicians differentiate between BD and MDD. This also demonstrates that machine learning approaches are very effective in dealing with the interactions between biomarkers of neuroinflammatory disorders. Therefore, based on the existing findings and previous research findings, we are confident enough to expect that machine learning party techniques will play a more effective role in disease neuroinflammation research and model building in BD and MDD compared to traditional statistical analysis methods, which may be a pleasing trend in the future.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding authors.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Huzhou Third Municipal Hospital. The patients/participants provided their written informed consent to participate in this study.

Author contributions

ZS, LS, and XS: study concept and design. XZ, YBS, YS, and SM: data collection. LS and YS: statistical analysis of the data. LS and ZS: manuscript preparation. ZS, XS, and YBS: critical revision of the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study has received funding from the Huzhou Public Welfare Research Project Social Development Category (2018GYB49, ZS) and the Social Development Project of Public Welfare Technology Application in Zhejiang Province in 2019 (LGF19H090002, ZS).

Conflict of interest

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2022.1017888/full#supplementary-material

References

1. Bauer ME, Teixeira AL. Neuroinflammation in mood disorders: role of regulatory immune cells. *Neuroimmunomodulation.* (2021). 28:99–107. doi: 10.1159/000515594

2. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, et al. National comorbidity survey replication. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *JAMA.* (2003) 289:3095–105. doi: 10.1001/jama.289.23.3095

3. Merikangas KR, Akiskal HS, Angst J, Greenberg PE, Hirschfeld RM, Petukhova M, et al. Lifetime and 12-month prevalence of bipolar spectrum disorder in the National Comorbidity Survey replication. *Arch Gen Psychiatry.* (2007) 64:543–52. doi: 10.1001/archpsyc.64.5.543

4. Judd LL, Akiskal HS, Schettler PJ, Endicott J, Maser J, Solomon DA, et al. The long-term natural history of the weekly symptomatic status of bipolar I disorder. *Arch Gen Psychiatry.* (2002) 59:530–7. doi: 10.1001/archpsyc.59.6.53

5. Muzina DJ, Kemp DE, McIntyre RS. Differentiating bipolar disorders from major depressive disorders: treatment implications. *Ann Clin Psychiatry.* (2007) 19:305–12. doi: 10.1080/10401230701653591

6. Liebers DT, Pirooznia M, Ganna A, Bipolar Genome Study [BiGS], Goes FS. Discriminating bipolar depression from major depressive disorder with polygenic risk scores. *Psychol Med.* (2021) 51:1451–8. doi: 10.1017/S003329172000015X

7. Fusar-Poli L, Natale A, Amerio A, Cimpoesu P, Grimaldi Filioli P, Aguglia E, et al. Neutrophil-to-luymphocyte, platelet-to-lymphocyte and monocyte-to-lymphocyte ratio in bipolar disorder. *Brain Sci.* (2021) 11:58. doi: 10.3390/brainsci1101005

8. Perlis RH, Ostacher MJ, Goldberg JF, Miklowitz DJ, Friedman E, Calabrese J, et al. Transition to mania during treatment of bipolar depression. *Neuropsychopharmacology.* (2010) 35:2545–52. doi: 10.1038/npp.2010.122

9. McIntyre RS, Zimmerman M, Goldberg JF, First MB. Differential diagnosis of major depressive disorder versus bipolar disorder: current status and best clinical practices. *J Clin Psychiatry.* (2019) 80:ot18043ah2. doi: 10.4088/JCP.ot18043ah2

10. Hirschfeld RM. Differential diagnosis of bipolar disorder and major depressive disorder. *J Affect Disord.* (2014) 169:S12–6. doi: 10.1016/S0165-032770004-7

11. Goghari VM, Harrow M. Anxiety symptoms across twenty-years in schizoaffective disorder, bipolar disorder, and major depressive disorder. *Psychiatry Res.* (2019) 275:310–4. doi: 10.1016/j.psychres.2019.03.050

12. Coryell W, Fiedorowicz JG, Solomon D, Leon AC, Rice JP, Keller MB. Effects of anxiety on the long-term course of depressive disorders. *Br J Psychiatry.* (2012) 200:210–5. doi: 10.1192/bjp.bp.110.081992

13. Yatham LN, Kennedy SH, Schaffer A, Parikh SV, Beaulieu S, O'Donovan C, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) and International Society for Bipolar Disorders (ISBD) collaborative update of CANMAT guidelines for the management of patients with bipolar disorder: update 2009. *Bipolar Disord.* (2009) 11:225–55. doi: 10.1111/j.1399-5618.2009.00672.x

14. Pacchiarotti I, Bond DJ, Baldessarini RJ, Nolen WA, Grunze H, Licht RW, et al. The International Society for Bipolar Disorders (ISBD) task force report on antidepressant use in bipolar disorder. *Am J Psychiatry.* (2013) 170:1249–62. doi: 10.1176/appi.ajp.2013.13020185

15. Baldessarini RJ, Undurraga J, Vázquez GH, Tondo L, Salvatore P, Ha K, et al. Predominant recurrence polarity among 928 adult international bipolar I disorder patients. *Acta Psychiatr Scand.* (2012) 125:293–302. doi: 10.1111/j.1600-0447.2011.01818.x

16. Nasrallah HA. Consequences of misdiagnosis: inaccurate treatment and poor patient outcomes in bipolar disorder. *J Clin Psychiatry.* (2015) 76:e1328. doi: 10.4088/JCP.14016tx2c

17. Valentí M, Pacchiarotti I, Bonnín CM, Rosa AR, Popovic D, Nivoli AM, et al. Risk factors for antidepressant-related switch to mania. *J Clin Psychiatry.* (2012) 73:e271–6. doi: 10.4088/JCP.11m07166

18. Maes M. Depression is an inflammatory disease, but cell-mediated immune activation is the key component of depression. *Prog Neuropsychopharmacol Biol Psychiatry.* (2011) 35:664–75. doi: 10.1016/j.pnpbp.2010.06.014

19. Becking K, Haarman BCM, Grosse L, Nolen WA, Claes S, Arolt V, et al. The circulating levels of CD4+ t helper cells are higher in bipolar disorder as compared to major depressive disorder. *J Neuroimmunol.* (2018) 319:28–36. doi: 10.1016/j.jneuroim.2018.03.004

20. Delves PJ, Roitt IM. The immune system. *N Engl J Med.* (2000) 343:37–49. doi: 10.1056/NEJM200007063430107

21. Rosser EC, Mauri C. Regulatory B cells: origin, phenotype, and function. *Immunity.* (2015) 42:607–12. doi: 10.1016/j.immuni.2015.04.005

22. Breunis MN, Kupka RW, Nolen WA, Suppes T, Denicoff KD, Leverich GS, et al. High numbers of circulating activated T cells and raised levels of serum IL-2 receptor in bipolar disorder. *Biol Psychiatry.* (2003) 53:157–65. doi: 10.1016/s0006-322301452-x

23. Vogels RJ, Koenders MA, van Rossum EFC, Spijker AT, Drexhage HA. T cell deficits and overexpression of hepatocyte growth factor in anti-inflammatory circulating monocytes of middle-aged patients with bipolar disorder characterized by a high prevalence of the metabolic syndrome. *Front Psychiatry.* (2017) 8:34. doi: 10.3389/fpsyt.2017.00034

24. Bulut NS, Yorguner N, Çarkaxhiu Bulut G. The severity of inflammation in major neuropsychiatric disorders: comparison of neutrophil-lymphocyte and platelet-lymphocyte ratios between schizophrenia, bipolar mania, bipolar depression, major depressive disorder, and obsessive compulsive disorder. *Nord J Psychiatry.* (2021) 75:624–32. doi: 10.1080/08039488.2021.1919201

25. Wu W, Zheng YL, Tian LP, Lai JB, Hu CC, Zhang P, et al. Circulating T lymphocyte subsets, cytokines, and immune checkpoint inhibitors in patients with bipolar II or major depression: a preliminary study. *Sci Rep.* (2017) 7:1–7. doi: 10.1038/srep40530

26. Barbosa IG, Rocha NP, Assis F, Vieira ÉLM, Soares JC, Bauer ME, et al. Monocyte and lymphocyte activation in bipolar disorder: a new piece in the puzzle of immune dysfunction in mood disorders. *Int J Neuropsychopharmacol.* (2014) 18:yu021. doi: 10.1093/ijnp/pyu021

27. Denney DR, Stephenson LA, Penick EC, Weller RA. Lymphocyte subclasses and depression. *J Abnorm Psychol.* (1988) 97:499. doi: 10.1037//0021-843x.97.4.499

28. Darko DF, Gillin JC, Risch SC, Bulloch K, Golshan S, Tasevska Z, et al. Immune cells and the hypothalamic-pituitary axis in major depression. *Psychiatry Res.* (1988) 25:173–9. doi: 10.1016/0165-178190048-0

29. Müller N, Hofschuster E, Ackenheil M, Mempel W, Eckstein R. Investigations of the cellular immunity during depression and the free interval: evidence for an immune activation in affective psychosis. *Prog Neuropsychopharmacol Biol Psychiatry.* (1993) 17:713–30. doi: 10.1016/0278-584690055-w

30. Maes M, Stevens WJ, DeClerck LS, Bridts CH, Peeters D, Schotte C, et al. A significantly increased number and percentage of B cells in depressed subjects: results of flow cytometric measurements. *J Affect Disord.* (1992) 24:127–34. doi: 10.1016/0165-032790060-j

31. Pietruczuk K, Lisowska KA, Grabowski K, Landowski J, Cubała WJ, Witkowski JM. Peripheral blood lymphocyte subpopulations in patients with bipolar disorder type II. *Sci Rep.* (2019) 9:5869. doi: 10.1038/s41598-019-42482-6

32. Benschop RJ, Rodriguez-Feuerhahn M, Schedlowski M. Catecholamine-induced leukocytosis: early observations, current research, and future directions. *Brain Behav Immun.* (1996) 10:77–91. doi: 10.1006/brbi.1996.0009

33. Patas K, Willing A, Demiralay C, Engler JB, Lupu A, Ramien C, et al. T cell phenotype and T cell receptor repertoire in patients with major depressive disorder. *Front Immunol.* (2018) 9:291. doi: 10.3389/fimmu.2018.00291

34. Suzuki H, Savitz J, Kent Teague T, Gandhapudi SK, Tan C, Misaki M, et al. Altered populations of natural killer cells, cytotoxic T lymphocytes, and regulatory T cells in major depressive disorder: association with sleep disturbance. *Brain Behav Immun.* (2017) 66:193–200. doi: 10.1016/j.bbi.2017.06.011

35. Zhou J, Zhou J, Sun Z, Feng L, Zhu X, Yang J, et al. Development and internal validation of a novel model to identify inflammatory biomarkers of a response to escitalopram in patients with major depressive disorder. *Front Psychiatry.* (2021) 12:593710. doi: 10.3389/fpsyt.2021.593710

36. Shen Z, Cui L, Mou S, Ren L, Yuan Y, Shen X, et al. Combining S100B and cytokines as neuro-inflammatory biomarkers for diagnosing generalized anxiety disorder: a proof-of-concept study based on machine learning. *Front Psychiatry.* (2022) 13:881241. doi: 10.3389/fpsyt.2022.881241

37. Reichling C, Taieb J, Derangere V, Klopfenstein Q, Le Malicot K, Gornet JM, et al. Artificial intelligence-guided tissue analysis combined with immune infiltrate assessment predicts stage III colon cancer outcomes in PETACC08 study. *Gut.* (2020) 69:681–90. doi: 10.1136/gutjnl-2019-319292

38. Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, et al. Development and validation of a radiomics nomogram for preoperative prediction of lymph node metastasis in colorectal cancer. *J Clin Oncol.* (2016) 34:2157–64. doi: 10.1200/JCO.2015.65.9128

39. Lei Z, Li J, Wu D, Xia Y, Wang Q, Si A, et al. Nomogram for preoperative estimation of microvascular invasion risk in hepatitis B virus-related hepatocellular carcinoma within the milan criteria. *JAMA Surg.* (2016) 151:356–63. doi: 10.1001/jamasurg.2015.4257

40. Hajian-Tilaki K. Sample size estimation in diagnostic test studies of biomedical informatics. *J Biomed Inform.* (2014) 48:193–204.

41. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. *BMJ.* (2015) 102:148–58. doi: 10.1136/bmj.g7594

42. Calotă DR, Nițescu C, Marinescu S, Cristescu C, Boiangiu I, Florescu IP, et al. Correlations between morphological appearance and psychosocial difficulties in patients with extensive burns who received allotransplant. *Rom J Morphol Embryol.* (2012) 53(3 Suppl.):703–11.

43. Matza LS, Morlock R, Sexton C, Malley K, Feltner D. Identifying HAM-A cutoffs for mild, moderate, and severe generalized anxiety disorder. *Int J Methods Psychiatr Res.* (2010) 19:223–32. doi: 10.1002/mpr.323

44. Sauerbrei W, Royston P, Binder H. Selection of important variables and determination of functional form for continuous predictors in multivariable model building. *Stat Med.* (2007) 26:5512–28. doi: 10.1002/sim.3148

45. Hastie T, Tibshirani R, Wainwright M. Statistical learning with sparsity: the lasso and generalizations. *Monogr Stat Appl Probab.* (2015) 143:143.

46. Liu L, Xie J, Wu W, Chen H, Li S, He H, et al. A simple nomogram for predicting failure of non-invasive respiratory strategies in adults with COVID-19: a retrospective multicentre study. *Lancet Digit Health.* (2021) 3:e166–74. doi: 10.1016/S2589-750030316-2

47. Kramer AA, Zimmerman JE. Assessing the calibration of mortality benchmarks in critical care: the Hosmer-Lemeshow test revisited. *Crit Care Med.* (2007) 35:2052–6. doi: 10.1097/01.CCM.0000275267.64078.B0

48. Vickers AJ, Cronin AM, Elkin EB, Gonen M. Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers. *BMC Med Inform Decis Mak.* (2008) 8:53. doi: 10.1186/1472-6947-8-53

49. Pencina MJ, D'Agostino RB. Overall C as a measure of discrimination in survival analysis: model specific population value and confidence interval estimation. *Stat Med.* (2004) 23:2109–23. doi: 10.1002/sim.1802

50. Bilska K, Pawlak J, Kapelski P, Narożna B, Zakowicz P, Szczepankiewicz A, et al. Differences in the clinical picture in women with a depressive episode in the course of unipolar and bipolar disorder. *J Clin Med.* (2021) 10:676. doi: 10.3390/jcm10040676

51. Joslyn C, Hawes DJ, Hunt C, Mitchell PB. Is age of onset associated with severity, prognosis, and clinical features in bipolar disorder? A meta-analytic review. *Bipolar Disord.* (2016) 18:389–403. doi: 10.1111/bdi.12419

52. Sánchez-Carro Y, Portella MJ, Leal-Leturia I, Salvat-Pujol N, Etxandi M, de Arriba-Arnau A, et al. Age at illness onset and physical activity are associated with cognitive impairment in patients with current diagnosis of major depressive disorder. *J Affect Disord.* (2021) 279:343–52. doi: 10.1016/j.jad.2020.10.032

53. Tondo L, Lepri B, Cruz N, Baldessarini RJ. Age at onset in 3014 Sardinian bipolar and major depressive disorder patients. *Acta Psychiatr Scand.* (2010) 121:446–52. doi: 10.1111/j.1600-0447.2009.01523.x

54. Kessing LV, Willer I, Andersen PK, Bukh JD. Rate and predictors of conversion from unipolar to bipolar disorder: a systematic review and meta-analysis. *Bipolar Disord.* (2017) 19:324–35. doi: 10.1111/bdi.12513

55. Eraydin IE, Mueller C, Corbett A, Ballard C, Brooker H, Wesnes K, et al. Investigating the relationship between age of onset of depressive disorder and cognitive function. *Int J Geriatr Psychiatry.* (2019) 34:38–46. doi: 10.1002/gps.4979

56. Baykara B, Koc D, Resmi H, Akan P, Tunca Z, Ozerdem A, et al. Brain-derived neurotrophic factor in bipolar disorder: associations with age at onset and illness duration. *Prog Neuropsychopharmacol Biol Psychiatry.* (2021) 108:110075. doi: 10.1016/j.pnpbp.2020.110075

57. Wieck A, Grassi-Oliveira R, do Prado CH, Rizzo LB, de Oliveira AS, Kommers-Molina J, et al. Differential neuroendocrine and immune responses to acute psychosocial stress in women with type 1 bipolar disorder. *Brain Behav Immun.* (2013) 34:47–55. doi: 10.1016/j.bbi.2013.07.005

58. Snijders G, Brouwer R, Kemner S, Bootsman F, Drexhage HA, Hillegers MHJ. Genetic and environmental influences on circulating NK and T cells and their relation to bipolar disorder. *Int J Bipolar Disord.* (2019) 7:4. doi: 10.1186/s40345-018-0139-3

59. Dong C, Flavell RA. Th1 and Th2 cells. *Curr Opin Hematol.* (2001) 8:47–51. doi: 10.1097/00062752-200101000-00009

60. Rosser EC, Blair PA, Mauri C. Cellular targets of regulatory B cell-mediated suppression. *Mol Immunol.* (2014) 62:296–304. doi: 10.1016/j.molimm.2014.01.014

61. Yoshizaki A, Miyagaki T, DiLillo DJ, Matsushita T, Horikawa M, Kountikov EI, et al. Regulatory B cells control T-cell autoimmunity through IL-21-dependent cognate interactions. *Nature.* (2012) 491:264–8. doi: 10.1038/nature11501

62. Mann MK, Maresz K, Shriver LP, Tan Y, Dittel BN. B cell regulation of CD4+CD25+ T regulatory cells and IL-10 via B7 is essential for recovery from experimental autoimmune encephalomyelitis. *J Immunol.* (2007) 178:3447–56. doi: 10.4049/jimmunol.178.6.3447

63. Leday GG, Vértes PE, Richardson S, Greene JR, Regan T, Khan S, et al. Replicable and coupled changes in innate and adaptive immune gene expression in two case-control studies of blood microarrays in major depressive disorder. *Biol Psychiatry.* (2018) 83:70–80. doi: 10.1016/j.biopsych.2017.01.021

64. Laumet G, Edralin JD, Chiang AC, Dantzer R, Heijnen CJ, Kavelaars A. Resolution of inflammation-induced depression requires T lymphocytes and endogenous brain interleukin-10 signaling. *Neuropsychopharmacology.* (2018) 43:2597–605. doi: 10.1038/s41386-018-0154-1

65. Abeer, El-Sayed A, Ramy HA. Immunological changes in patients with mania: changes in cell mediated immunity in a sample from Egyptian patients. *Egypt J Immunol.* (2006) 13:79–85.

66. Frank MG, Wieseler Frank JL, Hendricks SE, Burke WJ, Johnson DR. Age at onset of major depressive disorder predicts reductions in NK cell number and activity. *J Affect Disord.* (2002) 71:159–67. doi: 10.1016/s0165-032700395-0

67. Maes M. Evidence for an immune response in major depression: a review and hypothesis. *Prog Neuropsychopharmacol Biol Psychiatry.* (1995) 19:11–38. doi: 10.1016/0278-584600101-m

68. Caruso C, Candore G, Cigna D, Colucci AT, Modica MA. Biological significance of soluble IL-2 receptor. *Mediat Inflamm.* (1993) 2:3–21. doi: 10.1155/S0962935193000018

69. Darko DF, Rose J, Gillin JC, Golshan S, Baird SM. Neutrophilia and lymphopenia in major mood disorders. *Psychiatry Res.* (1988) 25:243–51. doi: 10.1016/0165-178190095-9

70. Teixeira AL, Barbosa IG, Machado-Vieira R, Rizzo LB, Wieck A, Bauer ME. Novel biomarkers for bipolar disorder. *Expert Opin Med Diagn.* (2013) 7:147–59. doi: 10.1517/17530059.2013.734807

71. Marazziti, D, Torrigiani S, Carbone MG, Mucci F, Flamini W, Ivaldi T, et al. Neutrophil/lymphocyte, platelet/lymphocyte, and monocyte/lymphocyte ratios in mood disorders. *Curr Med Chem.* (2022). 58:7–12. doi: 10.2174/0929867328666210922160116

72. O'Brien SM, Scully P, Scott LV, Dinan TG. Cytokine profiles in bipolar affective disorder: focus on acutely ill patients. *J Affect Disord.* (2006) 90:263–7. doi: 10.1016/j.jad.2005.11.015

73. Ortiz-Domínguez A, Hernández ME, Berlanga C, Gutiérrez-Mora D, Moreno J, Heinze G, et al. Immune variations in bipolar disorder: phasic differences. *Bipolar Disord.* (2007) 9:596–602. doi: 10.1111/j.1399-5618.2007.00493.x

74. Lee LF, Lih CJ, Huang CJ, Cao T, Cohen SN, McDevitt HO. Genomic expression profiling of TNF-α-treated BDC2.5 diabetogenic CD4+ T cells. *Proc Natl Acad Sci U.S.A.* (2008) 105:10107–12. doi: 10.1073/pnas.0803336105

75. Akkouh IA, Skrede S, Holmgren A, Ersland KM, Hansson L, Bahrami S, et al. Exploring lithium's transcriptional mechanisms of action in bipolar disorder: a multi-step study. *Neuropsychopharmacology.* (2020) 45:947–55. doi: 10.1038/s41386-019-0556-8

76. Chiu CT, Wang Z, Hunsberger JG, Chuang DM. Therapeutic potential of mood stabilizers lithium and valproic acid: beyond bipolar disorder. *Pharmacol Rev.* (2013) 65:105–42. doi: 10.1124/pr.111.005512

77. Pietruczuk K, Lisowska KA, Grabowski K, Landowski J, Witkowski JM. Proliferation and apoptosis of T lymphocytes in patients with bipolar disorder. *Sci Rep.* (2018) 8:3327. doi: 10.1038/s41598-018-21769-0

78. Bei E, Salpeas V, Pappa D, Anagnostara C, Alevizos V, Moutsatsou P. Phosphorylation status of glucocorticoid receptor, heat shock protein 70, cytochrome c and Bax in lymphocytes of euthymic, depressed and manic bipolar patients. *Psychoneuroendocrinology.* (2009) 34:1162–75. doi: 10.1016/j.psyneuen.2009.03.002

79. Müller N, Schwarz MJ. The immune-mediated alteration of serotonin and glutamate: towards an integrated view of depression. *Mol Psychiatry.* (2007) 12:988–1000. doi: 10.1038/sj.mp.4002006

80. Fallarino F, Grohmann U, Vacca C, Bianchi R, Orabona C, Spreca A, et al. T cell apoptosis by tryptophan catabolism. *Cell Death Differ.* (2002) 9:1069–77. doi: 10.1038/sj.cdd.4401073

81. Anderson G, Maes M. Bipolar disorder: role of immune-inflammatory cytokines, oxidative and nitrosative stress and tryptophan catabolites. *Curr Psychiatry Rep.* (2015) 17:1–9. doi: 10.1007/s11920-014-0541-1

82. Haarman BC, Riemersma-Van der Lek RF, Burger H, Netkova M, Drexhage RC, Bootsman F, et al. Relationship between clinical features and inflammation-related monocyte gene expression in bipolar disorder – Towards a better understanding of psychoimmunological interactions. *Bipolar Disord.* (2014) 16:137–50. doi: 10.1111/bdi.12142

83. Stertz L, Magalhães PV, Kapczinski F. Is bipolar disorder an inflammatory condition? The relevance of microglial activation. *Curr Opin Psychiatry.* (2013) 26:19–26. doi: 10.1097/YCO.0b013e32835aa4b4

84. Beumer W, Gibney SM, Drexhage RC, Pont-Lezica L, Doorduin J, Klein HC, et al. The immune theory of psychiatric diseases: a key role for activated microglia and circulating monocytes. *J Leukoc Biol.* (2012) 92:959–75. doi: 10.1189/jlb.0212100

85. Haarman BC, Burger H, Doorduin J, Renken RJ, Sibeijn-Kuiper AJ, Marsman JB, et al. Volume, metabolites and neuroinflammation of the hippocampus in bipolar disorder – A combined magnetic resonance imaging and positron emission tomography study. *Brain Behav Immun.* (2016) 56:21–33. doi: 10.1016/j.bbi.2015.09.004

86. Beurel E, Lowell JA. Th17 cells in depression. *Brain Behav Immun.* (2018) 69:28–34. doi: 10.1016/j.bbi.2017.08.001

87. do Prado CH, Rizzo LB, Wieck A, Lopes RP, Teixeira AL, Grassi-Oliveira R, et al. Reduced regulatory T cells are associated with higher levels of Th1/Th17 cytokines and activated MAPK in type 1 bipolar disorder. *Psychoneuroendocrinology.* (2013) 38:667–76. doi: 10.1016/j.psyneuen.2012.08.005

88. de Witte L, Tomasik J, Schwarz E, Guest PC, Rahmoune H, Kahn RS, et al. Cytokine alterations in first-episode schizophrenia patients before and after antipsychotic treatment. *Schizophr Res.* (2014) 154:23–9. doi: 10.1016/j.schres.2014.02.005

89. Drexhage RC, Hoogenboezem TA, Cohen D, Versnel MA, Nolen WA, van Beveren NJ, et al. An activated set point of T-cell and monocyte inflammatory networks in recent-onset schizophrenia patients involves both pro- and anti-inflammatory forces. *Int J Neuropsychopharmacol.* (2011) 14:746–55. doi: 10.1017/S1461145710001653

90. Drexhage RC, Hoogenboezem TH, Versnel MA, Berghout A, Nolen WA, Drexhage HA. The activation of monocyte and T cell networks in patients with bipolar disorder. *Brain Behav Immun.* (2011) 25:1206–13. doi: 10.1016/j.bbi.2011.03.013

91. Kang J, Choi YJ, Kim IK, Lee HS, Kim H, Baik SH, et al. LASSO-based machine learning algorithm for prediction of lymph node metastasis in T1 colorectal cancer. *Cancer Res Treat.* (2021) 53:773–83. doi: 10.4143/crt.2020.97

92. Kaufmann CN, Lee EE, Wing D, Sutherland AN, Christensen C, Ancoli-Israel S, et al. Correlates of poor sleep based upon wrist actigraphy data in bipolar disorder. *J Psychiatr Res.* (2021) 141:385–9. doi: 10.1016/j.jpsychires.2021.06.038

93. Xu X, Wang H, Du P, Zhang F, Li S, Zhang Z, et al. A predictive nomogram for individualized recurrence stratification of bladder cancer using multiparametric MRI and clinical risk factors. *J Magn Reson Imaging.* (2019) 50:1893–904. doi: 10.1002/jmri.26749

94. Tang XR, Li YQ, Liang SB, Jiang W, Liu F, Ge WX, et al. Development and validation of a gene expression-based signature to predict distant metastasis in locoregionally advanced nasopharyngeal carcinoma: a retrospective, multicentre, cohort study. *Lancet Oncol.* (2018) 19:382–93. doi: 10.1016/S1470-204530080-9

95. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol.* (2015) 16:e173–80. doi: 10.1016/S1470-204571116-7

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 07 November 2022
DOI 10.3389/fpsyt.2022.1019618

OPEN ACCESS

EDITED BY
Magdalena Sowa-Kucma,
University of Rzeszow, Poland

REVIEWED BY
Jing Sui,
Beijing Normal University, China
Grzegorz Kreiner,
Polish Academy of Sciences (IF PAS), Poland

*CORRESPONDENCE
Guanglei Xun
xungl@163.com
Weihua Yue
dryue@bjmu.edu.cn

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Mood Disorders, a section of the journal Frontiers in Psychiatry

RECEIVED 15 August 2022
ACCEPTED 11 October 2022
PUBLISHED 07 November 2022

CITATION
Gong Y, Lu Z, Kang Z, Feng X, Zhang Y, Sun Y, Chen W, Xun G and Yue W (2022) Peripheral non-enzymatic antioxidants as biomarkers for mood disorders: Evidence from a machine learning prediction model. *Front. Psychiatry* 13:1019618. doi: 10.3389/fpsyt.2022.1019618

Peripheral non-enzymatic antioxidants as biomarkers for mood disorders: Evidence from a machine learning prediction model

Yuandong Gong[1†], Zhe Lu[2,3,4†], Zhewei Kang[2,3,4], Xiaoyang Feng[2,3,4], Yuyanan Zhang[2,3,4], Yaoyao Sun[2,3,4], Weimin Chen[1], Guanglei Xun[1*] and Weihua Yue[2,3,4,5,6*]

[1]Shandong Mental Health Center, Shandong University, Jinan, China, [2]Peking University Sixth Hospital, Peking University Institute of Mental Health, Beijing, China, [3]National Clinical Research Center for Mental Disorders, Peking University Sixth Hospital, Beijing, China, [4]NHC Key Laboratory of Mental Health, Peking University, Beijing, China, [5]PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China, [6]Chinese Institute for Brain Research, Beijing, China

Background: Oxidative stress is related to the pathogenesis of mood disorders, and the level of oxidative stress may differ between bipolar disorder (BD) and major depressive disorder (MDD). This study aimed to detect the differences in non-enzymatic antioxidant levels between BD and MDD and assess the predictive values of non-enzymatic antioxidants in mood disorders by applying a machine learning model.

Methods: Peripheral uric acid (UA), albumin (ALB), and total bilirubin (TBIL) were measured in 1,188 participants (discover cohort: 157 with BD and 544 with MDD; validation cohort: 119 with BD and 95 with MDD; 273 healthy controls). An extreme gradient boosting (XGBoost) model and a logistic regression model were used to assess the predictive effect.

Results: All three indices differed between patients with mood disorders and healthy controls; in addition, the levels of UA in patients with BD were higher than those of patients with MDD. After treatment, UA levels increased in the MDD group, while they decreased in the BD group. Finally, we entered age, sex, UA, ALB, and TBIL into the XGBoost model. The area under the curve (AUC) of the XGBoost model for distinguishing between BD and MDD reached 0.849 (accuracy = 0.808, 95% CI = 0.719–0.878) and for distinguishing between BD with depression episode (BD-D) and MDD was 0.899 (accuracy = 0.891, 95% CI = 0.856–0.919). The models were validated in the validation cohort. The most important feature distinguishing between BD and MDD was UA.

Conclusion: Peripheral non-enzymatic antioxidants, especially the UA, might be a potential biomarker capable of distinguishing between BD and MDD.

KEYWORDS

non-enzymatic antioxidants, bipolar disorder, major depressive disorder, machine learning model, uric acid

Introduction

Bipolar disorders (BD) and major depressive disorder (MDD) are serious mental disorders that are characterized by diverse clinical symptoms, with a high prevalence, and impose a heavy disease burden (1). Due to the misdiagnosis between BD and MDD, patients with BD often receive an inappropriate treatment, especially for BD starting with a depressive episode, which leads to repeated attacks (2). It is challenging for clinicians to distinguish between BD and MDD based only on their clinical symptoms. Therefore, there is an urgent need to identify a reliable, objective biomarker to differentiate BD from MDD.

Oxidative stress is critical for the normal physiological functions of the human body, but excessive levels of peroxides contribute to deleterious oxidation, consequently leading to various pathological mechanisms. The brain is a lipid-rich organ with enormous oxygen consumption and an insufficient antioxidant barrier, which makes the brain highly susceptible to deleterious oxidation. Such oxidative imbalances are linked to various psychiatric disorders, including MDD and BD (3, 4).

The oxidative stress system is ensured by a complicated antioxidant defense system (5, 6). The main enzymatic antioxidants include superoxide dismutase, catalase, glutathione transferase, and glutathione peroxidase, which play important roles in cells. Non-enzymatic antioxidants constitute the antioxidant defense in extracellular fluid; these compounds include glutathione, certain vitamins, uric acid (UA), albumin (ALB), total bilirubin (TBIL), and some metal ions (7, 8).

Numerous abnormalities in antioxidant defense have been noted in association with mood disorders. Compared with healthy controls, patients with BD were found to have increased glutathione-transferase, catalase, UA, and decreased glutathione (9, 10). MDD was associated with increased superoxide dismutase and decreased UA (11, 12). However, the results of studies on the activity of antioxidants differ significantly, and there are few studies that compared the differences in antioxidant profiles between BD and MDD. Data on UA, ALB, and TBIL are easy to obtain because they are included in the routine bloodwork examinations performed in hospitals. In addition, the three indices, jointly accounting for approximately 85% of the antioxidant capacity of plasma (13, 14), effectively represent the level of peripheral antioxidants.

Extreme gradient boosting (XGBoost) is a gradient-boosting-based algorithm that employs increasing accurate approximations to find the best models and applies the advanced regularization technique. It enhances the model training speed and generalization and reduces the model complexity. However, the XGBoost method has some limitations. For example, XGBoost does not perform well on high-dimensional sparse features or unstructured data; in such a case, there are too many parameters, and the parameter optimization is too complex. In the present study, the XGBoost method is suitable because the clinical data are two-dimensional structured data. We first investigated the differences in non-enzymatic antioxidants between BD and MDD. Then, we assessed the predictive value of non-enzymatic antioxidants in mood disorders by applying a machine learning model, and we also used the logistic regression method to confirm the strength of the XGBoost prediction model.

In the present study, we have two hypotheses. First, we hypothesize that the differences in three peripheral non-enzymatic antioxidants between BD and MDD are significant. Second, we hypothesize that peripheral non-enzymatic antioxidants can efficiently distinguish between BD and MDD.

Materials and methods

Participants

All procedures involving human subjects were approved by the Clinical Research Ethics Committee of Shandong Mental Health Center ([2021] (Research) Ethical Review No. [66]), and the protocol was compliant with the Code of Ethics of the World Medical Association (Declaration of Helsinki). Informed written consent was obtained from all participants.

The inclusion criteria for the discovery cohort were as follows: 1) a diagnosis of BD or MDD based on the International Classification of Diseases, 10th version (ICD-10); 2) Having at least one completed test of UA, ALB, and TBIL during the first 3 days of hospitalization. The HCs had no mental disorders and a family history of mental disorders.

Subjects in the validation cohort were chosen from our previous study which has been published (15); the detailed information is shown in the Supplementary materials.

Data extraction

The data of the discovery cohort were obtained from the electronic medical record system of Shandong Mental Health Center. It included sex, age, diagnosis, and laboratory examination results (including ALB, TBIL, and UA). The discharge diagnosis was defined as the main diagnosis, and all the diagnoses were confirmed by third-level ward rounds (during the first week of hospitalization) according to the criteria of the hospitalization procedure. Any change in diagnosis was recorded, and the discharge diagnosis was consistent with the changed diagnosis. We extracted the first four consecutive test results of ALB, TBIL, and UA of patients with BD and MDD during a hospital stay. All the data were anonymous. During the data extraction, all the data were stripped of personal identifiers and instead labeled with ID numbers; personnel with access to the ID numbers could not identify any of the participants in this study.

Statistical analysis

The analyses were conducted by using IBM SPSS Statistics, Version 26 (Chicago Inc., USA) and R software. All measurement data were inspected for normality with the Kolmogorov–Smirnov test. A Chi-square test was conducted to analyze the sex distributions. Differences in ALB, TBIL, and UA among the BD, MDD, and HC groups were tested by analysis of covariance (ANCOVA), with age and sex as covariates; a Bonferroni test was used to identify the pairwise differences between groups. A generalized estimation equation was used to compare the different post-treatment trends of non-enzymatic antioxidants between BD and MDD. We used XGBoost to build a machine learning model. Then, we used 10-fold cross-validation repeated 10 times to reduce underfitting or overfitting and used the grid search method to optimize the model hyperparameters. The main metrics used to evaluate classification performance were accuracy and the area under the curve (AUC).

Results

Demographic data

A total of 157 patients with BD (55.4% male participants; mean age: 31.16 ± 14.25 years) and 544 patients with MDD (33.8% male participants; mean age: 21.15 ± 13.07 years) were included in the discovery cohort. A total of 119 patients with BD (51.26% men; mean age: 31.91 ± 11.68 years) and 95 patients with MDD (45.26% men; mean age: 37.63 ± 13.41 years) were included in the validation cohort. This study also included 273 HCs (34.1% men; mean age: 37.92 ± 9.22 years). All the participants received treatment, and the differences in medications (including antidepressants, antipsychotics, and mood stabilizers) between the BD and MDD groups were significant (Table 1). Detailed information on the medication use of the participants is provided in the Supplementary Table S1.

Peripheral non-enzymatic antioxidants at baseline in the BD, MDD, and HC groups

At baseline, the differences in three non-enzymatic antioxidants among the BD, MDD, and HC groups were significant (UA, $F = 23.512$, $p < 0.001$; ALB, $F = 131.385$, $p < 0.001$; TBIL, $F = 31.343$, $p < 0.001$). The UA levels in the BD group were higher than those in the MDD group and HC group ($p < 0.001$), and the UA levels in the MDD group were also higher than those in the HC group ($p < 0.001$). The ALB levels of the BD group and MDD group were lower than those of the HC group ($p < 0.001$), while there were no significant differences in ALB levels between the BD group and MDD group ($p = 0.179$). The TBIL levels of the BD group and MDD group were higher than those of the HC group ($p < 0.001$), while the differences in TBIL levels between the BD group and MDD group were not significant ($p = 0.565$; Figure 1).

Then, we divided the BD group into the bipolar disorder with mania/hypomania/mixed episode (BD-M) subgroup and the bipolar disorder with depression episode (BD-D) subgroup. The results showed that UA levels in the BD-M subgroup were higher than those in the BD-D subgroup and MDD group ($p < 0.001$), but the differences in UA levels between the BD-D subgroup and the MDD group were not significant. ALB levels of the BD-M subgroup were higher than the MDD group ($P = 0.043$), while the differences between the BD-D subgroup and the MDD group were not significant. There were no significant differences in TBIL among the BD-M subgroup, BD-D subgroup, and MDD group (Figure 1).

In the validation cohort, similar results were found, while the UA levels of the BD-D subgroup were higher than those of the MDD group in the validation cohort. The detailed statistical results are shown in Supplemetary Tables S2–S4.

Changes in three non-enzymatic antioxidants after treatment

The UA levels in the BD group decreased after treatment, while the UA levels in the MDD group increased after treatment. The changing trend of UA was significant (MDD and V1 as a reference, $Wald\ Chi\text{-}Square_{\ diagnosis} = 4.244$, $p = 0.039$, $Wald\ Chi\text{-}Square_{diagnosis*time} = 5.714$, $p = 0.017$; age, sex, number of antidepressants, antipsychotics, and mood stabilizers were set as covariates). The ALB levels in the MDD group decreased after treatment, while there was no significant change in ALB levels in the BD group after treatment. The TBIL levels of both groups decreased after treatment. The changing trends of ALB and TBIL levels between the MDD group and the BD group were not significant (age, sex, number of antidepressants, antipsychotics, and mood stabilizers were set as covariates, Figure 2). The detailed statistical results are shown in Supplementary Tables S5–S7.

Effectiveness of non-enzymatic antioxidants and clinical data in distinguishing between BD and MDD

First, we used age, sex, and three non-enzymatic antioxidants to distinguish between the mood disorders and HC groups. The results showed that the AUC of the XGBoost model for distinguishing between BD and HC groups was 0.943 (accuracy = 0.905, 95% CI = 0.804–0.964), and the AUC for distinguishing between MDD and HC groups

TABLE 1 Medication information of BD and MDD groups.

Medications		BD ($n = 157$)	MDD ($n = 544$)	X^2	P
Antidepressants	Prescribed	49	508	288.530	<0.001
	Non-prescribed	108	36		
Antipsychotics	Prescribed	151	461	14.375	<0.001
	Non-prescribed	6	83		
Mood stabilizers	Prescribed	151	319	77.708	<0.001
	Non-prescribed	6	225		

BD, bipolar disorders; MDD, major depressive disorder.

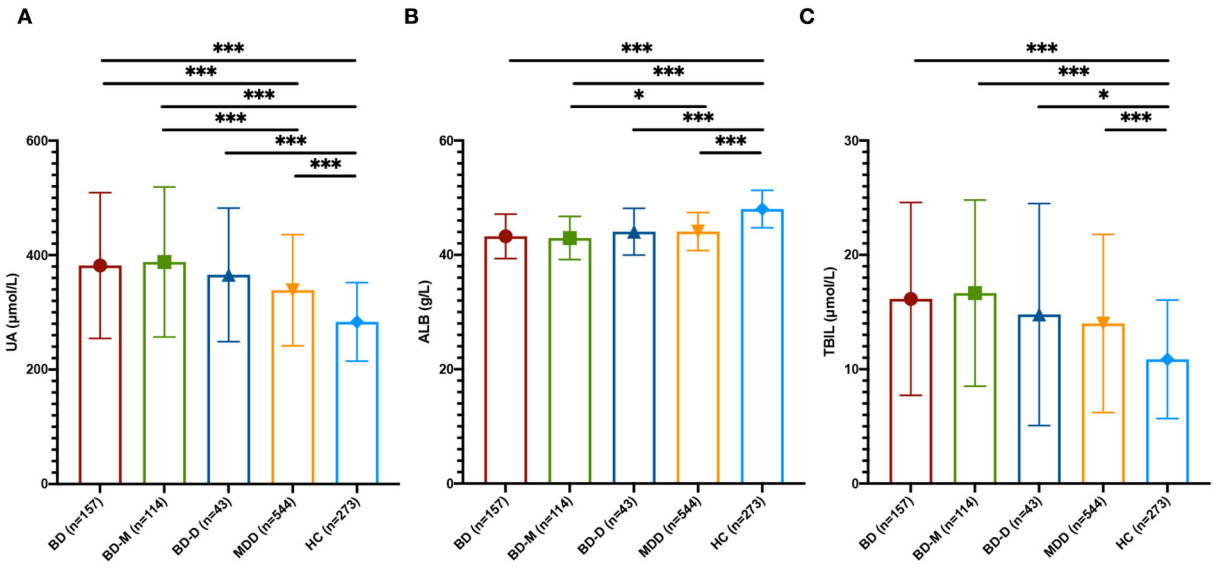


FIGURE 1
Peripheral non-enzymatic antioxidants at baseline among BD, MDD, and HC. **(A)** UA levels among BD, MDD, and HC groups. **(B)** ALB levels among BD, MDD, and HC groups. **(C)** TBIL levels among BD, MDD, and HC groups. UA, uric acid; ALB, albumin; TBIL, total bilirubin; BD, bipolar disorder; BD-M, bipolar disorder with mania/hypomania/mixed episode; BD-D, bipolar disorder with depression episode; MDD, major depressive disorder; HC, healthy control; *$p < 0.05$; ***$p < 0.001$.

was 0.990 (accuracy = 0.909, 95% CI = 0.843–0.954). Then, we used these data to distinguish between BD and MDD, as well as between BD-D and MDD. The AUC of the XGBoost model for distinguishing between BD and MDD groups was 0.849 (accuracy = 0.808, 95% CI = 0.719–0.878), and the AUC for distinguishing between BD-D subgroup and MDD group was 0.899 (accuracy = 0.891, 95% CI = 0.856–0.919). Finally, we used an independent cohort, including 119 subjects with BD, and 95 subjects with MDD, as a validation cohort to verify our results. Detailed information on the validation cohort has been published in our previous study (15). The above results were confirmed. The AUC of the XGBoost model in validation cohort for distinguishing between BD and HC groups was 0.982 (accuracy = 0.947, 95% CI = 0.854–0.989), the AUC for distinguishing between MDD and HC was groups 0.971 (accuracy = 0.889, 95% CI = 0.774–0.958), the AUC for distinguishing between BD and MDD groups was 0.781 (accuracy = 0.706, 95% CI = 0.597–0.800), the AUC for distinguishing between BD-D and MDD groups was 0.781 (accuracy = 0.633, 95% CI = 0.499–0.754). The most important feature for distinguishing between BD and MDD was UA in both the discovery cohort and validation cohort (Figure 3).

The XGBoost model showed better classifying ability than the logistic regression model (Table 2).

Discussion

In this study, we compared three peripheral non-enzymatic antioxidants among patients with BD, MDD, and HCs. The result showed that the differences in all 3 indices of interest

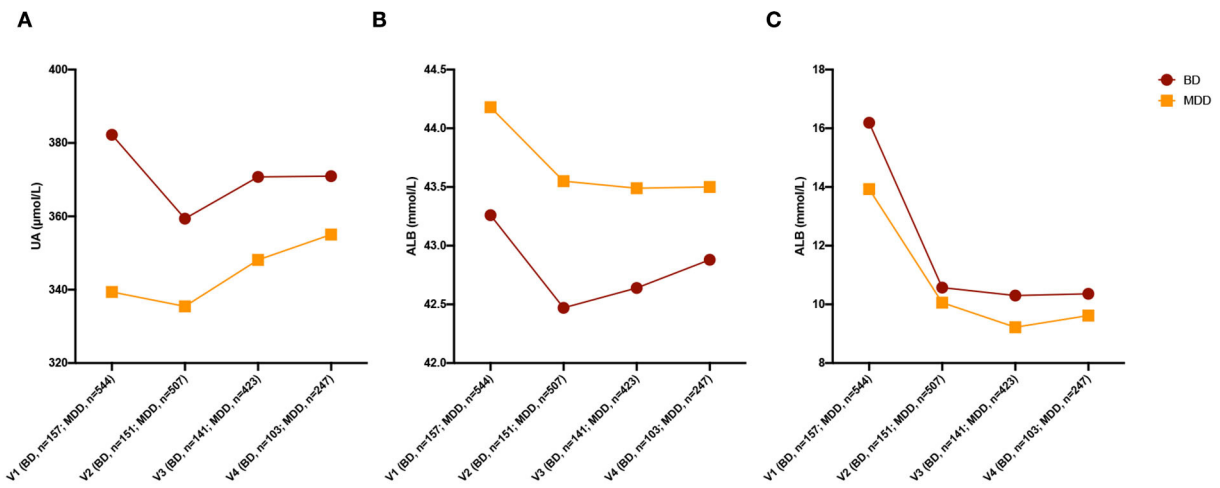


FIGURE 2
Changing of three non-enzymatic antioxidants after treatment. **(A)** Changing of UA levels after treatment. **(B)** Changing of ALB levels after treatment. **(C)** Changing of TBIL levels after treatment. UA, uric acid; ALB, albumin; TBIL, total bilirubin; BD, bipolar disorder; MDD, major depressive disorder; HC, healthy control.

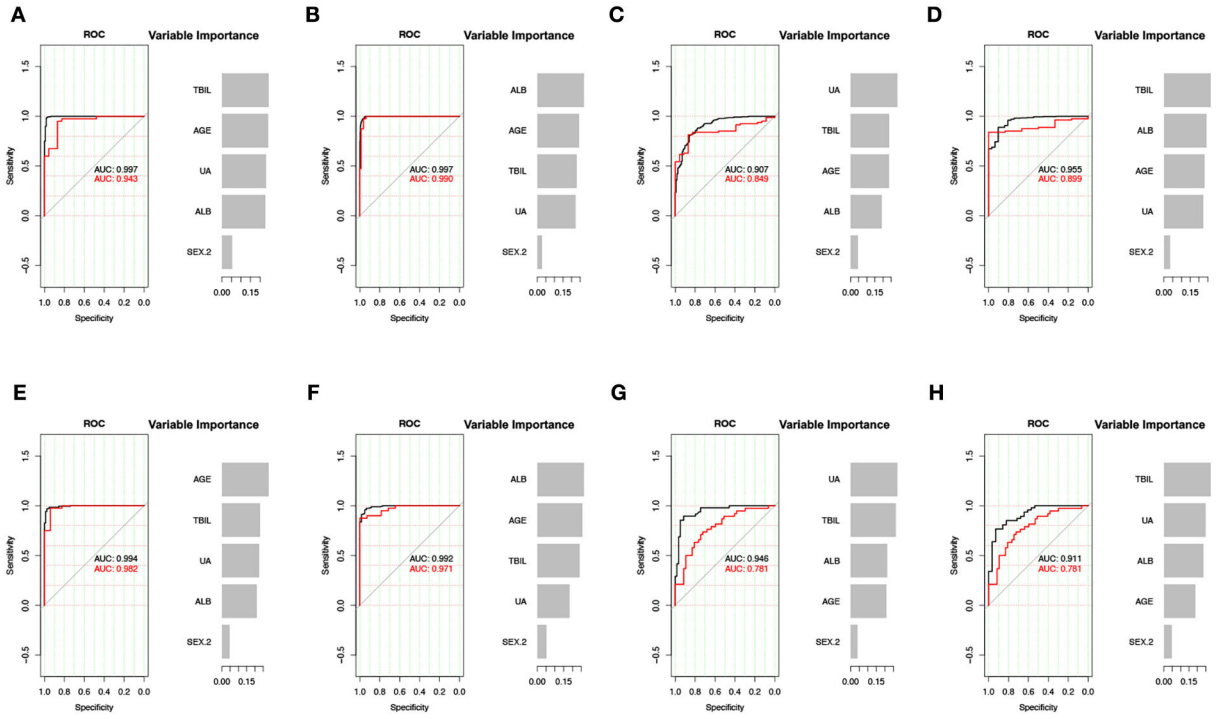


FIGURE 3
XGBoost model for the predictive effect of non-enzymatic antioxidants. **(A–D)** The result of discovery data. **(E–H)** The result of validation data. **(A,E)** BD vs. HC. **(B,F)** MDD vs. HC. **(C,G)** BD vs. MDD. **(D,H)** BD-D vs. MDD. UA, uric acid; ALB, albumin; TBIL, total bilirubin; BD, bipolar disorder; MDD, major depressive disorder; BD-D, bipolar disorder with depression episode; HC, healthy control; ROC, receiver operating characteristic curve; AUC, area under the curve.

among 3 groups were significant; the UA and TBIL levels of the BD and MDD groups were higher than those of the HC group, while their ALB levels were lower than those of the HC group. These results indicated that mood disorders were associated with dysfunction of the peripheral antioxidation system.

TABLE 2 AUC of two prediction models.

	BD vs. HC ($n = 157$ vs. $n = 273$)	MDD vs. HC (n = 544 vs. $n = 273$)	BD vs. MDD ($n = 157$ vs. $n = 544$)	BD-D vs. MDD ($n = 43$ vs. $n = 544$)
Discovery cohort				
Logistic regression	0.971	0.960	0.729	0.710
XGBoost	0.943	0.990	0.849	0.899
	BD vs. HC ($n = 119$ vs. $n = 273$)	**MDD vs. HC ($n = 95$ vs. $n = 273$)**	**BD vs. MDD ($n = 119$ vs. $n = 95$)**	**BD-D vs. MDD ($n = 55$ vs. $n = 95$)**
Validation cohort				
Logistic regression	0.966	0.963	0.759	0.698
XGBoost	0.982	0.971	0.781	0.781

BD, bipolar disorder; MDD, major depressive disorder; BD-D, bipolar disorder with depression episode; HC, healthy control; AUC, area under the curve.

Mitochondria are able to reduce deleterious oxidation, but this ability is not sufficient to neutralize all oxidative stress. Therefore, antioxidants are necessary to prevent excessive oxidative damage. Enzymatic antioxidants can inhibit the genesis of peroxide and remove excessive reactive oxygen species (16). Non-enzymatic antioxidants also play a key role in the antioxidant system and can chelate transition metals and interact with reactive oxygen species by breaking free radical chain reactions (17). There are many non-enzymatic antioxidants in plasma, such as UA, ALB, TBIL, zinc, tocopherol, ascorbate, and retinol. In addition to the three non-enzymatic antioxidants, other compounds also have important biological functions. For instance, retinol is essential for embryonic development, especially for the development of the brain, meanwhile, the ascorbate, as a neuroprotective compound, is highly concentrated in the brain and regulates the function of neurons and synapses (12, 18, 19). Abnormal levels of these antioxidants in mood disorders were widely discussed in previous studies, and some meta-analyses were conducted, which showed that their results were consistent with our findings (10, 11). However, many non-enzymatic antioxidants are trace elements and are not routinely measured in the clinic. In contrast, the three indices detected in this study, which accounted for approximately 85% of the antioxidant capacity of plasma, could represent the level of peripheral antioxidants. In addition, they are routinely examined during hospitalization and are easy to obtain (14).

Uric acid, as a selective antioxidant, can scavenge reactive oxygen or nitrogen and prevent the erythrocyte membrane from lipid peroxidation by reacting with peroxides (20). Moreover, it is the end-product of the purinergic system, which is involved in the pathophysiology of mental disorders *via* influencing cell proliferation, neuronal differentiation, and neuroglial cell inflammation (21, 22). UA is also associated with sleep, cognition, appetite, social interaction, etc. (23, 24). In this study, UA levels in the BD group were found to be higher than those in the MDD group, which suggested that BD might have a more severe imbalance of redox homeostasis and that UA might be a potential biomarker to distinguish between BD and MDD. Then, we divided the BD group into BD-M and BD-D subgroups. The UA levels of BD-M were markedly higher than those of MDD, while the differences in UA between BD-D and MDD were not significant, implying that UA might be a status indicator for BD. Therefore, we checked the changing trend of UA, which decreased after treatment in the BD group, in contrast, UA increased after treatment in the MDD group, which supported UA as a potential biomarker to distinguish BD from MDD. Previous studies also showed that the UA of the BD group was higher than MDD and HC groups (25–27), and the UA levels of BD-M were higher than those of BD-D and euthymic stage (28–30), while the results regarding UA in MDD were inconsistent. Several studies reported that MDD was associated with lower UA levels than HC and BD (31), while there was a study that showed no significant differences in UA between MDD and HC groups (32). This study showed that the UA levels of MDD were higher than those of HC, which might result from the heterogeneity of the participants. This study aimed to explore peripheral antioxidants of mood disorders in a real-world setting, and we set sex, age, and medication use as covariates to control for confounding factors. Although the diet might significantly affect the level of the UA, all the participants were hospitalized and their diets were essentially the same. In our previous study, age was negatively associated with UA levels, and the age of the MDD group was lower than that of HC, which might be a reason for higher levels of UA in MDD than in HC (15). The consistent results of UA in BD indicated that UA was a reliable biomarker for BD.

Albumin is an endogenous antioxidant with a radical scavenging function, binding metal ions, and responsible for reactive oxygenated radical species. Furthermore, ALB, which

plays a role in the inflammation and the immune system, has also been demonstrated to be involved in the pathogenesis of mental disorders. Bilirubin plays a role in antioxidative mechanisms by efficiently scavenging peroxyl radicals and acting as a chain-breaking antioxidant. Besides the antioxidant ability of TBIL, it has toxic effects on the brain, and a great deal of evidence indicates that it is related to cognitive function. Abnormal levels of TBIL were also found in mental disorders (33). Differences in ALB and TBIL between BD and MDD groups were neither significant nor were the changing trends of these two indices after treatment. However, the ALB in the BD-M subgroup was lower than those in the MDD group, implying that patients with mania episodes might have more severe dysfunction of the antioxidation system than patients with MDD.

Previous studies indicated that the antidepressants (such as escitalopram), mood stabilizers (such as lithium and valproate), and antipsychotics (such as olanzapine and clozapine) may have neuroprotective effects against oxidative stress (34–40), while antioxidative effects differed among different types of medications, and the conclusion was not consistent (41, 42). In this study, we set the numbers of antidepressants, mood stabilizers, and antipsychotics as covariates when detecting the changing trend after treatment to avoid confounding factors, and the results were still significant.

Misdiagnosis is common in BD since it often starts with a depression episode (2). It leads to inappropriate treatment, switching to mania/hypomania, and repeated attacks (43). Based on the above findings, non-enzymatic antioxidants might be potential biomarkers. We applied age, sex, and three non-enzymatic antioxidants to distinguish between BD and MDD. The AUC was 0.849, and UA was the most important feature for distinguishing between BD and MDD, which was used to confirm the results.

Some researchers believe that both MDD and BD belong to the same mood disorder spectrum; they used novel terms such as bipolar spectrum disorders to describe these conditions (44). Some researchers hold different views, and they attempted to distinguish between patients with BD and MDD by their personal characteristics using a "softer bipolar spectrum," including the onset age, temperament, and response to antidepressants, and their concepts were validated by several studies (45–48). In addition, in contrast to the fourth version of the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV), DSM-5 divides mood disorders into two chapters (bipolar disorders and depressive disorder), which confirms that BD is distinct from MDD.

There are some strengths and limitations to our study. The first advantage was the real-world measures used in the study that increased the clinical transferability of our results. Furthermore, the study also analyzed the changes in peripheral non-enzymatic antioxidants after treatment. The third advantage is that we applied a machine learning model, namely, the XGBoost model, to explore the predictive effect of non-enzymatic antioxidants, and the results showed that machine learning could improve classification ability. Nevertheless, some limitations should be discussed. First of all, the retrospective design of the study did not allow us to assess the severity of symptoms; in addition, the diagnosis could not be confirmed because of the short-observation period. The second limitation was, although all the participants were inpatients who received uniform diets provided by the hospital, we did not take into account certain confounding factors that may influence the levels of peripheral antioxidants, such as smoking and body mass index into account. Finally, we explored the change after treatment. Although we set the numbers of medications as covariates to prevent them from acting as confounders, the types of medications were too complex, so the detailed types of medications were not taken into account. In the future, a rigorously designed head-to-head randomized clinical trial should be conducted to explore the antioxidative effects of different medications.

In conclusion, the dysfunction of the peripheral non-enzymatic antioxidant system might be involved in the pathogenesis of mood disorders, and UA might be used as a potential biomarker to distinguish between BD and MDD.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Clinical Research Ethics Committee of Shandong Mental Health Center. The patients/participants provided their written informed consent to participate in this study.

Author contributions

YG, ZL, GX, and WY designed the research. ZL, ZK, XF, and WC extracted and cleaned the data. ZL, YZ, and YS performed the statistical analysis. ZL wrote the original draft. YG, GX, and WY reviewed and edited the final manuscript. All authors contributed to and have approved the final manuscript.

Funding

This study was supported by the National Natural Science Foundation of China (81825009); National Key R&D Program of China (2021YFF1201103); Academy of Medical Sciences

Research Unit (2019-I2M-5-006); Chinese Institute for Brain Research at Beijing (2020-NKX-XM-12); PKUHSC-KCL Joint Medical Research (BMU2020KCL001) (WY). This study was also supported by Shandong Province Key R&D Program (Science and Technology Demonstration Project) Project (2021SFGC0504) (YG). This study was also supported by the Innovation Fund for Outstanding Doctoral Students of Peking University Health Science Center (ZL).

Conflict of interest

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2022.1019618/full#supplementary-material

References

1. Huang Y, Wang Y, Wang H, Liu Z, Yu X, Yan J, et al. Prevalence of mental disorders in China: a cross-sectional epidemiological study. *Lancet Psychiatry.* (2019) 6:211–24. doi: 10.1016/S2215-0366(18)30511-X

2. Tondo L, Visioli C, Preti A, Baldessarini RJ. Bipolar disorders following initial depression: modeling predictive clinical factors. *J Affect Disord.* (2014) 167:44–9. doi: 10.1016/j.jad.2014.05.043

3. Cecerska-Heryć E, Polikowska A, Serwin N, Roszak M, Grygorcewicz B, Heryć R, et al. Importance of oxidative stress in the pathogenesis, diagnosis, and monitoring of patients with neuropsychiatric disorders, a review. *Neurochem Int.* (2022) 153:105269. doi: 10.1016/j.neuint.2021.105269

4. Lu Z, Pu C, Zhang Y, Sun Y, Liao Y, Kang Z, et al. Oxidative stress and psychiatric disorders: evidence from the bidirectional mendelian randomization study. *Antioxidants (Basel).* (2022) 11:1386. doi: 10.3390/antiox11071386

5. Nordberg J, Arnér ES. Reactive oxygen species, antioxidants, and the mammalian thioredoxin system. *Free Radic Biol Med.* (2001) 31:1287–312. doi: 10.1016/S0891-5849(01)00724-9

6. Poljsak B, Šuput D, Milisav I. Achieving the balance between ROS and antioxidants: when to use the synthetic antioxidants. *Oxid Med Cell Longev.* (2013) 2013:956792. doi: 10.1155/2013/956792

7. Yao JK, Keshavan MS. Antioxidants, redox signaling, and pathophysiology in schizophrenia: an integrative view. *Antioxid Redox Signal.* (2011) 15:2011–35. doi: 10.1089/ars.2010.3603

8. Pisoschi AM, Pop A. The role of antioxidants in the chemistry of oxidative stress: a review. *Eur J Med Chem.* (2015) 97:55–74. doi: 10.1016/j.ejmech.2015.04.040

9. Jiménez-Fernández S, Gurpegui M, Garrote-Rojas D, Gutiérrez-Rojas L, Carretero MD, Correll CU. Oxidative stress parameters and antioxidants in patients with bipolar disorder: results from a meta-analysis comparing patients, including stratification by polarity and euthymic status, with healthy controls. *Bipolar Disord.* (2021) 23:117–29. doi: 10.1111/bdi.12980

10. Brown NC, Andreazza AC, Young LT. An updated meta-analysis of oxidative stress markers in bipolar disorder. *Psychiatry Res.* (2014) 218:61–8. doi: 10.1016/j.psychres.2014.04.005

11. Jiménez-Fernández S, Gurpegui M, Díaz-Atienza F, Pérez-Costillas L, Gerstenberg M, Correll CU. Oxidative stress and antioxidant parameters in patients with major depressive disorder compared to healthy controls before and after antidepressant treatment: results from a meta-analysis. *J Clin Psychiatry.* (2015) 76:1658–67. doi: 10.4088/JCP.14r09179

12. Liu T, Zhong S, Liao X, Chen J, He T, Lai S, et al. A meta-analysis of oxidative stress markers in depression. *PLoS ONE.* (2015) 10:e0138904. doi: 10.1371/journal.pone.0138904

13. Miller NJ, Rice-Evans C, Davies MJ, Gopinathan V, Milner A. A novel method for measuring antioxidant capacity and its application to monitoring the antioxidant status in premature neonates. *Clin Sci (Lond).* (1993) 84:407–12. doi: 10.1042/cs0840407

14. Maxwell SR, Dietrich T, Chapple IL. Prediction of serum total antioxidant activity from the concentration of individual serum antioxidants. *Clin Chim Acta.* (2006) 372:188–94. doi: 10.1016/j.cca.2006.04.015

15. Lu Z, Wang Y, Xun G. Individuals with bipolar disorder have a higher level of uric acid than major depressive disorder: a case-control study. *Sci Rep.* (2021) 11:18307. doi: 10.1038/s41598-021-97955-4

16. Salim S. Oxidative stress and psychological disorders. *Curr Neuropharmacol.* (2014) 12:140–7. doi: 10.2174/1570159X11666131120230309

17. Salim S. Oxidative stress and the central nervous system. *J Pharmacol Exp Ther.* (2017) 360:201–5. doi: 10.1124/jpet.116.237503

18. Reay WR, Cairns MJ. The role of the retinoids in schizophrenia: genomic and clinical perspectives. *Mol Psychiatry.* (2020) 25:706–18. doi: 10.1038/s41380-019-0566-2

19. Moretti M, Rodrigues ALS. Functional role of ascorbic acid in the central nervous system: a focus on neurogenic and synaptogenic processes. *Nutr Neurosci.* (2021) 2021:1–11. doi: 10.1080/1028415X.2021.1956848

20. Sautin YY, Johnson RJ. Uric acid: the oxidant-antioxidant paradox. *Nucleosides Nucleotides Nucleic Acids.* (2008) 27:608–19. doi: 10.1080/15257770802138558

21. Burnstock G, Krügel U, Abbracchio MP, Illes P. Purinergic signalling: from normal behaviour to pathological brain function. *Prog Neurobiol.* (2011) 95:229–74. doi: 10.1016/j.pneurobio.2011.08.006

22. Burnstock G. Introductory overview of purinergic signalling. *Front Biosci (Elite Ed).* (2011) 3:896–900. doi: 10.2741/e298

23. Machado-Vieira R, Lara DR, Souza DO, Kapczinski F. Purinergic dysfunction in mania: an integrative model. *Med Hypotheses.* (2002) 58:297–304. doi: 10.1054/mehy.2001.1543

24. Sutin AR, Cutler RG, Camandola S, Uda M, Feldman NH, Cucca F, et al. Impulsivity is associated with uric acid: evidence from humans and mice. *Biol Psychiatry.* (2014) 75:31–7. doi: 10.1016/j.biopsych.2013.02.024

25. Bartoli F, Crocamo C, Gennaro GM, Castagna G, Trotta G, Clerici M, et al. Exploring the association between bipolar disorder and uric acid: a mediation analysis. *J Psychosom Res.* (2016) 84:56–9. doi: 10.1016/j.jpsychores.2016.03.014

26. Kesebir S, Süner O, Yaylaci ET, Bayrak A, Turan C. Increased uric acid levels in bipolar disorder: is it trait or state? *J Biol Regul Homeost Agents.* (2013) 27:981–8. Available online at: https://www.biolifesas.org/EN/Y2013/V27/I4/981

27. Salvadore G, Viale CI, Luckenbaugh DA, Zanatto VC, Portela LV, Souza DO, et al. Increased uric acid levels in drug-naïve subjects with bipolar disorder during a first manic episode. *Prog Neuropsychopharmacol Biol Psychiatry.* (2010) 34:819–21. doi: 10.1016/j.pnpbp.2010.02.027

28. De Berardis D, Conti CM, Campanella D, Carano A, Di Giuseppe B, Valchera A, et al. Evaluation of plasma antioxidant levels during different phases

of illness in adult patients with bipolar disorder. *J Biol Regul Homeost Agents.* (2008) 22:195–200. Available online at: https://www.biolifesas.org/EN/Y2008/V22/I3/195

29. Muti M, Del Grande C, Musetti L, Marazziti D, Turri M, Cirronis M, et al. Serum uric acid levels and different phases of illness in bipolar I patients treated with lithium. *Psychiatry Res.* (2015) 225:604–8. doi: 10.1016/j.psychres.2014.11.038

30. Albert U, De Cori D, Aguglia A, Barbaro F, Bogetto F, Maina G. Increased uric acid levels in bipolar disorder subjects during different phases of illness. *J Affect Disord.* (2015) 173:170–5. doi: 10.1016/j.jad.2014.11.005

31. Wen S, Cheng M, Wang H, Yue J, Wang H, Li G, et al. Serum uric acid levels and the clinical characteristics of depression. *Clin Biochem.* (2012) 45:49–53. doi: 10.1016/j.clinbiochem.2011.10.010

32. Wiener C, Rassier GT, Kaster MP, Jansen K, Pinheiro RT, Klamt F, et al. Gender-based differences in oxidative stress parameters do not underlie the differences in mood disorders susceptibility between sexes. *Eur Psychiatry.* (2014) 29:58–63. doi: 10.1016/j.eurpsy.2013.05.006

33. Huang F, Pariante CM, Borsini A. From dried bear bile to molecular investigation: a systematic review of the effect of bile acids on cell apoptosis, oxidative stress and inflammation in the brain, across pre-clinical models of neurological, neurodegenerative and neuropsychiatric disorders. *Brain Behav Immun.* (2022) 99:132–46. doi: 10.1016/j.bbi.2021.09.021

34. Rizak J, Tan H, Zhu H, Wang JF. Chronic treatment with the mood-stabilizing drug lithium up-regulates nuclear factor E2-related factor 2 in rat pheochromocytoma PC12 cells in vitro. *Neuroscience.* (2014) 256:223–9. doi: 10.1016/j.neuroscience.2013.10.036

35. Khairova R, Pawar R, Salvadore G, Juruena MF, de Sousa RT, Soeiro-de-Souza MG, et al. Effects of lithium on oxidative stress parameters in healthy subjects. *Mol Med Rep.* (2012) 5:680–2. doi: 10.3892/mmr.2011.732

36. Frey BN, Valvassori SS, Réus GZ, Martins MR, Petronilho FC, Bardini K, et al. Effects of lithium and valproate on amphetamine-induced oxidative stress generation in an animal model of mania. *J Psychiatry Neurosci.* (2006) 31:326–32. doi: 10.1007/s11064-006-9070-6

37. Brinholi FF, Farias CC, Bonifácio KL, Higachi L, Casagrande R, Moreira EG, et al. Clozapine and olanzapine are better antioxidants than haloperidol, quetiapine, risperidone and ziprasidone in *in vitro* models. *Biomed Pharmacother.* (2016) 81:411–5. doi: 10.1016/j.biopha.2016.02.047

38. Dietrich-Muszalska A, Kolińska-Łukaszuk J. Comparative effects of aripiprazole and selected antipsychotic drugs on lipid peroxidation in plasma. *Psychiatry Clin Neurosci.* (2018) 72:329–36. doi: 10.1111/pcn.12631

39. Dionisie V, Ciobanu AM, Toma VA, Manea MC, Baldea I, Olteanu D, et al. Escitalopram targets oxidative stress, caspase-3, BDNF and MeCP2 in the hippocampus and frontal cortex of a rat model of depression induced by chronic unpredictable mild stress. *Int J Mol Sci.* (2021) 22:7483. doi: 10.3390/ijms22147483

40. Shetty S, Hariharan A, Shirole T, Jagtap AG. Neuroprotective potential of escitalopram against behavioral, mitochondrial and oxidative dysfunction induced by 3-nitropropionic acid. *Ann Neurosci.* (2015) 22:11–8. doi: 10.5214/ans.0972.7531.220104

41. Dietrich-Muszalska A, Kontek B, Rabe-Jabłońska J. Quetiapine, olanzapine and haloperidol affect human plasma lipid peroxidation in vitro. *Neuropsychobiology.* (2011) 63:197–201. doi: 10.1159/000321623

42. Jajoo A, Donlon C, Shnayder S, Levin M, McVey M. Sertraline induces DNA damage and cellular toxicity in Drosophila that can be ameliorated by antioxidants. *Sci Rep.* (2020) 10:4512. doi: 10.1038/s41598-020-61362-y

43. Stahl SM, Morrissette DA, Faedda G, Fava M, Goldberg JF, Keck PE, et al. Guidelines for the recognition and management of mixed depression. *CNS Spectr.* (2017) 22:203–19. doi: 10.1017/S1092852917000165

44. Ghaemi SN, Ko JY, Goodwin FK. "Cade's disease" and beyond: misdiagnosis, antidepressant use, and a proposed definition for bipolar spectrum disorder. *Can J Psychiatry.* (2002) 47:125–34. doi: 10.1177/070674370204700202

45. Perugi G, Akiskal HS. The soft bipolar spectrum redefined: focus on the cyclothymic, anxious-sensitive, impulse-dyscontrol, and binge-eating connection in bipolar II and related conditions. *Psychiatr Clin North Am.* (2002) 25:713–37. doi: 10.1016/S0193-953X(02)00023-0

46. Akiskal HS, Pinto O. The evolving bipolar spectrum. Prototypes I, II, III, IV. *Psychiatr Clin North Am.* (1999) 22:517–34, vii. doi: 10.1016/S0193-953X(05)70093-9

47. Akiskal HS, Akiskal KK, Lancrenon S, Hantouche E. Validating the soft bipolar spectrum in the French National EPIDEP Study: the prominence of BP-II 1/2. *J Affect Disord.* (2006) 96:207–13. doi: 10.1016/j.jad.2006.03.011

48. Akiskal HS, Akiskal KK, Lancrenon S, Hantouche EG, Fraud JP, Gury C, et al. Validating the bipolar spectrum in the French National EPIDEP Study: overview of the phenomenology and relative prevalence of its clinical prototypes. *J Affect Disord.* (2006) 96:197–205. doi: 10.1016/j.jad.2006.05.015

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 15 December 2022
DOI 10.3389/fpsyt.2022.1070198

OPEN ACCESS

EDITED BY
Zhi Xu,
Southeast University, China

REVIEWED BY
Feng Liu,
Tianjin Medical University General Hospital, China
Renping Yu,
Zhengzhou University, China

*CORRESPONDENCE
Zhengxia Wang
zxiawang@hainanu.edu.cn

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Computational Psychiatry, a section of the journal Frontiers in Psychiatry

RECEIVED 14 October 2022
ACCEPTED 28 November 2022
PUBLISHED 15 December 2022

CITATION
Guo L, Zhang Y, Liu Q, Guo K and Wang Z (2022) Multi-band network fusion for Alzheimer's disease identification with functional MRI. *Front. Psychiatry* 13:1070198. doi: 10.3389/fpsyt.2022.1070198

Multi-band network fusion for Alzheimer's disease identification with functional MRI

Lingyun Guo†, Yangyang Zhang†, Qinghua Liu, Kaiyu Guo and Zhengxia Wang* for the Alzheimer's Disease Neuroimaging Initiative

School of Computer Science and Technology, Hainan University, Haikou, China

Introduction: The analysis of functional brain networks (FBNs) has become a promising and powerful tool for auxiliary diagnosis of brain diseases, such as Alzheimer's disease (AD) and its prodromal stage. Previous studies usually estimate FBNs using full band Blood Oxygen Level Dependent (BOLD) signal. However, a single band is not sufficient to capture the diagnostic and prognostic information contained in multiple frequency bands.

Method: To address this issue, we propose a novel multi-band network fusion framework (MBNF) to combine the various information (e.g., the diversification of structural features) of multi-band FBNs. We first decompose the BOLD signal adaptively into two frequency bands named high-frequency band and low-frequency band by the ensemble empirical mode decomposition (EEMD). Then the similarity network fusion (SNF) is performed to blend two networks constructed by two frequency bands together into a multi-band fusion network. In addition, we extract the features of the fused network towards a better classification performance.

Result: To verify the validity of the scheme, we conduct our MBNF method on the public ADNI database for identifying subjects with AD/MCI from normal controls.

Discussion: Experimental results demonstrate that the proposed scheme extracts rich multi-band network features and biomarker information, and also achieves better classification accuracy.

KEYWORDS

functional brain networks, signal decomposition, network fusion, resting state fMRI, Alzheimer diagnosis

1 Introduction

Alzheimer's disease (AD) is an irreversible neurodegenerative disease that severely impacts the quality of life for patients (1). As a non-invasive measure for detecting brain abnormalities, functional brain network (FBN), derived from resting state magnetic resonance imaging (rs-fMRI), provides a valuable opportunity for early intervention and control of AD disease. Previous studies usually divide the brain of patients into several regions of interest (ROI) through a certain brain template. Then, the FBN is constructed by calculating the full band Blood Oxygen Level Dependent (BOLD) signals correlation coefficients among these ROIs. However, a single band is not sufficient to capture the diagnostic and prognostic information contained in multiple frequency bands.

In practice, BOLD signals based on different frequencies have different physiological significance. As early as 1995, researchers found that there is a correlation between low-frequency BOLD signals in certain brain regions (2). In 2011, Baria et al. divided the BOLD signal into four frequency bands to study the energy of each band and its distribution in the whole brain. They found that the signals in the 0.01–0.05 Hz frequency band are mainly distributed in the prefrontal, parietal, and occipital cortices; the signals in the 0.05–0.1 Hz frequency band are mainly distributed in the thalamus and basal ganglia; the signals in the 0.1–0.15 Hz frequency band are mainly distributed in the insula and temporal cortex; the signals in the 0.15–0.2 Hz frequency band are also distributed in the insula and temporal cortex (3). Most studies focused on the BOLD signal at (0.01–0.08) Hz, a range in which frequencies vary between brain regions.

In addition to the different physiological significance, many studies found that the use of frequency division in estimating FBNs with different frequency bands can achieve a variety of descriptions of FBN structures. For example, Zhang et al. calculate the node statistics (e.g., node degree, node path length, and betweenness centrality) of FBNs estimated by different bands and discover that the structural characteristics of different frequency bands are significantly different (4). Song et al. decomposed the time series of each voxel and found that ReHo in cortical areas was higher and more frequency-dependent than those in the subcortical regions (5). Li et al. found that compared with the healthy control group, the functional connectivity of patients with temporal lobe epilepsy in δ, θ, low α, and β bands was significantly increased, and the value of the weighted small-world measure in θ band was significantly decreased (6). Besides, studies have found that different band-based FBNs used for disease diagnosis achieved different classification results (7). The explanation is that FBNs based on different frequencies have different discrimination abilities.

Since the different information brought by different frequency bands, it is a good perspective to decompose the BOLD signal into multiple bands for constructing multiple FBNs and fuse the features of every FBN. For example, Zou et al. extract the temporal, spatial, and spatial-temporal variability features of functional networks in each frequency band and fused them into a set of feature vectors for schizophrenia classification (8). Zuo et al. proposed a deep multi-fusion framework with classifier-based feature synthesis to automatically fuse multi-modal medical images. They validated the approach for brain disease classification using the fused images and illustrated that the improvement in classification performance is due to the adoption of the fusion strategy (9). However, these feature fusion methods have limited interpretability, which does not provide a good biomarker for the diagnosis of brain diseases. More important, both global-and local-level features extracted from FBNs tend to capture different network properties, which requires prior knowledge and thus makes the feature design an intractable problem.

Different from feature fusion, network fusion can obtain the diverse information of multiband-based FBNs and eliminate the redundant information caused by the correlation between different feature sets. Considering the varied characteristics of FBNs in different frequency bands, we propose a novel multi-band network fusion framework (MBNF) to estimate information-rich multi-frequency FBNs. Specifically, our framework can be summarized in the following steps: (1) using ensemble empirical mode decomposition (EEMD) to decomposed the bold signal into high and low-frequency bands adaptively; (2) fusing FBNs constructed by the two frequency bands into a multi-band fusion network by similarity network fusion (SNF); (3) extracting the features of the fused networks and employing the Support Vector Machines (SVM) for classification.

The rest of the paper is organized as follows. In Section "2 Material and methods," we present the experimental data and the proposed method. In Section "3 Experiment," we design the experiment and compare it with other methods. In Section "4 Discussion," we discuss the effect of different signal decomposition methods, and different fusion methods on the classification results. Then, we propose the limitations of the work and future research directions. In Section "5 Conclusion," we conclude this article.

2 Materials and methods

In this section, we first introduce data acquisition and preprocessing in detail. Then the overall process of brain disease classification based on the MBNF framework is presented in the following parts.

2.1 Dataset description and image preprocessing

In this paper, we evaluate our proposed scheme based on the dataset from the Alzheimer's Disease Neuroimaging Initiative (ADNI), which divides MCI into two subcategories, early MCI (eMCI) and late MCI (lMCI), and previous studies have shown that lMCI has a high potential for transition to AD. The datasets contain 154 normal controls (NCs), 165 eMCI, 145 lMCI, and 99 AD. The scan parameters of these data as listed below: in-plane image resolution is 2.29–3.31 mm and the thickness of each slice is 3.31 mm. The Echo time (TE) of the slice is 30 ms, and the repetition time (TR) is 2.2–3.1 s. Each subject's scan consisted of 140 volumes. The detailed demographic information is shown in Table 1.

We used FSL FEAT software which is a standard pipeline to process the rs-fMRI scans (10). We first cast aside the first 3 volumes to allow signal stabilization. For the remaining 137 volumes, we corrected the slice time and motion to avoid interference with the data and eliminate the impact of head motion. Then, we striped the structural skull according to the T1-weighted MRI. We use the processed image to align with the Montreal Neurological Institute (MNI) space. All subjects are processed with band-pass filtering at frequency intervals of [0.015, 0.15 Hz]. And then we regress the nuisance signals which contain motion parameters, white matter, and cerebrospinal fluid. Furthermore, a Gaussian kernel with full-width-at half-maximum (FWHM) of 6 mm is used to smooth the data. It is worth noting that we did not perform scrubbing to data because this would introduce additional artifacts. At last, the brain space of fMRI scans is partitioned into 116 pre-defined ROIs using the Automated Anatomical Labeling (AAL) template (11). For each subject, the bold signals are extracted from each ROI, and then normalized as following:

$$\mathrm{r}(x) = \frac{(x - \mu_i)}{\sigma_i} \tag{1}$$

where x denotes the time point signal from the i-th ROI. μ_i represents the mean of the x and σ_i denote the standard deviation of the x.

TABLE 1 Demographic information of the involved 563 rs-fMRI subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database.

Category	Scan #	Age (Years)	Gender (M/F)
AD	99	75.04± 7.71	55/44
eMCI	165	72.03± 7.26	73/92
lMCI	145	71.99± 7.67	95/50
NC	154	75.36± 6.16	67/87

The values are denoted as mean ± standard deviation. M/F: male/female.

2.2 The multi-band network fusion framework

In this section, we introduce the multi-band network fusion framework (MBNF) scheme for brain disease diagnosis. As shown in Figure 1, the MBNF contains three major parts: (1) BOLD signal decomposition based on EEMD; (2) FBN construction and fusion; and (3) feature selection and classification.

2.2.1 BOLD signal decomposition based on EEMD

Previous studies typically used band-pass filters (e.g., wavelet transform) to acquire multi-band signals. However, since the frequency characteristics of the BOLD signal are complex, traditional band-pass filters are unsuitable. Therefore, Huang et al. propose a novel adaptive signal time-frequency processing method called empirical mode decomposition (EMD) (12). Different from wavelet transform which needs to set the feasible decomposition layers in advance, EMD can decompose signals adaptively according to the time characteristics of data. Specifically, EMD can decompose the non-stationary time series into a group of Intrinsic Mode Functions (IMF) components, which are oscillatory functions with time-varying frequencies and can reflect the local characteristics of non-stationary signals (13).

In practice, the mode aliasing problem can occur during the execution of EMD, which leads to mistakes for subsequent feature extraction, model training, and pattern recognition. To solve this problem, the ensemble empirical mode decomposition (EEMD), an improved method of EMD, is performed for signal decomposition in the proposed MBNF method (14). Specifically, EEMD adds different white noises with the same amplitude to alter the extreme point characteristics of signals (15). Figure 2 presents the algorithm flowchart. In Figure 2, x is the original signal, n_m represents the $m-th$ additive white noise sequence, $c_{m,f}$ represents the $f-th$ IMF component obtained by decomposition after adding white noise for the $m-th$ time, f is the number of IMF components, $r_{m,f}$ is the residual function, and M is the average number of corresponding IMF components after multiple decomposition.

After obtaining the IMF components, we transformed the IMF time-domain signals of each brain region into frequency-domain signals to display the frequency-domain range of each IMF component. Note, since EEMD decomposition is adaptive, the number of IMF components after signal decomposition in each brain region may be different. Specifically, we calculate the average frequency of IMF components in every brain region and show the total results of all subjects in different categories (i.e., eMCI, lMCI, AD, and NC) in Figure 3. We can observe in Figure 3 that no matter in which category, the IMF1 component is about 0.06–0.16 Hz, while the average frequency of other IMF components is less than 0.1 Hz. In order

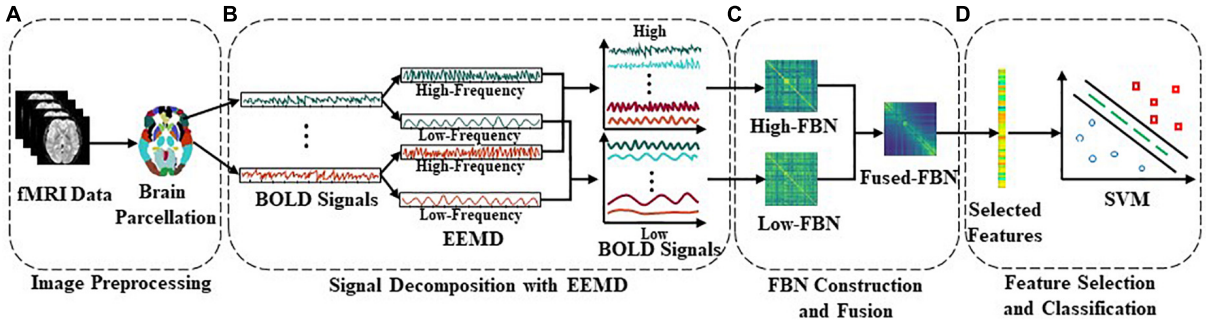


FIGURE 1
Flowchart of the proposed multi-band network fusion framework (MBNF) scheme for brain disease classification, including four major parts: **(A)** data acquisition and image pre-processing; **(B)** blood oxygen level dependent (BOLD) signal decomposition based on ensemble empirical mode decomposition (EEMD); **(C)** functional brain networks (FBN) construction and fusion; and **(D)** feature selection and classification.

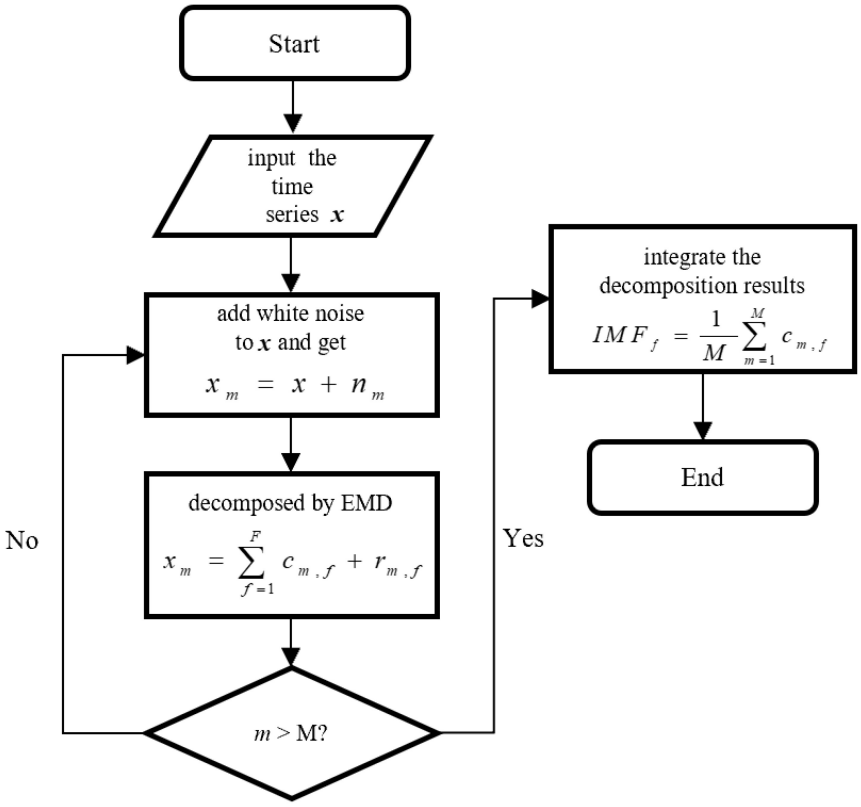


FIGURE 2
The algorithm flowchart of ensemble empirical mode decomposition (EEMD).

to facilitate the construction of a FBN for subsequent analysis, IMF1 components are used as high-frequency BOLD signals, and the remaining IMFs components are integrated together as low-frequency BOLD signals.

2.2.2 FBN construction and fusion

Once we obtain the high/low-frequency BOLD signals of each ROI, we utilize the two types of signals to estimate different FBNs, which provides an effective tool to compare different subjects and to mine biomarkers of neurological/mental disorders. Note, we perform different methods to construct FBNs in the follow-up experiment for verifying the robustness of our method. In recent decades, a number of methods have been developed for constructing FBNs, among which the representative is Pearson's correlation (PC) and sparse representation (SR) (16–18).

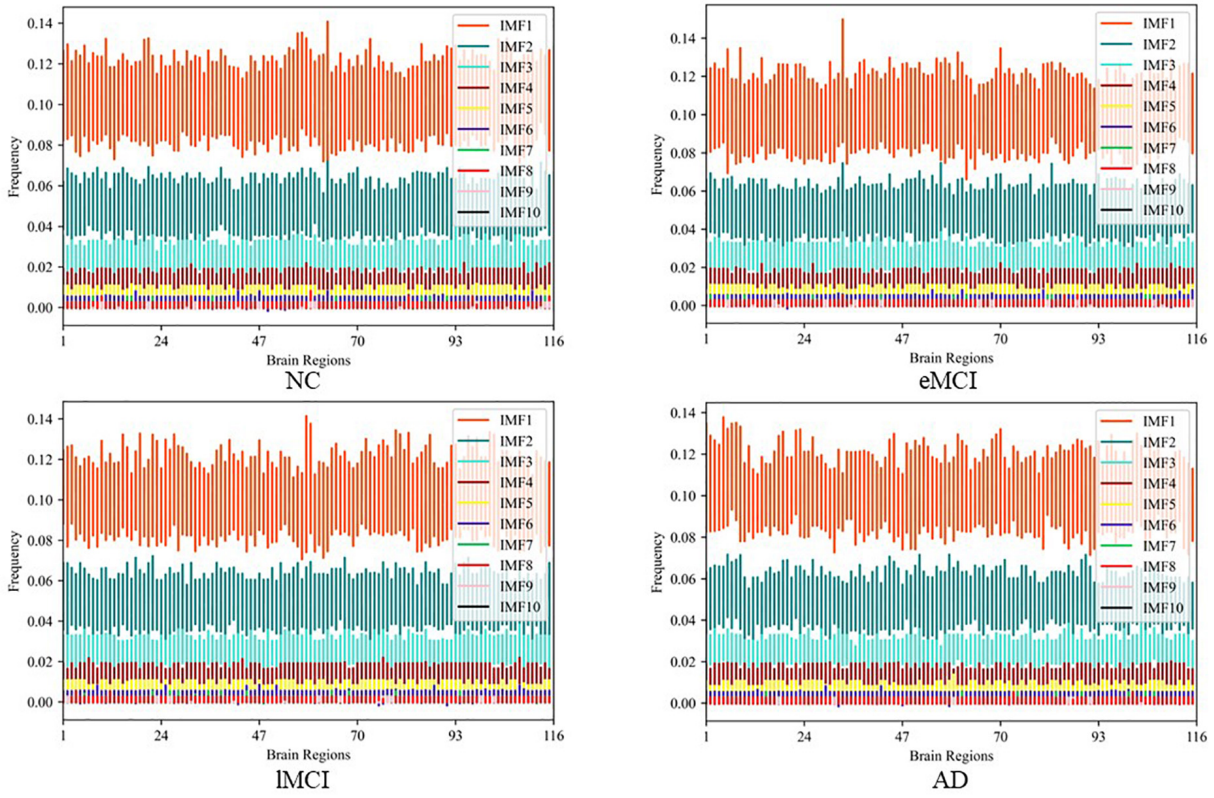


FIGURE 3
The average frequency of intrinsic mode functions (IMF) components in every brain region in different categories.

Specifically, denote P ($P = 116$ in this work) is the number of ROIs and T ($T = 137$ in this work) is the total number of temporal image volumes. For any i, j ($i, j = 1, \cdots, P$), W_{ij} is the functional connectivity between a pair of ROIs i and j. The calculation formula of PC-based functional connectivity is as follows:

$$W_{ij}^{PC} = \frac{(x_i - \bar{x}_i)^T (x_j - \bar{x}_j)}{\sqrt{(x_i - \bar{x}_i)^T (x_i - \bar{x}_i)} \sqrt{(x_j - \bar{x}_j)^T (x_j - \bar{x}_j)}} \quad (2)$$

where $x_i \in R^T$ represents the time series of the ith ROI, $\bar{x}_i \in R^T$ is the corresponding mean vector of x_i . Another FBN construction method is SR, which is an l_1 -regularized linear regression. The mathematical model can be obtained by the following objective function:

$$\min_{W^{SR}} \sum_{i=1}^{P} (||x_i - \sum_{j \neq i} W_{ij}^{SR} x_j||^2 + \lambda \sum_{j \neq i} |W_{ij}^{SR}|) \quad (3)$$

where λ is a regularized parameter. Note, the same methods are performed for high-frequency FBN and low-frequency FBN.

After FBN construction, we perform the similarity network fusion (SNF) method to fuse high/low-frequency FBNs for obtaining complementary information of multi-frequency bands. The similarity fusion network is robust to noise and can obtain useful information from fewer samples (19, 20). For high-frequency FBN (W^{High}) and low-frequency FBN (W^{Low}), we construct similarity matrix S^{High} and S^{Low} separately. Note, similarity matrix is a sparse kernel matrix encoding its own sparse strong connections. For every similarity matrix S, we use the K-nearest neighbors (KNN) to measure the local affinity, and set the similarity between non-adjacent points to zero. The calculation formula of similarity value between a pair of ROIs i and j is as follows:

$$S_{ij} = \begin{cases} W_{ij}, & if\ i \in KNN_j \\ 0, & otherwise \end{cases} \quad (4)$$

where KNN_j represents a set of K-nearest neighbors of the ROI j in W. Similar to previous study (21), we set the number of nearest neighbors to 11.

Based on the sparse kernel matrixes S^{High} and S^{Low}, we fuse them into a single network using nonlinear methods. Each similar network needs to be updated iteratively to make it

more similar to another network. For example, S^{High} could be iteratively updated as follows:

$$(W^{High})^{g+1} = S^{High} \times (W^{Low})^{(g)} \times (S^{High})^{T} \quad (5)$$

where g is the number of iterations, $(W^{Low})^{(g)}$ represents the W^{Low} after gth iteration.

Because different networks carry distinct frequency information, W^{High} can integrate the information provided by W^{Low} after several iterative learnings. At the same time, the sparse kernel matrix S guides the iterative process through the strongest connections of W, and thus can reduce the noise effectively. Iteration stops when the converged network is close to stopping changing. Because different networks carried different frequency information, fusion networks integrated the information provided by different frequency networks. When the iterative fusion network was almost constant, the network stopped iterating. Specifically, the fusion network stops updating in the process of iteration when it satisfies the formula (6):

$$||(W^{High})^{g+1} - (W^{High})^{g}|| \leq 0.01 \quad (6)$$

Finally, we obtain the fusion network by averaging two networks. The fusion network is as follows:

$$W^{Fused} = \frac{\left(W^{High}\right)^{'} + \left(W^{Low}\right)^{'}}{2} \quad (7)$$

where $(\cdot)^{'}$ represents the last updated matrix.

2.2.3 Feature selection and classification

Once we obtain the fused FBNs for all subjects, the subsequent task is to extract/select the most discriminative features according to the FBNs for disease classification. Currently, there are two categories of features based on different granularities in FBN analysis, including node-level and edge-level features. Since the node-level features tend to capture different network properties that caused the extra prior knowledge to design effective features, we use the edge-level feature (i.e., functional connectivity between ROIs) in our experiment. As shown in Figure 4, we concatenate the upper triangle of the obtained fused FBNs into an edge vector (removing the redundant part if the adjacent matrix is symmetric), and then pile up the edge vectors from all subjects into a feature matrix for subsequent classification tasks. Besides, in order to remove redundant information in these features, t-test is used for feature selection ($P < 0.05$).

Finally, considering that small changes in different steps (FBN construction, feature selection, and classification) will have an impact on the end results, it is difficult to conclude which step contributes further to the final accuracy. Therefore, the simplest and most popular classifier support vector machine ($C = 1$) is performed to classify the AD/MCI from NC. In addition, the reason for using SVM instead of deep learning is the latter often requires very large data sets. It is challenging to train a good model and tune the hyper-parameters when there are not enough training samples (subjects).

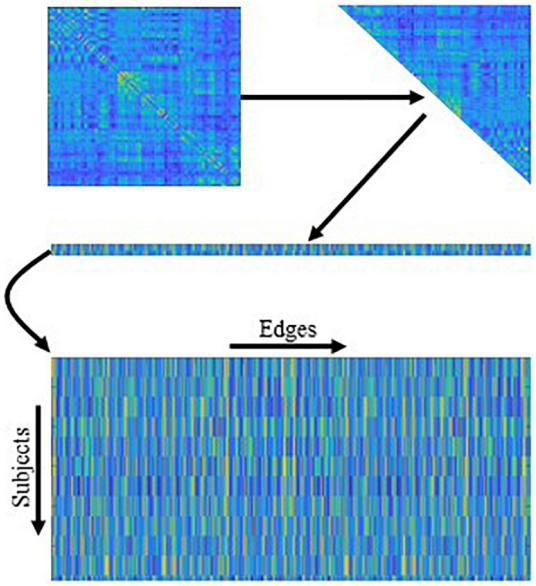


FIGURE 4
The mechanism of traditional edge feature extraction in functional brain network (FBN). The network adjacency matrix from each subject is first mapped onto a vector by removing the redundant part if the matrix is symmetric, and then the vectors from all subjects are rearranged together as an input of the following feature selection methods.

3 Experiment

In this section, we first introduce the competing methods with our proposed scheme and the settings of our experiment. Then the experiment result is analyzed in detail.

3.1 Competing methods

In the experiments, we compare our proposed MBNF with several schemes, including (1) Full-Band, a scheme based on FBN construction by a full BOLD signal; (2) Low-Band, a scheme based on FBN construction by the low band after BOLD signal decomposition; (3) High-Band, a scheme based on FBN construction by the high band after BOLD signal decomposition; (4) MBNF, our proposed scheme. For a fair comparison, we employ t-test ($p < 0.05$) to select discriminative features and then use SVM ($C = 1$) for brain disease classification for all competing schemes. Besides, two FBN construction methods mentioned in 2.2.2

TABLE 2 Classification performance of four schemes in four classification tasks based on Pearson's correlation (PC) construction method (mean ± standard deviation).

Task	Scheme	ACC (%)	SEN (%)	SPE (%)	AUC (%)
eMCI vs. NC	Full-band	84.29 ± 1.94	88.60 ± 1.47	79.92 ± 1.52	90.94 ± 0.47
	Low-band	78.98 ± 1.04	81.82 ± 1.07	76.94 ± 2.31	87.86 ± 0.96
	High-band	83.94 ± 1.28	88.61 ± 1.31	78.24 ± 1.17	89.58 ± 0.83
	MBNF	**90.60 ± 1.56***	**92.72 ± 1.95***	**88.48 ± 2.14***	**97.50 ± 0.25***
lMCI vs. NC	Full-band	86.28 ± 2.02	87.95 ± 2.13	83.40 ± 1.22	82.43 ± 0.32
	Low-band	78.25 ± 1.18	80.48 ± 1.72	74.89 ± 1.83	87.03 ± 0.61
	High-band	84.61 ± 1.58	85.95 ± 1.52	80.03 ± 1.64	92.02 ± 1.05
	MBNF	**91.98 ± 1.66***	**93.34 ± 2.29***	**90.23 ± 1.47***	**97.12 ± 0.71***
eMCI vs. lMCI	Full-band	81.93 ± 2.11	**91.78 ± 1.81**	73.41 ± 2.23	88.19 ± 0.57
	Low-band	76.77 ± 0.98	74.67 ± 1.32	80.06 ± 1.37	82.43 ± 1.04
	High-band	75.80 ± 1.74	68.24 ± 2.01	83.32 ± 0.95	87.05 ± 1.11
	MBNF	**90.64 ± 1.44***	86.02 ± 1.92	**94.97 ± 2.02***	**96.98 ± 0.46***
AD vs. NC	Full-band	90.49 ± 2.01	**87.95 ± 1.23**	83.40 ± 1.78	82.43 ± 0.41
	Low-band	80.63 ± 0.94	79.03 ± 1.44	90.97 ± 0.89	86.79 ± 0.52
	High-band	90.89 ± 1.27	81.38 ± 1.56	95.39 ± 2.01	96.90 ± 0.79
	MBNF	**93.08 ± 1.85***	86.96 ± 1.35	**96.73 ± 1.32***	**98.58 ± 0.74***

*Denotes that the result of MBNF is significantly better than other competing schemes. Bold values indicate the best results in each task.

are performed in our experiment to further indicate the effectiveness of our method.

3.2 Experimental settings

We designed four classification tasks to evaluate the performance of our method and four competing schemes, which are as follows: (1) eMCI vs. NC (2) lMCI vs. NC (3) AD vs. NC (4) eMCI vs. lMCI. Then, three evaluation metrics are employed for evaluating the classification performance of all methods, including classification accuracy (ACC), sensitivity (SEN), and specificity (SPE), which are defined as follows:

$$ACC = \frac{TP + TN}{TP + FP + TN + FN} \times 100\% \quad (8)$$

$$SEN = \frac{TP}{TP + FN} \times 100\% \quad (9)$$

$$SPE = \frac{TN}{TN + FP} \times 100\% \quad (10)$$

where TP, TN, FP, and FN represent true positive, true negative, false positive, and false negative, respectively. In addition to the above, we also add the area under the receiver operating characteristic curve (AUC) as another metric.

In our experiment, a 5-fold cross-validation (CV) is adopted to evaluate the generalization capability of the different methods. Besides, considering the hyper-parameters (i.e., sparsity) involved in the FBN construction methods may significantly affect the ultimate classification results, we select optimal parametric values by a grid search in a large range. For the regularized parameter λ in SR, we use 20 candidate values in [0.1, 0.15, 0.2, . . ., 0.95, 1]. Although PC is parameter-free. For a fair comparison, we perform a thresholding parameter in PC by preserving a percentage of connectivity with strongest correlation. To be consistent with other methods, we set up 20 sparsity from a candidate set [5%, 10%, ...95%, 99%]. For example, 100% means all edges are preserved, and 90% means 10% weak edges are removed. Then, an inner-5-fold CV on the training data to determine the optimal sparsity, which is based on the classification accuracy in each inner loop. For fairness, we also employed inner-5-fold CV strategy in other competitive methods compared with MBNF. Note, we perform the 5-fold CV process 1,000 times independently to avoid random errors introduced in cross-validation, and the mean and standard deviation of the classification results are reported in Table 2. To illustrate the statistical significance of the results, we perform a paired t-test ($p < 0.05$) on the results of the methods involved and then use "*" to mark the results better than the other methods.

3.3 Classification results and analysis

Tables 2, 3 provide the classification results of four schemes in four tasks based on two FBN construction methods, and also shows some intriguing findings.

(1) The proposed scheme with multi-band fusion networks is significantly superior to other three competing schemes. This indicates that combining the various information

TABLE 3 Classification performance of four schemes in four classification tasks based on sparse representation (SR) construction method (mean ± standard deviation).

Task	Scheme	ACC (%)	SEN (%)	SPE (%)	AUC (%)
eMCI vs. NC	Full-band	84.31 ± 2.31	86.30 ± 1.85	81.84 ± 1.39	93.24 ± 0.56
	Low-band	77.43 ± 1.80	79.69 ± 1.36	76.01 ± 1.35	83.02 ± 0.68
	High-band	82.14 ± 1.53	85.85 ± 1.09	78.17 ± 1.91	90.76 ± 0.51
	MBNF	**89.91 ± 1.75***	**93.51 ± 1.75***	**88.59 ± 1.84***	**96.97 ± 0.41***
lMCI vs. NC	Full-band	89.94 ± 1.53	90.93 ± 1.80	89.15 ± 1.76	95.53 ± 0.87
	Low-band	79.24 ± 1.69	78.52 ± 1.56	80.37 ± 1.77	85.85 ± 0.89
	High-band	86.59 ± 1.27	87.18 ± 1.41	85.42 ± 1.39	93.31 ± 0.55
	MBNF	**91.96 ± 1.75***	**94.35 ± 1.65***	**90.41 ± 1.33**	**98.02 ± 0.33***
eMCI vs. lMCI	Full-band	80.96 ± 1.34	79.51 ± 2.07	82.26 ± 2.14	89.52 ± 0.31
	Low-band	67.09 ± 1.35	64.80 ± 2.03	69.52 ± 1.57	75.09 ± 0.64
	High-band	79.35 ± 1.52	77.64 ± 1.59	82.15 ± 0.98	88.23 ± 0.57
	MBNF	**90.96 ± 1.43***	**92.26 ± 1.35***	**90.81 ± 1.98***	**97.65 ± 0.49***
AD vs. NC	Full-band	89.34 ± 1.30	84.04 ± 1.32	93.62 ± 1.75	96.45 ± 0.39
	Low-band	86.17 ± 1.80	81.13 ± 1.67	88.66 ± 0.78	90.17 ± 0.48
	High-band	86.96 ± 1.31	80.33 ± 1.25	92.78 ± 1.82	95.83 ± 0.53
	MBNF	**92.86 ± 1.69**	**88.82 ± 1.44***	**95.98 ± 1.55***	**97.77 ± 0.58**

*Denotes that the result of MBNF is significantly better than other competing schemes. Bold values indicate the best results in each task.

of multi-band FBNs helps boost the classification performance for brain disease classification.

(2) The low-band scheme achieves a worse performance when compared with the high-band scheme in every classification task. Combined with previous researches (22, 23), the possible reason is that the features of high band-based FBNs are more robust and discriminative. For example, Zuo et al. have shown that the test–retest reliability of high-band-based fluctuations is greater and more widely distributed than that of the low-band (24).

(3) Regarding four tasks of classification based on two FBN construction methods, the task of identifying subjects with AD from normal controls is relatively easier. The underlying reason is that brain function degeneration in AD subjects could be more serious than MCI and NC.

3.4 Discriminative functional connections and brain regions

As the most important step in FBN analysis, selecting the discriminative features is meaningful to search for the biomarkers used to determine brain disease. A rising corpus of research indicates that many mental diseases emerge from interactions between various brain regions rather than being restricted to just one particular area of the brain. Therefore, we employ *t*-test to select the most discriminative functional connections for our MBFN method in four tasks of classification. As shown in **Figure 5**, the color of each arc is chosen at random for better visualization, and its thickness represents the discriminative power of connection (rather than the actual connectivity strength).

Besides, we also visualized the discriminative brain regions based on the functional connections in **Figure 6**. This visualization is drawn by BrainNet Viewer toolbox[1] and these stably selected brain regions are mapped onto the International Consortium for Brain Mapping (ICBM) 152 surface based on AAL atlas. For MCI classification (eMCI vs. NC and lMCI vs. NC), we can observe that frontal lobe, Cingulum, Postcentral, Fusiform and inferior temporal gyrus are the most discriminative brain regions. Previous research has shown that abnormal changes in these brain regions accelerates the conversion of people with mild cognitive impairment to Alzheimer's disease (25–30). Similarly, for AD classification, the regions of the posterior cingulate gyrus, postcentral gyrus, c, hippocampus, middle temporal gyrus, and inferior temporal gyrus are the most discriminative brain areas, which have been previously documented to be involved in AD (31–34).

Many brain disorders are not isolated to specific brain regions, but result from the interaction of different brain regions. For example, the frontal lobe plays a key role in non-task long-term memory (35), the hippocampus is responsible for storage and transformation of long-term memory and spatial memory and localization (36), and the posterior cingulate gyrus is involved in processes such as emotion and self-evaluation (37). Memory loss, cognitive decline and frequent mood swings are hallmarks of Alzheimer's disease (38). Previous studies have shown differences in the connections between these brain regions between AD patients and normal controls. These

1 https://www.nitrc.org

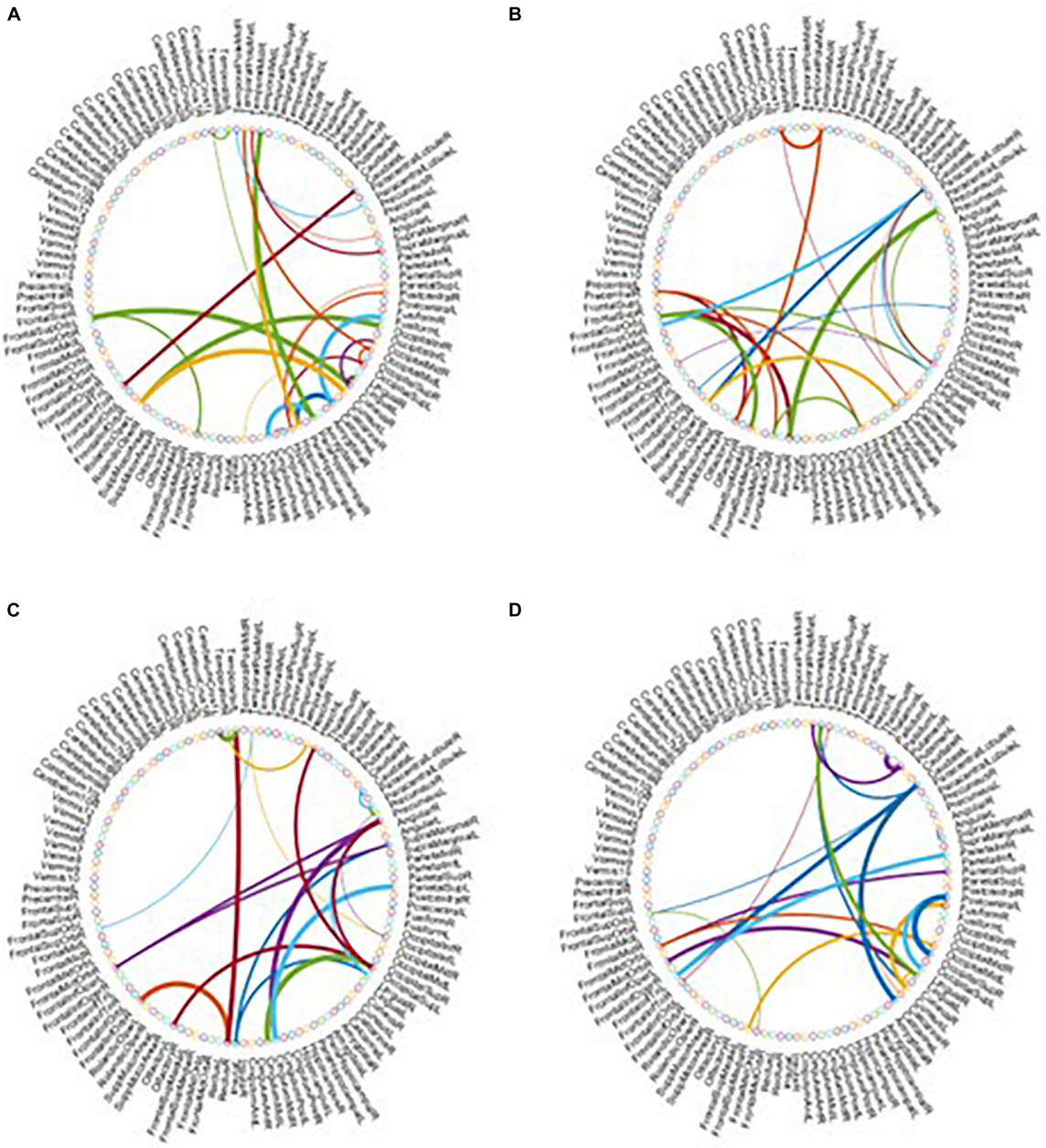


FIGURE 5
Most discriminative functional connections in four classification tasks: **(A)** eMCI vs. NC, **(B)** lMCI vs. NC, **(C)** AD vs. NC, and **(D)** eMCI vs. lMCI.

characteristics could be considered as biomarkers of Alzheimer's disease (39).

3.5 Frequency variability of brain regions

To visually illustrate the difference between high- and low-frequency BOLD signals, we employ frequency variability (FV) to assess changes in different brain regions at different frequency bands (40). FV is defined as follows:

$$FV_i = 1 - \frac{\sum_{f=1,g\neq f}^{N_F} corrcoef\left(FC_{f,i}, FC_{g,i}\right)}{N_F \times \frac{N_F-1}{2}} \tag{11}$$

where $FC_{f,i}$ is the functional connection of node $i\,(i = 1, \cdots, 116)$ to other ROIs in frequency band f, N_F is the total number of frequency bands (here $N_F = 2$). The

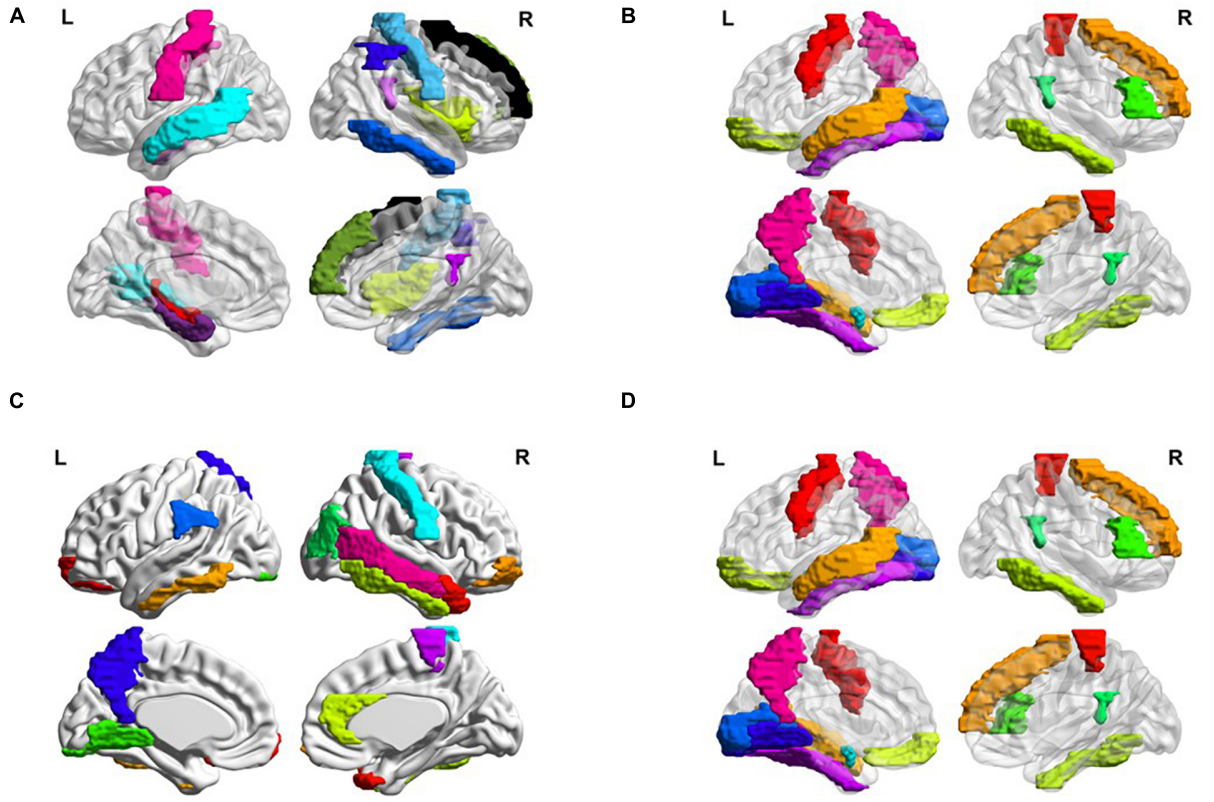


FIGURE 6
Most discriminative regions of interests (ROIs) identified by our proposed MBFN method in four tasks of **(A)** eMCI vs. NC, **(B)** lMCI vs. NC, **(C)** AD vs. NC, and **(D)** eMCI vs. lMCI.

higher the value of FV, the greater the difference of brain regions in different frequency bands. Figure 7 shows the FV of all ROIs in four categories (i.e., NC, eMCI, lMCI, and AD).

We can observe that the FV of normal people is relatively lower than other patient categories. The probable reason is that the disease of MCI/AD can cause disturbance of normal neuronal behavior and destruction of neuronal networks, which leads to unstable BOLD signals. In addition, for patient categories (i.e., eMCI, lMCI, and AD), the amygdala, middle temporal gyrus, and superior frontal gyrus showed relatively high FV, which may be biologically associated with MCI/AD.

4 Discussion

In this section, we first analyze the effect of different signal decomposition methods, the impact of different fusion methods on classification performance, the effect of Different Datasets, and the effect of Connection Variations in FBNs. Then we present the limitations of this work as well as several future research directions.

4.1 Effect of different signal decomposition methods

In our proposed MBNF scheme, the EEMD signal decomposition method is used to extract different frequency band signals. To verify the effectiveness of the EEMD method and the effect of different signal decomposition methods on our experiment, we employ three different signal decomposition competing methods, including (1) discrete wavelet transform (DWT) (41), (2) local mean decomposition (LMD) (42), and (3) empirical mode decomposition (EMD). For a fair comparison, all competing schemes are performed in consistent steps (i.e., same data pre-processing, FBN construction and fusion, feature selection, and classification) except for the signal decomposition step.

Table 4 summarizes the results of four signal decomposition methods in two classification tasks. We can observe that our proposed MBNF using the EEMD decomposition method provides the best results. The probable reason is that EEMD can decompose signals adaptively according to the time characteristics of data, which has the advantage of obtaining good results in processing BOLD signals.

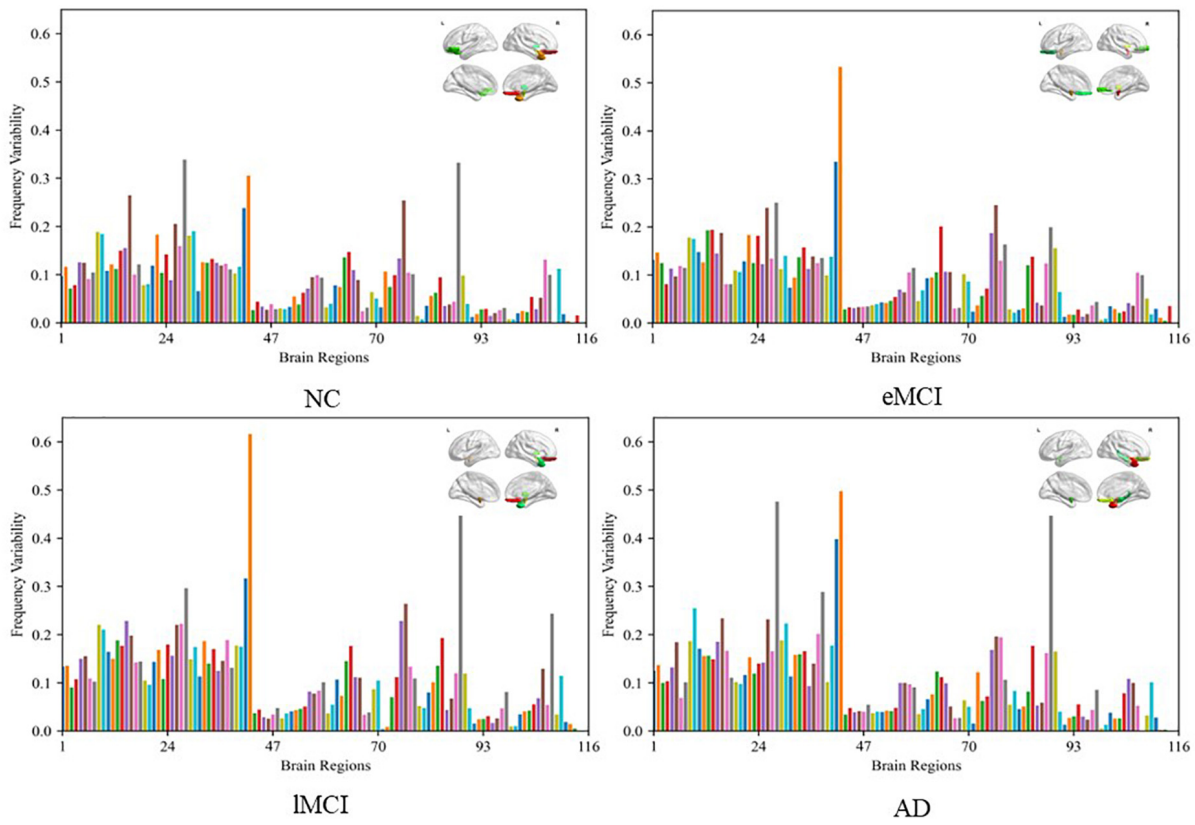


FIGURE 7
Frequency variability of all regions of interests (ROIs) in four categories (i.e., NC, eMCI, lMCI, and AD). The top right in every figure represents the ROI with the highest frequency variability (FV).

TABLE 4 Classification results of four signal decomposition methods in two tasks.

Task	Method	ACC (%)	SEN (%)	SPE (%)	AUC (%)
AD vs. NC	DWT	88.10	78.77	95.13	95.77
	LMD	86.15	73.13	95.71	93.26
	EMD	88.13	83.74	92.67	97.20
	Ours	**93.08**	**86.96**	**96.73**	**98.58**
eMCI vs. lMCI	DWT	85.80	84.19	87.86	93.96
	LMD	85.80	**87.89**	85.87	92.11
	EMD	87.74	85.49	90.76	96.21
	Ours	**90.64**	86.02	**94.97**	**96.98**

Bold values indicate the best results in each task.

4.2 Effect of different fusion methods

We use the SNF method to combine FBNs based on different frequency bands in the proposed MBNF scheme. To verify the effectiveness of the SNF method and the effect of different network fusion methods in our experiment, two methods are used to compare the SNF method, including (1) Concatenate, a scheme for splicing FBNs based on different bands into a feature vector; (2) Canonical Correlation Analysis (CCA), a typical fusion method (43). For a fair comparison, all competing schemes are performed in consistent steps (i.e., same data pre-processing, signal decomposition, FBN construction, feature selection, and classification) except for the FBN fusion step.

In Table 5, we can observe that the performance of CCA is worse than the SNF techniques. The underlying reason is that CCA can only determine the linear correlation and ignore the nonlinear correlation in the interaction between the high-frequency FBN and the low-frequency FBN. Besides, the reason why the SNF achieves better performance than concatenating is that the concatenate method ignores the structural properties of FBNs by the splicing technique. To explore the impact of noise on FBN, we added random white Gaussian noise with varying standard deviation to the FBN (44). It can be seen in the Figure 7 and table that with the increasing noise level, the classification accuracy was decreasing. We used a bootstrapping method to enhance the robustness of our method. We resampled the data and created several training sets which were the same size as the original data. The experimental results are shown in the Figure 8.

TABLE 5 Classification results of three network fusion methods in two tasks.

Task	Method	ACC (%)	SEN (%)	SPE (%)	AUC (%)
AD vs. NC	Concatenate	90.88	84.04	96.02	97.40
	CCA	85.36	75.03	92.97	90.79
	Ours	**93.08**	**86.96**	**96.73**	**98.58**
eMCI vs. lMCI	Concatenate	87.74	**88.52**	87.53	94.43
	CCA	78.38	67.66	88.23	85.15
	Ours	**90.64**	86.02	**94.97**	**96.98**

Bold values indicate the best results in each task.

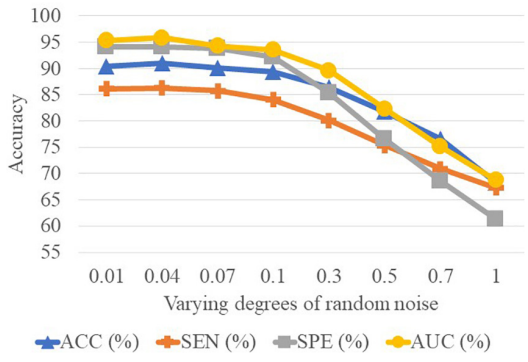


FIGURE 8
Results achieved by the proposed method with varying degrees of functional brain network (FBN) random noise of Pearson's correlation (PC) in lMCI vs. eMCI classification.

4.3 Effect of different datasets

Since different distributed datasets may affect the experimental results, we perform three independent datasets to confirm our conclusions, including Schizophrenia (SZ), Major Depressive Disorder (MDD), and Autism Spectrum Disorder (ASD). Specifically, the dataset of SZ (45), from publicly shared online datasets by the Mind Research Network and the University of New Mexico, includes 57 patients with chronic schizophrenia patients and 64 NCs. Besides, we also perform our proposed scheme on the ABIDE database (46) collected from the New York University site. The ABIDE dataset includes 184 subjects, of which 79 are from ASD and 105 are from NC. The MDD dataset is from the ninth site of the REST-meta-MDD Consortium (47), which contains 49 MDD patients and 47 NCs. Note that due to the fact that the MDD database used in this study is provided as preprocessed by the REST meta-MDD project, we have no control over the preprocessing pipeline. Therefore, we process the other two databases *via* the same pipeline as the MDD database for fairness.

As shown in Table 6, our MBNF method achieves the overall best performance regardless of which database is used. These results imply that combining the structural information of functional brain networks in different frequency bands helps to improve the accuracy of identifying patients from NCs. In addition, the other three databases give lower performance compared to the ADNI database. The probable reason is that the lesions of brain structure caused by AD/MCI are more severe than mental disease (e.g., MDD and ASD) and neurodevelopmental disorders (e.g., SZ).

4.4 Effect of connection variations in FBNs

It is well-known that PC based functional connectivity tends to be sensitive to noise. To investigate whether variations in connectivity affect our proposed method, we performed a set of experiments by adding white Gaussian random noise of varying degrees to the FBN estimated by the PC, and present

TABLE 6 Classification result of three data sets on MBNF method.

Task	Scheme	ACC (%)	SEN (%)	SPE (%)	AUC (%)
Schizophrenia vs. NC	Full-band	60.52	58.37	65.83	68.57
	Low-band	56.74	55.56	63.15	65.55
	High-band	55.81	52.19	60.23	63.08
	MBNF	**61.34**	**59.89**	**66.71**	**69.83**
ASD vs. NC	Full-band	64.86	60.24	69.28	71.86
	Low-band	57.21	55.18	58.36	60.86
	High-band	60.73	56.83	62.61	63.73
	MBNF	**65.58**	**61.36**	**69.72**	**72.59**
MDD vs. NC	Full-band	59.67	61.64	58.73	62.27
	Low-band	54.46	54.28	53.61	57.95
	High-band	56.69	56.39	54.41	59.34
	MBNF	**60.93**	**59.36**	**60.79**	**63.82**

Bold values indicate the best results in each task.

the experimental results in Figure 8. It can be observed that the classification results only show a slight fluctuation when the noise degree (standard deviation) is less than 0.1. However, the classification accuracy decreases substantially as the noise level increases. This side-fact indicates that low degrees of noise have little effect on our method and implies that the MBNF scheme already has a relatively good robustness.

4.5 Limitation and future work

Although our proposed framework has a good effect on disease diagnosis, there are still several limitations that need to be noted. The steps of signal decomposition, FBN construction, and fusion in our proposed MBNF scheme are performed separately, which probably leads to potential noise in each step. In addition, the extracted features based on the way of separate-step are not necessarily optimal for the subsequent classification task. Therefore, an end-to-end method like deep learning improves experimental performance, which is also the direction of our future work.

5 Conclusion

In this paper, we propose a multi-frequency network Fusion framework (MBNF) to combine the structural information of functional brain networks in different frequency bands. Specifically, we first use EEMD to decompose the BOLD signal into high-frequency signal and low-frequency signal. Then we construct a high-frequency functional network and a low-frequency functional network, respectively. Finally, the similarity network fusion is employed to fuse high-frequency network and low-frequency network for classification. The validation on the ADNI dataset shows that our proposed multi-band network fusion framework is effective.

Data availability statement

The original contributions presented in this study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

LG and ZW designed the study. LG and YZ downloaded and analyzed the data, performed experiments, and drafted the manuscript. QL and KG preprocessed the data and performed some experiments. YZ and ZW revised the manuscript. All authors read and approved the final manuscript.

Funding

This work was partly supported by Key Research and Development Program of Hainan province (No. ZDYF2021GXJS017), National Natural Science Foundation of China (No. 82160345), Key Science and Technology plan project of Haikou (2021-016), and Hainan Provincial Natural Science Foundation of China (No. 620RC558). Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012).

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Gauthier S, Rosa-Neto P, Morais J, Webster C. World Alzheimer report 2021: journey through the diagnosis of dementia. *World Alzheimer Rep.* (2021) 1:3–6.

2. Biswal B, Yetkin FZ, Haughton VM, Hyde JS. Functional connectivity in the motor cortex of resting human brain using echo-planar mri. *Magn Reson Med.* (1995) 34:537–41. doi: 10.1002/mrm.1910340409

3. Baria AT, Baliki MN, Parrish T, Apkarian AV. Anatomical and functional assemblies of brain BOLD oscillations. *J Neurosci.* (2011) 31:7910. doi: 10.1523/JNEUROSCI.1296-11.2011

4. Zhang TT, Liao Q, Zhang DM, Zhang C, Yan J, Ngetich R, et al. Predicting MCI to AD conversation using integrated sMRI and rs-fMRI: machine learning and graph theory approach. *Front Aging Neurosci.* (2021) 13:688926. doi: 10.3389/fnagi.2021.688926

5. Song X, Zhang Y, Liu YJPO. Frequency specificity of regional homogeneity in the resting-state human brain. *PLoS One.* (2014) 9:e86818. doi: 10.1371/journal.pone.0086818

6. Li Y, Zhu H, Chen Q, Yang L, Bao X, Chen F, et al. Evaluation of brain network properties in patients with mri-negative temporal lobe epilepsy: an MEG study. *Brain Topogr.* (2021) 34:618–31. doi: 10.1007/s10548-021-00856-y

7. Wang Z, Xin J, Wang X, Wang Z, Zhao Y, Qian W. Voxelwise-based brain function network using multi-graph model. *Sci Rep.* (2018) 8:17754. doi: 10.1038/s41598-018-36155-z

8. Zou H, Yang J. Multi-frequency dynamic weighted functional connectivity networks for schizophrenia diagnosis. *Appl Magn Reson.* (2019) 50:847–59. doi: 10.1007/s00723-019-01117-9

9. Zuo Q, Zhang J, Yang Y. DMC-fusion: deep multi-cascade fusion with classifier-based feature synthesis for medical multi-modal images. *IEEE J Biomed Health Inform.* (2021) 25:3438–49. doi: 10.1109/JBHI.2021.3083752

10. Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FSL. *NeuroImage.* (2012) 62:782–90. doi: 10.1016/j.neuroimage.2011.09.015

11. 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) 15:273–89. doi: 10.1006/nimg.2001.0978

12. Huang NE, Shen Z, Long SR, Wu MC, Shih HH, Zheng Q, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc R Soc Lond Ser A Mathemat Phys Eng Sci.* (1998) 454:903–95. doi: 10.1098/rspa.1998.0193

13. Wang YH, Yeh CH, Young HWV, Hu K, Lo MT. On the computational complexity of the empirical mode decomposition algorithm. *Phys Statist Mech Appl.* (2014) 400:159–67. doi: 10.1016/j.physa.2014.01.020

14. Wu Z, Huang NE. Ensemble empirical mode decomposition: a noise-assisted data analysis method. *Adv Adapt Data Anal.* (2009) 1:1–41. doi: 10.1142/S1793536909000047

15. Moncada F, Gonzalez VM, Alvarez V, Garcia B, Villar JR. A preliminary study on automatic detection and filtering of artifacts from EEG signals. *Proceedings of the 2021 IEEE 34th International Symposium on Computer-based Medical Systems (CBMS)*. Electr Network (2021). doi: 10.1109/CBMS52027.2021.00046

16. Sun L, Xue YF, Zhang YN, Qiao LS, Zhang LM, Liu MX. Estimating sparse functional connectivity networks via hyperparameter-free learning model. *Artif Intellig Med.* (2021) 111:32–45. doi: 10.1016/j.artmed.2020.102004

17. Li WK, Chen YC, Xu XW, Wang X, Gao X. Human-guided functional connectivity network estimation for chronic tinnitus identification: a modularity view. *IEEE J Biomed Health Inform.* (2022) 26:4849–58. doi: 10.1109/JBHI.2022.3190277

18. Su H, Zhang LM, Qiao LS, Liu MX. Estimating high-order brain functional networks by correlation-preserving embedding. *Med Biol Eng Comput.* (2022) 60:2813–23. doi: 10.1007/s11517-022-02628-7

19. Liu WY, Zhang WH, Han JG, Wang GF. A new wind turbine fault diagnosis method based on the local mean decomposition. *Renew Energy.* (2012) 48:411–5. doi: 10.1016/j.renene.2012.05.018

20. Shen LL, Xia Y, Li YP, Sun MY. A multiscale siamese convolutional neural network with cross-channel fusion for motor imagery decoding. *J Neurosci Methods.* (2022) 367:44–56. doi: 10.1016/j.jneumeth.2021.109426

21. Ma T, Zhang A. Integrate multi-omic data using affinity network fusion (ANF) for cancer patient clustering. *Proceedings of the 2017 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. Kansas City, MI (2017). p. 398–403. doi: 10.1109/BIBM.2017.8217682

22. Chen H, Duan XJ, Liu F, Lu FM, Ma XJ, Zhang YX, et al. Multivariate classification of autism spectrum disorder using frequency-specific resting-state functional connectivity-a multi-center study. *Prog Neuropsychopharmacol Biol Psychiatry.* (2016) 64:1–9. doi: 10.1016/j.pnpbp.2015.06.014

23. Salsabilian S, Bibineyshvili Y, Margolis D, Najafizadeh L. Identifying mild traumatic brain injury using measures of frequency-specified networks. *J Neural Eng.* (2022) 19:23–34. doi: 10.1088/1741-2552/ac954e

24. Zuo X-N, Di Martino A, Kelly C, Shehzad ZE, Gee DG, Klein DF, et al. The oscillating brain: complex and reliable. *NeuroImage.* (2010) 49:1432–45. doi: 10.1016/j.neuroimage.2009.09.037

25. Qi ZG, Wu X, Wang ZQ, Zhang N, Dong HQ, Yao L, et al. Impairment and compensation coexist in amnestic MCI default mode network. *NeuroImage.* (2010) 50:48–55. doi: 10.1016/j.neuroimage.2009.12.025

26. Zhang YY, Xue YF, Wu X, Qiao LS, Wang ZX, Shen DG, et al. Selecting multiple node statistics jointly from functional connectivity networks for brain disorders identification. *Brain Topogr.* (2022) 2:19–32.

27. Zhou B, Liu Y, Zhang ZQ, An NY, Yao HX, Wang P, et al. Impaired functional connectivity of the thalamus in Alzheimer's disease and mild cognitive impairment: a resting-state fMRI study. *Curr Alzheimer Res.* (2013) 10:754–66. doi: 10.2174/15672050113109990146

28. Cai SP, Huang LY, Zou J, Jing LL, Zhai BZ, Ji GJ, et al. Changes in thalamic connectivity in the early and late stages of amnestic mild cognitive impairment: a resting-state functional magnetic resonance study from ADNI. *PLoS One.* (2015) 10:e0115573. doi: 10.1371/journal.pone.0115573

29. Gilligan TM, Sibilia F, Farrell D, Lyons D, Kennelly SP, Bokde ALW. No relationship between fornix and cingulum degradation and within-network decreases in functional connectivity in prodromal Alzheimer's disease. *PLoS One.* (2019) 14:e0222977. doi: 10.1371/journal.pone.0222977

30. Li WK, Xu XW, Wang ZX, Peng LL, Wang PJ, Gao X. Multiple connection pattern combination from single-mode data for mild cognitive impairment identification. *Front Cell Dev Biol.* (2021) 9:782727. doi: 10.3389/fcell.2021.782727

31. Dai ZJ, Yan CG, Li KC, Wang ZQ, Wang JH, Cao M, et al. Identifying and mapping connectivity patterns of brain network hubs in Alzheimer's disease. *Cereb Cortex.* (2015) 25:3723–42. doi: 10.1093/cercor/bhu246

32. Cheng NN, Elazab A, Yang P, Liu DD, Yu SZ, Wang TF, et al. Low rank self-calibrated brain network estimation and auto-weighted centralized multi-task learning for early mild cognitive impairment diagnosis. *Proceedings of the 2019 41st Annual International Conference of The IEEE Engineering in Medicine and Biology Society (EMBC)*. Berlin: (2019). doi: 10.1109/EMBC.2019.8856310

33. Sendi MSE, Zendehrouh E, Miller RL, Fu ZN, Du YH, Liu JY, et al. Alzheimer's disease projection from normal to mild dementia reflected in functional network connectivity: a longitudinal study. *Front Neural Circuits.* (2021) 14:593263. doi: 10.3389/fncir.2020.593263

34. Meng XL, Liu JL, Fan X, Bian CY, Wei QP, Wang ZW, et al. Multi-modal neuroimaging neural network-based feature detection for diagnosis of Alzheimer's disease. *Front Aging Neurosci.* (2022) 14:911220. doi: 10.3389/fnagi.2022.911220

35. Hu ZS, Zhang J, Zhang LY, Xiang YT, Yuan Z. Linking brain activation to topological organization in the frontal lobe as a synergistic indicator to characterize the difference between various cognitive processes of executive functions. *Neurophotonics.* (2019) 6:111–24. doi: 10.1117/1.NPh.6.2.025008

36. Henson RN, Greve A, Cooper E, Gregori M, Simons JS, Geerligs L, et al. The effects of hippocampal lesions on MRI measures of structural and functional connectivity. *Hippocampus.* (2016) 26:1447–63. doi: 10.1002/hipo.22621

37. Li ZL, Zhu QF, Geng ZJ, Song ZH, Wang LX, Wang Y. Study of functional connectivity in patients with sensorineural hearing loss by using resting-state fMRI. *Int J Clin Exp Med.* (2015) 8:569–78.

38. Epperly T, Dunay MA, Boice JL. Alzheimer disease: pharmacologic and nonpharmacologic therapies for cognitive and functional symptoms. *Am Fam Phys.* (2017) 95:771–8.

39. Gomar JJ, Bobes-Bascaran MT, Conejero-Goldberg C, Davies P, Goldberg TE, Alzheimer's Dis N. Utility of combinations of biomarkers, cognitive markers, and risk factors to predict conversion from mild cognitive impairment to alzheimer disease in patients in the Alzheimer's disease neuroimaging initiative. *Arch Gen Psychiatry.* (2011) 68:961–9. doi: 10.1001/archgenpsychiatry.2011.96

40. Ma J, Lin Y, Hu C, Zhang J, Yi Y, Dai Z. Integrated and segregated frequency architecture of the human brain network. *Brain Struct Funct.* (2021) 226:335–50. doi: 10.1007/s00429-020-02174-8

41. Alickovic E, Kevric J, Subasi A. Performance evaluation of empirical mode decomposition, discrete wavelet transform, and wavelet packed decomposition for automated epileptic seizure detection and prediction. Biomed Signal Proc Control. (2018) 39:94–102.

42. Smith JS. The local mean decomposition and its application to EEG perception data. *J R Soc Interface.* (2005) 2:443–54. doi: 10.1016/j.bspc.2017.07.022

43. Guo C, Wu D. Canonical correlation analysis (CCA) based multi-view learning: an overview. *IEEE Trans Syst Man Cybernet.* (2019) 31:1863–83.

44. Zhang YY, Jiang X, Qiao LS, Liu MX. Modularity-guided functional brain network analysis for early-stage dementia identification. *Front Neurosci.* (2021) 15:720909. doi: 10.3389/fnins.2021.720909

45. Chavarria-Siles I, White T, De Leeuw C, Goudriaan A, Lips E, Ehrlich S, et al. Myelination-related genes are associated with decreased white matter integrity in schizophrenia. *Eur J Hum Genet.* (2016) 24:381–6. doi: 10.1038/ejhg.2015.120

46. Chu Y, Wang GY, Cao L, Qiao LS, Liu MX. Multi-scale graph representation learning for autism identification with functional MRI. *Front Neuroinform.* (2022) 15:802305. doi: 10.3389/fninf.2021.802305

47. Yan CG, Chen X, Li L, Castellanos FX, Bai TJ, Bo QJ, et al. Reduced default mode network functional connectivity in patients with recurrent major depressive disorder. *Proc Natl Acad Sci USA.* (2019) 116:9078–83.

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 09 January 2023
DOI 10.3389/fpsyt.2022.1105496

OPEN ACCESS

EDITED BY
Weikai Li,
Chongqing Jiaotong University, China

REVIEWED BY
Liling Peng,
Universal Medical Imaging Diagnostic, China
Yanyan Xu,
China-Japan Friendship Hospital, China

*CORRESPONDENCE
Ning Wu
✉ ning.wu@ccmu.edu.cn

SPECIALTY SECTION
This article was submitted to Computational Psychiatry, a section of the journal Frontiers in Psychiatry

RECEIVED 22 November 2022
ACCEPTED 12 December 2022
PUBLISHED 09 January 2023

CITATION
Yu H, Wang Z, Sun Y, Bo W, Duan K, Song C, Hu Y, Zhou J, Mu Z and Wu N (2023) Prognosis of ischemic stroke predicted by machine learning based on multi-modal MRI radiomics. *Front. Psychiatry* 13:1105496. doi: 10.3389/fpsyt.2022.1105496

Prognosis of ischemic stroke predicted by machine learning based on multi-modal MRI radiomics

Huan Yu[1], Zhenwei Wang[1], Yiqing Sun[1], Wenwei Bo[1], Kai Duan[1], Chunhua Song[1], Yi Hu[1], Jie Zhou[1], Zizhang Mu[2] and Ning Wu[3]*

[1]Department of Radiology, Liangxiang Hospital, Beijing, China, [2]Department of Neurology, Liangxiang Hospital, Beijing, China, [3]Department of Medical Imaging, Yanjing Medical College, Capital Medical University, Beijing, China

Objective: Increased risk of stroke is highly associated with psychiatric disorders. We aimed to conduct the machine learning model based on multi-modal magnetic resonance imaging (MRI) radiomics predicting the prognosis of ischemic stroke.

Methods: This study retrospectively analyzed 148 patients with acute ischemic stroke due to anterior circulation artery occlusion. Based on the modified Rankin Scale (mRS) score, patients were divided into good (mRS $\leq$ 2) and poor (mRS $>$ 2) outcome groups. Segmentation of the infarct region was performed by manually outlining a mask of the lesion on diffusion-weighted images (DWI) using MRIcron software. The apparent diffusion coefficient (ADC), fluid decay inversion recoverage (FLAIR), susceptibility weighted imaging (SWI) and T1-weighted (T1w) images were aligned to the DWI images and the radiomic features within the lesion area were extracted for each image modality. The calculations were done using pyradiomics software and a total of 4,744 stroke-related imaging features were automatically calculated. Next, feature selection based on recursive feature elimination was used for each modality and three radiomic features were extracted from each modality plus one feature from the lesion mask, for a total of 16 radiomic features. At last, five machine learning (ML) models were trained and tested to predict stroke prognosis, calculate the received operating characteristic (ROC) curves and other parameters, evaluate the performance of the models and validate their predictive efficacy by five-fold cross-validation.

Results: Sixteen radiomic features were selected to construct the ML models for prognostic classification. By five-fold cross-validation, light gradient boosting machine (LightGBM) model-based muti-modal MRI radiomic features performed best in binary prognostic classification with accuracy of 0.831, sensitivity of 0.739, specificity of 0.902, F1-score of 0.788 and an area under the curve (AUC) of 0.902.

Conclusion: The ML models based on muti-modal MRI radiomics are of high value for predicting clinical outcomes in acute stroke patients.

KEYWORDS

diffusion-weighted imaging, radiomics, machine learning, ischemic stroke, magnetic resonance imaging

1. Introduction

Psychiatric comorbidities, such as depression (1), anxiety (2) and dementia (3), are the frequent consequences of stroke, which is one of the leading causes of disability and death worldwide (4). The psychiatric disorders make the prognosis of stroke complicated, and the prognosis varies greatly depending on the time of consultation and treatment, which lead to a challenge in deciding of "when to treat" and "how to treat" during rehabilitation of stroke patients (5). The accurate prediction of rehabilitation outcomes will do great help to propose the appropriate treatment strategies and rehabilitation goals based on each patient's condition (6).

The combination of the multi-modal magnetic resonance imaging (MRI) techniques provides a powerful tool for stroke diagnosis. Mitra et al. (7) used the information from multimodal [T1-weighted, T2-weighted, fluid attenuated inversion recovery (FLAIR), and apparent diffusion coefficient (ADC)] MRI images to extract areas with high likelihood of being classified as stroke lesions. Radiomics is an emerging approach that combines imaging and artificial intelligence to extract quantitative features from images in high throughput. Zhang et al. (5) developed the machine learning model-based diffusion weighed imaging (DWI)/ADC radiomic features to classify ischemic stroke onset time. Quan et al. (8) constructed the unfavorable outcome model based on the radiomic feature extracted from FLAIR and ADC image. Moreover, susceptibility weighted imaging (SWI), reflecting the oxygen extraction fraction of brain tissues, has been demonstrated as a useful predictor of early infarct size and early-stage clinical prognosis in acute ischemic stroke (9).

In this study, we constructed five machine learning (ML) models that aimed to predict the prognosis of ischemic stroke patients based on muti-modal MRI radiomics. In addition, we assessed predictive value of the models for ischemic stroke treatment decision-making.

2. Materials and methods

2.1. Participants

This study was a retrospective analysis of 180 patients diagnosed with acute ischemic stroke at Liangxiang Hospital (Beijing, China) from October 2020 to May 2022, of which 148 were included in analysis (**Figure 1**). The inclusion criteria were: (1) acute ischemic stroke due to anterior circulation artery occlusion; (2) MRI completed within 48 h of admission; (3) complete set of MRI sequences; (4) complete data on demographics and clinical characteristics; and (5) signed informed consent. The exclusion criteria were: (1) cerebral hemorrhage; (2) traumatic brain injury; (3) previous neurological or psychiatric disease; and (4) significant artifacts in MRI data. This study was approved by the ethics committee of Liangxiang Hospital (approval number 2016126).

Demographic characteristics as well as the clinical and imaging data were collected under the permission of patients. The National Institutes of Health Stroke Scale (NIHSS) score was collected to evaluate the degree of neurological deficit in stroke patients, which represented the level of consciousness, eye movements, integrity of visual fields, facial movements, arm and leg muscle strength, sensation, coordination, language, speech and neglect (10). Arranging from 0 to 42, the higher the NIHSS score, the more severe the neurological impairment: score 0 was normal neurological function, 1 to 4 was mild stroke, 5 to 15 was moderate stroke, 16 to 20 was moderate-severe stroke and 21 to 42 was severe stroke.

The modified Rankin Scale (mRS) score was used as a prognostic judgment index, with good prognosis defined as mRS scores of 0, 1, and 2, and poor prognosis defined as mRS scores of 3, 4, and 5. Of the 148 patients included in the analysis, 83 were in the good prognosis group and 65 were in the poor prognosis group.

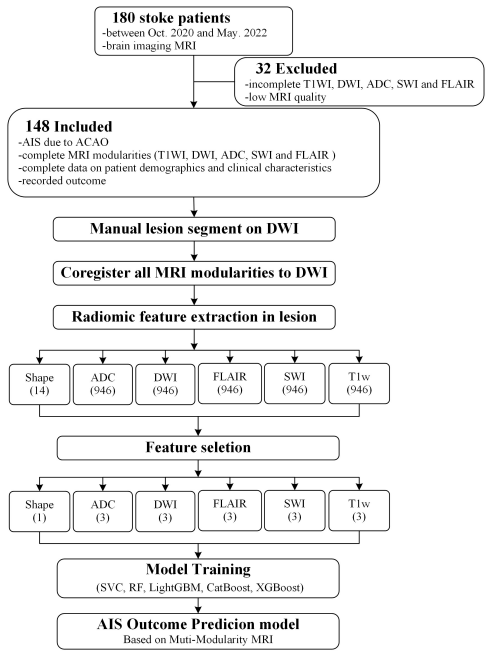


FIGURE 1
Flow chart of participant selection and analysis. DWI, diffusion-weighted imaging; ADC, apparent diffusion coefficient; FLAIR, fluid attenuated inversion recovery; SWI, susceptibility weighted imaging; ML, machine learning; SVM, support vector machine, RF, random forest; LightGBM, light gradient boosting; CatBoost, category boosting, and XGBoost, eXtreme gradient boosting.

2.2. MR image acquisition

MRI scans were performed within 3 days of stroke onset, using a Magnetom Skyra 3.0T MRI scanner (Siemens, Germany) with a 20-channel phased-array head coil.

All participants underwent the following scans:

(1) T1-weighted image scan. Scan parameters: T1w sequence, repetition time (TR) = 2,000 ms, inversion time (TI) = 900 ms, echo time (TE) = 8.8 ms, matrix = 209 × 256, field of view (FOV) = 220 mm^2 × 196 mm^2, thickness = 5 mm. number of layers = 24 layers, and parallel imaging factor = 2.

(2) Cerebrospinal fluid suppression image. Scan parameters: T2-FLAIR sequence, TR = 6,000 ms, TI = 2,028 ms, TE = 72 ms, matrix = 320 × 261, FOV = 220 mm^2 × 196 mm^2, thickness = 5 mm, number of layers = 24 layers, and parallel imaging factor = 2.

(3) SWI imaging scan sequence. Scan parameters: 3D-GRE sequence, TR = 27 ms, TE = 20 ms, flip angle (FA) = 15°, matrix = 256 × 256, FOV = 220 mm^2 × 196 mm^2, layer thickness = 2.5 mm, number of layers = 44, repetition number = 1, fat suppression on, and parallel imaging factor = 2.

(4) DWI imaging scan sequence. Scan parameters: EPI-Resolve sequence, b-value b = 1,000 and 0 scan, TR = 500 ms, TE1 = 63 ms, TE2 = 103 ms, FA = 180°, matrix = 160 × 160, FOV = 220 mm^2 × 220 mm^2, layer thickness = 5 mm, number of layers = 24, fat suppression on, and parallel imaging factor = 2.

2.3. Image processing and segmentation

Image analysis was performed independently by two MR diagnosticians blinded to the groups. The region of interest (ROI) of acute ischemic lesions was manually outlined layer by layer on the DWI images using MRIcron software.[1]

For each patient, data including all modalities (ADC, FLAIR, SWI, T1w) were aligned to the DWI images using SPM12 software[2] so that the outlined lesions could be directly used for texture feature extraction in the different modal images (**Figure 2**).

2.4. Radiomic feature extraction

For each patient, the MRI data of the five modalities was analyzed by pyradiomics software[3] for radiomiscs feature extraction using the recommended settings and steps for MRI data: (1) Images were resampled to 3 mm^2 × 3 mm^2 × 3 mm. (2) DWI, FLAIR, SWI, and T1w images are weighted images and thus needed to be numerically standardization. A scale of 100 was used and the binwith was set to 5. Since ADC images are quantitative, no numerical standardization was done and binwith was set to 20. The above settings ensured that the total number of bins is between 16 and 128. (3) Texture feature extraction was performed on the original images and filtered images, where the filter consisted of Laplacian of Gaussian filter based on sigma = 3 and 5 mm, edge enhancement filter, and 8 wavelet transforms (combination of high pass and low pass in three dimensions). (4) Finally, the texture features of radiomics were extracted, including 18 first order, 22 glcm, 16 glrlm, 16 glszm, and 14 gldm features. Thus, a total of 946 features were extracted per image. The shape features of 14 lesion regions were also extracted.

2.5. Feature selection and model training based on ML

The scikit-learn package[4] was used for feature selection of the MRI data of the five modalities. The recursive feature elimination (RFE) feature extraction method was used and

1 http://www.itk-snap.org
2 https://www.fil.ion.ucl.ac.uk/spm/
3 https://www.radiomics.io/pyradiomics.html
4 https://scikit-learn.org/

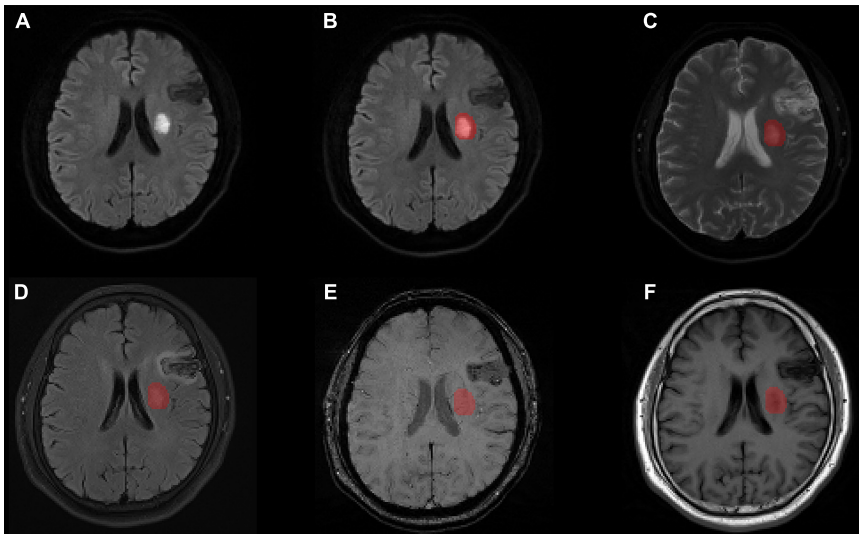


FIGURE 2
Lesion segmentation results for one patient with manual outlining of the lesion using MRIcron on **(A)** the original DWI image, **(B)** the lesion superimposed onto the DWI image, **(C)** the lesion superimposed onto the ADC image, **(D)** the lesion superimposed onto the FLAIR image aligned to the DWI image, **(E)** the lesion superimposed onto the SWI image aligned to the DWI image, and **(F)** the lesion superimposed onto the T1w image aligned to the DWI image. MRI, magnetic resonance imaging; ADC, apparent diffusion coefficient; DWI, diffusion-weighted imaging; FLAIR, fluid attenuated inversion recovery; SWI, susceptibility weighted imaging; T1w, T1-weighted.

TABLE 1 Baseline demographic and clinical characteristics.

Characteristics	Good prognosis (mRS ≤ 2) (*n* = 83)	Poor prognosis (mRS > 2) (*n* = 65)	t/χ^2	*P*-value
Age, year, mean ± SD	59.21 ± 10.94	70.80 ± 10.92	−6.404	0.001
Male, *n* (%)	62 (74.70)	38 (58.46)	4.386	0.036
NIHSS score, mean ± SD	2.33 ± 1.75	9.75 ± 5.65	−10.222	0.001
Hypertension, *n* (%)	61 (73.49)	56 (86.15)	3.528	0.060
Diabetes, *n* (%)	43 (51.81)	27 (41.54)	1.542	0.214
History of coronary heart disease, *n* (%)	4 (4.82)	13 (20.00)	8.263	0.004
History of atrial fibrillation, *n* (%)	3 (3.61)	11 (16.92)	7.539	0.006
Smoking, *n* (%)	55 (66.27)	31 (47.69)	5.166	0.023
Drinking, *n* (%)	45 (54.22)	22 (33.85)	6.105	0.013
Complications, *n* (%)	1 (1.20)	36 (55.38)	57.069	0.001

NIHSS, National Institutes of Health Stroke Scale; mRS, modified Rankin Scale.

TABLE 2 Results of feature selection for each MRI modality.

	Feature 1	Feature 2	Feature 3
Shape	MeshVolume		
DWI	log-sigma-3-0-mm-3D_glrlm_ LowGrayLevelRunEmphasis	log-sigma-3-0-mm-3D_glrlm_ ShortRunLowGrayLevelEmphasis	wavelet-LLH_glcm_Idn
ADC	log-sigma-5-0-mm-3D_firstorder_Maximum	log-sigma-5-0-mm-3D_firstorder_TotalEnergy	wavelet-HLL_gldm_ LargeDependenceHighGrayLevelEmphasis
FLAIR	original_firstorder_90Percentile	log-sigma-3-0-mm-3D_glszm_LargeAreaEmphasis	log-sigma-3-0-mm-3D_glszm_ LargeAreaHighGrayLevelEmphasis
SWI	log-sigma-5-0-mm-3D_glcm_Idmn	wavelet-HHH_glcm_Imc1	wavelet-HHH_glcm_Imc2
T1w	original_glszm_ZoneVariance	log-sigma-5-0-mm-3D_glszm_ LargeAreaLowGrayLevelEmphasis	wavelet-LLL_glcm_Imc1

ADC, apparent diffusion coefficient; DWI, diffusion-weighted imaging; FLAIR, fluid attenuated inversion recovery; SWI, susceptibility weighted imaging; T1w, T1-weighted.

TABLE 3 Results of different models for the test set (average results of five-fold).

Model	Accuracy	Sensitivity	Specificity	Precision	Recall	F1-score
SVM	0.791	0.631	0.915	0.858	0.631	0.722
RF	0.818	0.723	0.891	0.838	0.723	0.773
LightGBM	0.831	0.739	0.902	0.875	0.739	0.787
CatBoost	0.812	0.662	0.928	0.876	0.662	0.748
XGBoost	0.804	0.708	0.878	0.833	0.708	0.753

SVM, support vector machine Classifier; RF, random forest; LightGBM, light gradient boosting machine; XGB, extreme gradient boosting; F1-score = 2 × (precision × recall)/(precision + recall).

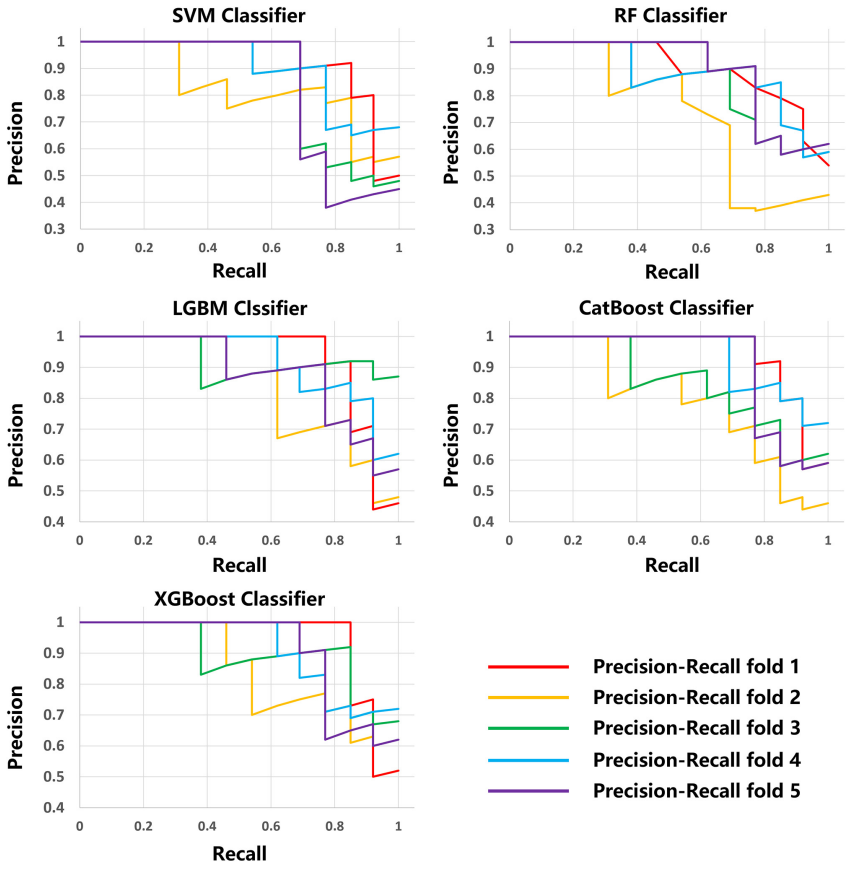


FIGURE 3
Precision-recall curves of the five models on five-fold cross-validation.

only the three best features were retained for each modality. For lesion shape features, we used the same method to retain the one best feature. Finally, a total of 16 image features were retained and used to train the ML model with the scikit-learn tool. A total of five methods, including Support Vector Machine (SVM) Classifier, Random Forest (RF) Classifier, Light Gradient Boosting Machine (LightGBM) Classifier, Category Boosting (CatBoost) Classifier, and eXtreme Gradient Boosting (XGBoost) Classifier, were used to build the models. Model performance was evaluated by a five-fold stratified cross-validation process. The evaluation metrics included accuracy, precision, recall, F1 score, receiver operating characteristic (ROC) curve, area under the curve (AUC), precision recall curve.

2.6. Statistical analysis

Statistical analysis was performed using SPSS 21.0 software. The independent samples *t*-test was used to compare the measurement data. Results with $p < 0.05$ were considered to be statistically significant differences.

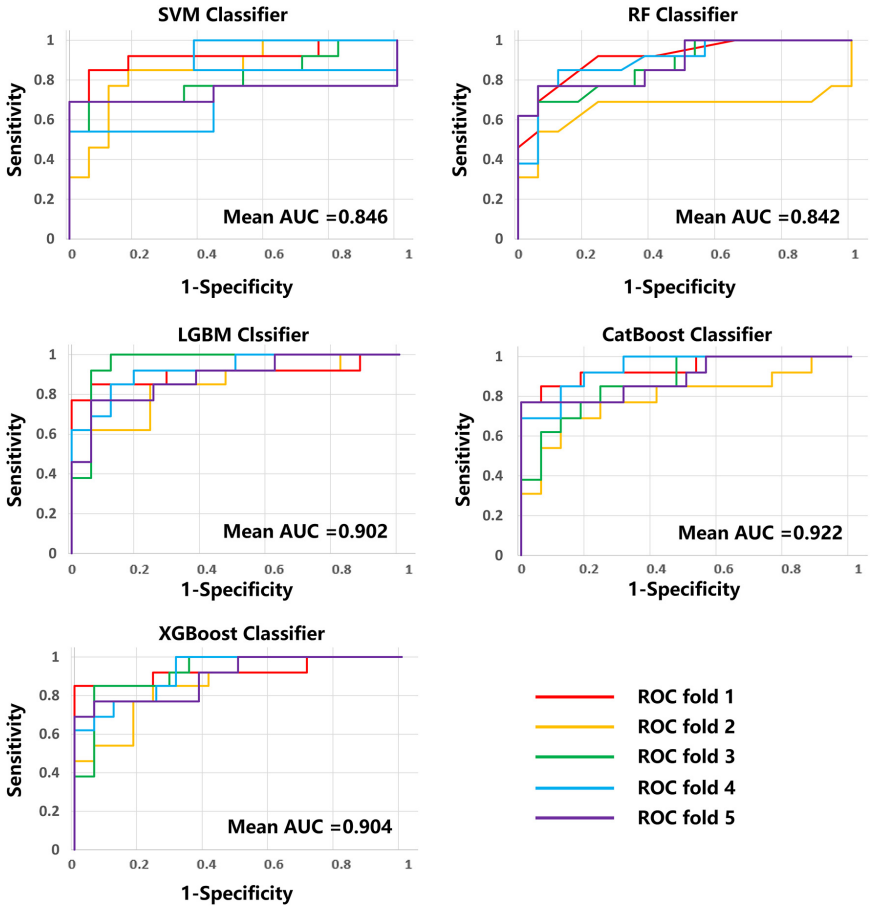


FIGURE 4
ROC curves and AUC values of the five models on five-fold cross-validation. ROC, received operating characteristic; AUC, area under the curve.

3. Results

3.1. Demographic characteristics of patients

Of the 148 patients, 83 (56.1%) had a good prognosis and 65 (43.9%) had a poor prognosis. The training set comprised 104 patients and the remaining 44 were used to test the ML model. For the overall sample, the mean age was 64.29 years and the mean NIHSS score was 5.59. These two variables differed significantly between groups ($p < 0.05$). No significant differences were found between patient groups for other baseline clinical characteristics (all $p > 0.05$, Table 1).

3.2. Radiomic feature extraction and selection

Three best radiomic features for each MRI modality (DWI, ADC, FLAIR, SWI and T1w) and one best feature for lesion shape were selected as features in the ML models. The detailed information about the features is presented in Table 2.

3.3. Training and evaluation of ML prediction models

The average results obtained for the test set using the different classification models after five-fold cross-validation are as follows: SVM model with 79% accuracy, RF model with 82% accuracy, LightGBM model with 83% accuracy, CatBoost model with 81% accuracy, and XGBoost model with 80% accuracy. The full model evaluation results are shown in Table 3 and Figures 3, 4.

4. Discussion

Early and accurate determination of disease progression may be important for new stroke patients, allowing timely

targeted treatment and effective improvement. To predict the prognosis of AIS early and accurately, this paper investigated five ML models based on multi-modal (T1w, ADC, DWI, FLAIR, and SWI) MRI radiomic features to predict AIS prognosis. The results showed that LightGBM model performed best in binary prognostic classification with accuracy of 0.831, sensitivity of 0.739, specificity of 0.902, F1-score of 0.788 and an area under the curve (AUC) of 0.902.

At present, most studies on the prognosis of ischemic stroke have used retrospective cohort studies to perform traditional statistical analysis of stroke prognosis models using Cox regression and logistic regression (11–13). Previous studies have failed to make full use of MRI data, resulting in low prediction accuracy (14–16). Several studies support that ML can predict stroke prognosis more accurately (17–19). Wang et al. (20) showed that, despite variability, current ML-based prognosis prediction of stroke patients has great potential. Qu et al. (21) used ML of retinal images to assess risk in 771 patients with ischemic and hemorrhagic stroke, achieving sensitivity and specificity of ischemic stroke risk assessment values of 91.0 and 94.8%, respectively. The area under the ROC curve for ischemic stroke was 0.929. Cui et al. (22) applied ML to develop and validate the incidence and severity of acute ischemic stroke in 1,100 patients. The combination of ML methods (e.g., complex neural networks) with imaging omics seems particularly promising, especially for the identification and segmentation of small lesions (23–25). Macciocchi et al. (26) performed a 3 month systematic evaluation of ischemic stroke and concluded that characteristics, such as age, previous stroke, initial neurological deficit, and lesion location, were highly correlated with functional outcome. The current results are consistent with those of previous studies, suggesting that imaging histology scores, hemorrhage, age, and NIHSS at 24 h are independent indicators of clinical outcome in patients with ischemic stroke. By combining these independent risk factors to generate a new imaging histology line graph, several studies have reported an association of DWI-derived ADC changes with functional outcome in ischemic stroke (27). A previous study reported that DWI had a 90% probability of identifying a lesion within 3 h prior to symptom onset (28). The present study suggested that radiomic features based on multi-modal MRI could predict clinical outcomes in acute stroke patients with accuracy of 0.831.

This study has provided new clues for predicting the prognosis of AIS and demonstrated the ability of multi-modal based radiomics to accurately predict the clinical functional outcome of AIS, contributing to the prevention of post-stroke psychiatric diseases. However, this study still has several limitations. First, this was a retrospective study with selection bias. Studies using larger samples are needed to further validate the predictive efficacy of the model. Second, this study did not differentiate the etiology and site of stroke, and manual outlining of ROI was affected by individual subjective factors. These clinical and imaging data should be considered in further study in the next step.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

This study was approved by the Ethics Committee of Liangxiang Hospital (approval number: 2016126). Written informed consent to participate was obtained from all subjects.

Author contributions

HY and ZW designed the study and drafted the manuscript. YS, WB, KD, and CS collected the MRI data. YH and JZ analyzed and interpreted the results of the data. ZM and NW revised the manuscript. All authors approved the final manuscript.

Funding

This study was supported by the Capital's Funds for Health Improvement and Research (grant number: 2020-4-7073).

Conflict of interest

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2022.1105496/full#supplementary-material

References

1. Hackett, M, Pickles K. The frequency of depression after stroke: an updated systematic review and meta-analysis of observational studies. *Int J Stroke.* (2013) 8:15. doi: 10.1111/ijs.12357

2. Rafsten L, Danielsson A, Sunnerhagen KS. Anxiety after stroke: a systematic review and meta-analysis. *J Rehabil Med.* (2018) 50:769–78.

3. Leys D, Hénon H, Mackowiak-Cordoliani M-A, Pasquier F. Poststroke dementia. *Lancet Neurol.* (2005) 4:752–9.

4. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, et al. Heart disease and stroke statistics—2016 update: a report from the American heart association. *Circulation.* (2016) 133:e38–360.

5. Zhang YQ, Liu AF, Man FY, Zhang YY, Li C, Liu YE, et al. MRI radiomic features-based machine learning approach to classify ischemic stroke onset time. *J Neurol.* (2022) 269:350–60. doi: 10.1007/s00415-021-10638-y

6. Su YY, Xiao SY, Haupt WF, Zhang Y, Zhao H, Pang Y, et al. Parameters and grading of evoked potentials: prediction of unfavorable outcome in patients with severe stroke. *J Clin Neurophysiol.* (2010) 27:25–9. doi: 10.1097/WNP.0b013e3181cb4282

7. Mitra J, Bourgeat P, Fripp J, Ghose S, Rose S, Salvado O, et al. Lesion segmentation from multimodal MRI using random forest following ischemic stroke. *NeuroImage.* (2014) 98:324–35. doi: 10.1016/j.neuroimage.2014.04.056

8. Quan G, Ban R, Ren JL, Liu Y, Wang W, Dai S, et al. FLAIR and ADC image-based radiomics features as predictive biomarkers of unfavorable outcome in patients with acute ischemic stroke. *Front Neurosci.* (2021) 15:730879. doi: 10.3389/fnins.2021.730879

9. Luo S, Yang L, Luo Y. Susceptibility-weighted imaging predicts infarct size and early-stage clinical prognosis in acute ischemic stroke. *Neurol Sci.* (2018) 39:1049–55. doi: 10.1007/s10072-018-3324-3

10. Kwah, LK, Diong J. National institutes of health stroke scale (NIHSS). *J Physiother.* (2014) 60:61.

11. Kaplan RC, Tirschwell D, Longstreth W, Manolio T, Heckbert S, Lefkowitz D, et al. Vascular events, mortality, and preventive therapy following ischemic stroke in the elderly. *Neurology.* (2005) 65:835–42.

12. Song B, Liu Y, Nyame L, Chen X, Jiang T, Wang W, et al. A COACHS nomogram to predict the probability of three-month unfavorable outcome after acute ischemic stroke in Chinese patients. *Cerebrovasc Dis.* (2019) 47:80–7. doi: 10.1159/000497243

13. Matsuo R, Ago T, Kiyuna F, Sato N, Nakamura K, Kuroda J, et al. Smoking status and functional outcomes after acute ischemic stroke. *Stroke.* (2020) 51:846–52.

14. Yu Y, Guo D, Lou M, Liebeskind D, Scalzo F. Prediction of hemorrhagic transformation severity in acute stroke from source perfusion MRI. *IEEE Trans Biomed Eng.* (2017) 65:2058–65. doi: 10.1109/TBME.2017.2783241

15. Tang TY, Jiao Y, Cui Y, Zeng CH, Zhao DL, Zhang Y, et al. Development and validation of a penumbra-based predictive model for thrombolysis outcome in acute ischemic stroke patients. *EBioMedicine.* (2018) 35:251–9. doi: 10.1016/j.ebiom.2018.07.028

16. Jiang L, Zhou L, Yong W, Cui J, Geng W, Chen H, et al. A deep learning-based model for prediction of hemorrhagic transformation after stroke. *Brain Pathol.* (2021). [Epub ahead of print]. doi: 10.1111/bpa.13023

17. Benzakoun J, Charron S, Turc G, Hassen WB, Legrand L, Boulouis G, et al. Tissue outcome prediction in hyperacute ischemic stroke: comparison of machine learning models. *J Cereb Blood Flow Metab.* (2021) 41:3085–96.

18. Kuang H, Qiu W, Boers AM, Brown S, Muir K, Majoie CB, et al. Computed tomography perfusion–based machine learning model better predicts follow-up infarction in patients with acute ischemic stroke. *Stroke.* (2021) 52:223–31. doi: 10.1161/STROKEAHA.120.030092

19. He Y, Luo Z, Zhou Y, Xue R, Li J, Hu H, et al. U-net models based on computed tomography perfusion predict tissue outcome in patients with different reperfusion patterns. *Transl Stroke Res.* (2022) 13:707–15. doi: 10.1007/s12975-022-00986-w

20. Wang X, Fan Y, Zhang N, Li J, Duan Y, Yang B. Performance of machine learning for tissue outcome prediction in acute ischemic stroke: a systematic review and meta-analysis. *Front Neurol.* (2022) 13:910259. doi: 10.3389/fneur.2022.910259

21. Qu Y, Zhuo Y, Lee J, Huang X, Yang Z, Yu H, et al. Ischemic and haemorrhagic stroke risk estimation using a machine-learning-based retinal image analysis. *Front Neurol.* (2022) 13:916966. doi: 10.3389/fneur.2022.916966

22. Cui J, Yang J, Zhang K, Xu G, Zhao R, Li X, et al. Machine learning-based model for predicting incidence and severity of acute ischemic stroke in anterior circulation large vessel occlusion. *Front Neurol.* (2021) 12:749599. doi: 10.3389/fneur.2021.749599

23. Azer SA. Deep learning with convolutional neural networks for identification of liver masses and hepatocellular carcinoma: a systematic review. *World J Gastrointest Oncol.* (2019) 11:1218. doi: 10.4251/wjgo.v11.i12.1218

24. Munir K, Elahi H, Ayub A, Frezza F, Rizzi A. Cancer diagnosis using deep learning: a bibliographic review. *Cancers.* (2019) 11:1235.

25. Sollini M, Antunovic L, Chiti A, Kirienko M. Towards clinical application of image mining: a systematic review on artificial intelligence and radiomics. *Eur J Nucl Med Mol Imaging.* (2019) 46:2656–72. doi: 10.1007/s00259-019-04372-x

26. Macciocchi SN, Diamond PT, Alves WM, Mertz T. Ischemic stroke: relation of age, lesion location, and initial neurologic deficit to functional outcome. *Arch Phys Med Rehabil.* (1998) 79:1255–7. doi: 10.1016/s0003-9993(98)90271-4

27. Miles J. Wiley statsref: statistics reference online. In: Balakrishnan N, Colton T, Everitt B, Piegorsch W, Ruggeri F, Teugels J, editors. *R squared, adjusted R squared*. Chichester: Wiley (2014).

28. Huisa BN, Liebeskind DS, Raman R, Hao Q, Meyer BC, Meyer DM, et al. Diffusion-weighted imaging–fluid attenuated inversion recovery mismatch in nocturnal stroke patients with unknown time of onset. *J Stroke Cerebrovasc Dis.* (2013) 22:972–7. doi: 10.1016/j.jstrokecerebrovasdis.2012.01.004

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 10 January 2023
DOI 10.3389/fpsyt.2022.1100266

OPEN ACCESS

EDITED BY
Weikai Li,
Chongqing Jiaotong University, China

REVIEWED BY
Haifeng Chen,
Nanjing Drum Tower Hospital, China
Weiwei Wang,
Fudan University, China

*CORRESPONDENCE
Xin Gao
✉ gaoxin@uvclinic.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Computational Psychiatry, a section of the journal Frontiers in Psychiatry

RECEIVED 16 November 2022
ACCEPTED 15 December 2022
PUBLISHED 10 January 2023

CITATION
Ma D, Peng L and Gao X (2023) Adaptive noise depression for functional brain network estimation. *Front. Psychiatry* 13:1100266. doi: 10.3389/fpsyt.2022.1100266

Adaptive noise depression for functional brain network estimation

Di Ma[1,2†], Liling Peng[2†] and Xin Gao[2*]

[1]College of Information Science and Technology, Nanjing Forestry University, Nanjing, China, [2]Department of PET/MR, Shanghai Universal Medical Imaging Diagnostic Center, Shanghai, China

Autism spectrum disorder (ASD) is one common psychiatric illness that manifests in neurological and developmental disorders, which can last throughout a person's life and cause challenges in social interaction, communication, and behavior. Since the standard ASD diagnosis is highly based on the symptoms of the disease, it is difficult to make an early diagnosis to take the best cure opportunity. Compared to the standard methods, functional brain network (FBN) could reveal the statistical dependence among neural architectures in brains and provide potential biomarkers for the early neuro-disease diagnosis and treatment of some neurological disorders. However, there are few FBN estimation methods that take into account the noise during the data acquiring process, resulting in poor quality of FBN and thus poor diagnosis results. To address such issues, we provide a brand-new approach for estimating FBNs under a noise modeling framework. In particular, we introduce a noise term to model the representation errors and impose a regularizer to incorporate noise prior into FBNs estimation. More importantly, the proposed method can be formulated as conducting traditional FBN estimation based on transformed fMRI data, which means the traditional methods can be elegantly modified to support noise modeling. That is, we provide a plug-and-play noise module capable of being embedded into different methods and adjusted according to different noise priors. In the end, we conduct abundant experiments to identify ASD from normal controls (NCs) based on the constructed FBNs to illustrate the effectiveness and flexibility of the proposed method. Consequently, we achieved up to 13.04% classification accuracy improvement compared with the baseline methods.

KEYWORDS

Autism spectrum disorder, functional brain network, Pearson's correlation, adaptive noise depression, functional magnetic resonance imaging

1. Introduction

Autism spectrum disorder (ASD) is one of the most common psychiatric illnesses characterized by repetitive behaviors and persistent impairments in communication and interaction (1, 2). Referring to a current report provided by the CDC of the USA (3), the overall ASD prevalence is rapidly increasing, rising from 6.7 per 1,000 children aged 8 years in surveillance years 2000 and 2002 to 23.0 in the surveillance year 2018.

However, doctor's diagnosis of ASD highly depends on people's developmental history and behavior, which may result in delayed final diagnosis and thus missed early help (4). To get an early diagnosis, gene-level measurements can help, but their high cost and complexity impedes the spread (5, 6). Recently, researchers have shown that unusual brain activity and abnormal functional disruptions in brain regions highly correlate with ASD, making it capable to discover informative biomarkers and analyze brain activity to help diagnose ASD (7–9).

As a successful non-invasive technique for measuring brain activity, functional magnetic resonance imaging (fMRI) has been successfully used to aid in the early diagnosis of ASD (10–12). Functional brain network (FBN) is one of the most popular tools to help diagnosis based on the fMRI data (13–15). Generally, FBN is constructed by the brain regions of interest (ROIs) and their correlations, such as the statistical dependence between different ROIs (16, 17). Compared to the current study, which directly utilizes the fMRI data to identify the ASD from normal controls (NCs), the FBN-based methods can provide more stable measurements among neural time series of the brain that highly relates to some neurological diseases, including ASD, mild cognitive impairment (MCI) (18), Alzheimer's disease (AD) (19), and chronic tinnitus (CT) (20).

Second-order statistics are most commonly used to estimate FBN, and typical models include Pearson's correlation (PC) (21) and sparse representation (SR) (22). PC enjoys an efficient and robust FBN estimation that captures the full correlation among ROIs, but it has dense connections and is affected by confounding effects from other brain regions (16). Instead, SR can capture the partial correlation that eliminates the potential effects of other brain regions. However, the computation of the inverse covariance matrix is involved in partial correlation, which is ill-posed (23). Therefore, SR is equipped with an L1-norm regularizer to obtain a more stable partial correlation, leading to relatively lower computation efficiency than PC.

Despite their successful applications, the existing methods rarely take into account the data acquisition noise, which usually leads to poor FBN estimation and thus poor performance on disease identification. Before FBN estimation, the data preprocessing follows a standard pipeline to avoid influence caused by noisy signals (24), which is still not easy to filter out all the artifacts/noises from the data due to the weak fMRI signals. For example, some preprocessing steps may further introduce the notorious noisy time points into the data (e.g., spatial normalization) (25).

To address the earlier issues, in this study, we propose a novel FBN estimation strategy by embedding a noise modeling term to depress the effects of noise on FBN estimation. The noise term measures the noises and their correlation among time series to capture the noise pattern implied in time series. We show that such a term appears to be a modification of the data-fitting term, which has a great influence on FBN estimation. Then we introduce a noise prior to constraining the noise pattern from a practical view. Consequently, our proposed method realizes to automatically and simultaneously model the noise and estimate FBN under a unified framework. In summary, the contributions of our proposed method are highlighted as follows:

1. Our method combines the noise depression and FBN estimation into a unified framework, making it capable of obtaining a clearer FBN estimation and higher accuracy on disease diagnosis.
2. The modification of the data-fitting term can be interpreted as a traditional fitting term on a transformed data series, making it possible to modify a series of traditional methods to support noise modeling. Meanwhile, such modifications can make good use of existing optimization methods, which are both cheap and convenient.
3. The noise pattern can be fitted with the help of extra prior knowledge, which can be independently adjusted to adapt to the current task.
4. The earlier two points jointly constitute a plug-and-play noise module capable of being embedded into different methods and adjusted according to different noise priors.

2. Materials and methods

2.1. Data acquisition and preprocessing

For the participants in this study, we simply utilize a well-known and publicly available dataset, that is, Autism Brain Imaging Data Exchange (ABIDE) as in several recent studies (16, 26). The autism criteria sets in the Diagnostic and Statistical Manual of Mental Disorders, 4th Edition, Text Revision (DSM-IV-TR) (27) are adopted to diagnose ASD from NC. Above all, 45 ASD subjects (36 males and nine females) and 47 NC subjects (36 males and 11 females) between 7 and 15 year of age are included, with gender, age, and full intelligence quotient (FIQ) not differing significantly from ASD to NC. The detailed demographic information of the participant is given in Table 1.

The resting-state fMRI scanning of all subjects is conducted using a 3T Siemens Allegra scanner within 6 min. During the scanning procedure, all subjects were asked to relax with their eyes open and focus on a white fixed cross in the middle of a black background projected on a screen. These requirements ensure the subjects to focus their attention and prevent meditation with the eyes closed, thereby avoiding violent neural activity. The parameters for acquiring images include: flip angle = 90°, 180 volumes per scan, 33 slices per volume, TR/TE = $2,000/15$ *ms*, and 4.0-*mm* voxel thickness (28).

After data acquisition, we conduct the statistical parametric mapping (SPM8) (http://www.fil.ion.ucl.ac.uk/spm/software/spm8/) as the preprocessing toolbox to preprocess the fMRI

TABLE 1 Demographic information of the subjects.

	ASD (N = 45)	NC (N = 47)	p-value
FIQ (mean ± SD)	106.8 ± 17.4	113.3 ± 14.1	0.0510
Age (year ± SD)	11.1 ± 2.3	11.1 ± 2.3	0.7773†
Gender (M/F)	36/9	36/11	0.2135*
ADOS (mean ± SD)	13.7 ± 5.0	-	-
ADI-R (mean ± SD)	32.2 ± 14.3‡	-	-

FIQ, full intelligence quotient; ADOS, Autism Diagnostic Observation Schedule; ADI-R, Autism Diagnostic Interview-Revised.
*The p-value was obtained by the chi-squared test.
†The p-value was obtained by a two-sample two-tailed t-test.
‡Two patients do not have the ADI-R score.

data. In particular, we remove the first 10 RS-fMRI images of each subject. The remaining images were spatially normalized into the Montreal Neurological Institute (MNI) template space with the resolution 3 × 3 × 3 *mm*. Then, the regression of nuisance signals (ventricle, white matter, global signals, and head motion with Friston 24-parameter model), signal detrending, and band-pass filtering (0.01–0.08 Hz) (29–31) are included for further corrections. After that, each preprocessed image was parcellated into 116 ROIs according to the automated anatomical labeling (AAL) atlas (32). Finally, these time series were as the data matrix, $\mathbf{X} \in \mathbb{R}^{170\times116}$, where 170 denotes the total number of temporal image volumes and 116 denotes the total number of ROIs. In our study, we focus on the first 90 ROIs that belong to the cerebrum as our regions of interest, which are utilized in most studies using AAL. Thus, the data matrix size is reformed as $\mathbf{X} \in \mathbb{R}^{170\times90}$.

2.2. Related methods

After data preparation, the next task is the FBN construction. In this study, we briefly review two FBN estimation approaches, that is, PC (33) and SR (22), which are all closely related to this study.

The notations used in the rest article are presented beforehand as follows. Bold uppercase letters are used to denote matrices, bold lowercase letters are used to denote vectors, and normal italic letters to denote scalars. The ith column of matrix $\mathbf{X}$ is denoted as $\mathbf{x}_i$, and the element of $\mathbf{X}$ at ith row and jth column is denoted as x_{ij}. $\|\cdot\|_F$, $\|\cdot\|_2$, and $\|\cdot\|_1$ denote the Frobenius norm, L2-norm, and L1-norm, respectively. $|\cdot|$ denotes the determinant of a matrix or the absolute value of a scalar. We further denote the transpose operator, the trace operator, and the inverse of a matrix $\mathbf{X}$ as $\mathbf{X}^T$, tr($\mathbf{X}$), and $\mathbf{X}^{-1}$.

2.2.1. Pearson's correlation

The Pearson's correlation is the simplest and most commonly used method for estimating FBNs (33). $\mathbf{X} \in \mathbb{R}^{T\times N}$ denote the fMRI data matrix (i.e., the BOLD signals), where T and N denote the number of time points in each series and the number of ROIs, respectively. $\mathbf{x}_i \in \mathbb{R}^T (i = 1, \cdots, N)$ denote the time series of the ith ROI, and then we can calculate the weight w_{ij} of the network connection between the ith and jth ROIs by PC as follows:

$$w_{ij} = \frac{(\mathbf{x}_i - \bar{\mathbf{x}}_i)^T(\mathbf{x}_j - \bar{\mathbf{x}}_j)}{\sqrt{(\mathbf{x}_i - \bar{\mathbf{x}}_i)^T(\mathbf{x}_i - \bar{\mathbf{x}}_i)}\sqrt{(\mathbf{x}_j - \bar{\mathbf{x}}_j)^T(\mathbf{x}_j - \bar{\mathbf{x}}_j)}}, \quad (1)$$

where $\bar{\mathbf{x}}_i \in \mathbb{R}^T$ denotes a mean vector with all entries being the mean value of all elements in $\mathbf{x}_i$.

Without the loss of generality, suppose that the fMRI data have been centralized and normalized by $\mathbf{x}_i = (\mathbf{x}_i - \bar{\mathbf{x}}_i)/\sqrt{(\mathbf{x}_i - \bar{\mathbf{x}}_i)^T(\mathbf{x}_i - \bar{\mathbf{x}}_i)}$, the weight can be simplified as the form $w_{ij} = \mathbf{x}_i^T\mathbf{x}_j$, which corresponds to the optimal solution of the following problem:

$$\min_{\mathbf{W}} \|\mathbf{W} - \mathbf{X}^T\mathbf{X}\|_F^2, \quad (2)$$

where $\mathbf{W} = (w_{ij}) \in \mathbb{R}^{N\times N}$ is the edge weight matrix of the estimated FBN. The above remodels PC in a perspective of optimization and benefits to develop new flexible FBN estimation methods based on PC (16).

In general, PC-based estimation methods produce a dense FBN, in which the ROIs are fully connected, with some connections being noisy or uninformative. In order to filter out such connections, thresholding or sparsity-induced constraint is generally used to sparsify the estimated FBN. For more details of the thresholding scheme, see Fornito et al. (34).

2.2.2. Sparse representation

Although PC is simple and empirically effective in building FBN, it can only measure the full correlation and neglect the interaction among multiple ROIs. Instead of measuring full correlation, PC-based methods aim to estimate more reliable connections between two ROIs by regressing out the confounding effect from other ROIs (22). Nevertheless, such an approach might be ill-posed due to the inverse calculation of a singular sample covariance matrix. To address this issue, an L1-norm regularizer is incorporated into the partial correlation model, resulting in the SR-based FBN estimation (22) as follows:

$$\min_{w_{ij}} \sum_{i=1}^{N} \left\| \mathbf{x}_i - \sum_{\substack{j=1 \\ j\neq i}}^{N} w_{ij}\mathbf{x}_j \right\|_2^2 + \lambda_1 \sum_{\substack{j=1 \\ j\neq i}}^{N} |w_{ij}| \quad (3)$$

which can be further rewritten as the equivalent matrix form:

$$\begin{aligned} \min_{\mathbf{W}} \quad & \|\mathbf{X} - \mathbf{X}\mathbf{W}\|_F^2 + \lambda_1 \|\mathbf{W}\|_1 \\ \text{s.t.} \quad & w_{ii} = 0, \forall i = 1, \cdots, N \end{aligned} \tag{4}$$

where constraint $w_{ii} = 0$ is used to remove $\mathbf{x}_i$ from $\mathbf{X}$ to avoid the trivial solution, λ_1 is a regularized parameter that controls the sparsity of the estimated FBN and benefits for achieving a stable solution (35).

2.3. Proposed methods

As two typical examples, PC and SR have been demonstrated to be more sensitive than some complex higher-order methods (13). Nevertheless, since the original data points in the time series possibly contain "noise", the estimated FBNs are often heavily influenced by the quality of the observed data and result in the representation noise of the network connections (25). To address this issue, in this section, we mainly focus on the baseline method SR and introduce a noise modeling scheme for FBN estimation.

2.3.1. SRAND: Sparse representation with adaptive noise depression

Suppose $\mathbf{X}$ to be the noisy fMRI data (assumed to be centralized and normalized), then the first term in Equation (3) can be rewritten as follows:

$$\begin{aligned} & \sum_{i=1}^{N} \left\| \mathbf{x}_i - \sum_{\substack{j=1 \\ j\neq i}}^{N} w_{ij}\mathbf{x}_j \right\|_2^2 \\ & = \sum_{i=1}^{N} \left\| \left(\mathbf{x}_i^{clear} + \mathbf{n}_i\right) - \sum_{\substack{j=1 \\ j\neq i}}^{N} w_{ij} \left(\mathbf{x}_j^{clear} + \mathbf{n}_j\right) \right\|_2^2 \\ & = \sum_{i=1}^{N} \left\| \left(\mathbf{x}_i^{clear} - \sum_{\substack{j=1 \\ j\neq i}}^{N} w_{ij}\mathbf{x}_j^{clear} \right) + \left(\mathbf{n}_i - \sum_{\substack{j=1 \\ j\neq i}}^{N} w_{ij}\mathbf{n}_j \right) \right\|_2^2 \end{aligned} \tag{5}$$

where the superscript *clear* denotes the clean data and $\mathbf{n}_i$ denotes the noise term of the ith ROI.

We can see that the total representation error in the objective function can be rewritten by the sum of the representation error on clean fMRI data and that caused by the noise of the fMRI data. However, traditional SR takes no account of the influence brought by such noise-caused error, and simply assumes that the representation error terms are identically independently distributed [i.i.d., equivalent to the L2-norm in Equation (3)]. To this end, we introduce to model of the error to depress the noise influence on FBN estimation. Specifically, we consider measuring the partial correlation and assuming the representation error between the ith ROI and the jth ROI follows Gaussian distribution with a non-diagonal precision matrix, that is,

$$\mathbf{x}_i \sim N\left(\mathbf{x}_i \middle| \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbf{x}_j w_{ij}, \boldsymbol{\Omega}^{-1} \right), \tag{6}$$

where $\boldsymbol{\Omega}$ denotes the precision matrix (i.e., the inverse covariance matrix). Here, non-diagonal condition on $\boldsymbol{\Omega}$ indicates the noise term to be non-i.i.d, which is more practical to measure the dependent relationship between noises among time series. Moreover, we assume that the precision matrix is identical for noise terms between different ROIs so that to capture the noise pattern implied in time series rather than ROIs.

Taking the negative logarithm of Equation (6), we obtain the maximum-likelihood estimation (MLE) of w_{ij} and $\boldsymbol{\Omega}$ by minimizing

$$\min_{w_{ij}, \boldsymbol{\Omega}} \quad \left(\mathbf{x}_i - \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbf{x}_j w_{ij} \right)^T \boldsymbol{\Omega} \left(\mathbf{x}_i - \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbf{x}_j w_{ij} \right) - \ln|\boldsymbol{\Omega}|, \tag{7}$$

In subsequent, we consider the optimization problem corresponding to w_{ij} by fixing $\boldsymbol{\Omega}$. We define a new transformed fMRI time series $\mathbf{y}_i = \boldsymbol{\Omega}^{\frac{1}{2}}\mathbf{x}_i$, then the objective term related to w_{ij} in the Equation (7) can be rewritten as follows:

$$\left(\mathbf{x}_i - \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbf{x}_j w_{ij} \right)^T \boldsymbol{\Omega} \left(\mathbf{x}_i - \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbf{x}_j w_{ij} \right) = \left\| \mathbf{y}_i - \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbf{y}_j w_{ij} \right\|_2^2. \tag{8}$$

Consequently, the optimization problem of w_{ij} embedded with the newly added noise depression module can be formulated as follows:

$$\min_{w_{ij}} \sum_{i=1}^{N} \left\| \mathbf{y}_i - \sum_{\substack{j=1 \\ j\neq i}}^{N} \mathbf{y}_j w_{ij} \right\|_2^2 + \lambda_1 \sum_{\substack{j=1 \\ j\neq i}}^{N} |w_{ij}| \tag{9}$$

or equivalently as its matrix form

$$\begin{aligned} \min_{\mathbf{W}} \quad & \|\mathbf{Y} - \mathbf{Y}\mathbf{W}\|_F^2 + \lambda_1 \|\mathbf{W}\|_1. \\ \text{s.t.} \quad & w_{ii} = 0, \forall i = 1, \cdots, N \end{aligned} \tag{10}$$

We can see that the above optimization problem coincides with the traditional SR on the transformed fMRI data series $\mathbf{y}_i$, which can be easily solved by existing SR-based optimization methods (22). In contrast, the coincidence promotes us to embed such noise modeling into other FBN estimation methods following a similar pipeline. In other words, such transformation on

fMRI data can be applied as a plug-and-play noise module to elegantly modify the traditional methods, making them capable of depressing noise.

2.3.2. Adaptive noise modeling with different priors

After the optimization of the edge weights w_{ij}, we turn to the noise modeling term $\mathbf{\Omega}$. We first reformulate Equation (7) into matrix form about $\mathbf{\Omega}$ as follows:

$$\min_{\mathbf{\Omega}} \quad \mathrm{tr}\left[\mathbf{\Omega}(\mathbf{X}-\mathbf{X}\mathbf{W})(\mathbf{X}-\mathbf{X}\mathbf{W})^T\right] - N\ln|\mathbf{\Omega}| \tag{11}$$

The element at the ith row and jth column of $\mathbf{\Omega}$ measures the noise relationship between the ith and jth time points. If the value approaches to zero, then the two time points are conditionally independent and vice versa. Such modeling can capture the noise dependence among time series, which is more practical than i.i.d. assumption.

However, similar to the estimation of FBN, due to the computation of $\mathbf{\Omega}$ involving inversing the covariance matrix, directly optimizing $\mathbf{\Omega}$ based on objective function (12) is often ill-posed. Thus, it makes sense to impose a prior to constrain the structure of $\mathbf{\Omega}$. We embed such prior *via* a regularizer on $\mathbf{\Omega}$ and present the following universal form for optimizing $\mathbf{\Omega}$,

$$\min_{\mathbf{\Omega}} \quad \mathrm{tr}\left[\mathbf{\Omega}(\mathbf{X}-\mathbf{X}\mathbf{W})(\mathbf{X}-\mathbf{X}\mathbf{W})^T\right] - N\ln|\mathbf{\Omega}| + \lambda_2 R(\mathbf{\Omega}), \tag{12}$$

where $R(\mathbf{\Omega})$ is a regularized term, λ_2 is a trade-off parameter.

Considering that only the noises which change with time series regularly appear to be correlated in $\mathbf{\Omega}$, while irregular noises among time series are usually independent of each other, it is natural to impose an L1-norm penalty to model such sparse structure, resulting in the following:

$$\min_{\mathbf{\Omega}} \quad \mathrm{tr}\left[\mathbf{\Omega}(\mathbf{X}-\mathbf{X}\mathbf{W})(\mathbf{X}-\mathbf{X}\mathbf{W})^T\right] - N\ln|\mathbf{\Omega}| + \lambda_2\|\mathbf{\Omega}\|_1, \tag{13}$$

which can be solved by existing optimization methods, for example, the method of Meinshausen and Bühlmann (36) or the classical graphical lasso (37).

By combining the objective functions and joining all constraints in Equations (10) and (13), the sparse representation based on adaptive noise depression (SRAND) with L1-norm constraint can be summarized as follows:

$$\begin{aligned} \min_{\mathbf{W},\mathbf{\Omega}} \quad & \mathrm{tr}\left[(\mathbf{X}-\mathbf{X}\mathbf{W})^T\mathbf{\Omega}(\mathbf{X}-\mathbf{X}\mathbf{W})\right] \\ & - N\ln|\mathbf{\Omega}| + \lambda_1\|\mathbf{W}\|_1 + \lambda_2\|\mathbf{\Omega}\|_1 \\ \text{s.t.} \quad & w_{ii} = 0, \forall i = 1, \cdots, N. \end{aligned} \tag{14}$$

Consequently, the procedure of SRAND with L1-norm is listed in Algorithm 1.

In addition to the above kind of prior that directly imposes specific structure on the noise pattern, in the following, we introduce another prior, Wishart distribution as the prior of $\mathbf{\Omega}$ to embed structure implicitly. In particular, the Wishart distribution is given by the following:

```
Input: Data matrix X, parameters λ1 and λ2
Output: Constructed FBN W
  Initialize Ω = I
  while not converge do
    Update W by solving the problem in Equation
    (10)
    Update Ω by solving the problem in Equation
    (13)
  end while
```

Algorithm 1. SRAND with L1-norm.

$$W(\mathbf{\Omega}|\mathbf{\Sigma},\nu) = B(\mathbf{\Sigma},\nu)|\mathbf{\Omega}|^{(\nu-T-1)/2}\exp\left(-\frac{1}{2}\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\mathbf{\Omega}\right)\right), \tag{15}$$

where

$$B(\mathbf{\Sigma},\nu) = |\mathbf{\Sigma}|^{-\nu/2}\left(2^{\nu T/2}\pi^{T(T-1)/4}\prod_{i=1}^{T}\Gamma\left(\frac{\nu+1-i}{2}\right)\right)^{-1}$$

Is a scaler irrelevant to $\mathbf{\Omega}$, ν is the *number of degrees of freedom* and restricted to $\nu > T-1$, $\mathbf{\Sigma}$ is a $N \times N$ symmetric, positive definite matrix.

Then we assume that the precision matrix $\mathbf{\Omega}$ is subject to Wishart distribution,

$$\mathbf{\Omega} \sim W(\mathbf{\Omega}|\mathbf{\Sigma},\nu). \tag{16}$$

By combining Equations (6) and (15), we obtain the conditional distribution of $\mathbf{\Omega}$ satisfying,

$$\begin{aligned} & p(\mathbf{\Omega}|\mathbf{\Sigma},\nu,\mathbf{W}) \\ & \propto \prod_{i=1}^{N} p(\mathbf{x}_i|\mathbf{X},\mathbf{w}_i,\mathbf{\Omega})p(\mathbf{\Omega}|\mathbf{\Sigma},\nu) \\ & \propto \prod_{i=1}^{N} |\mathbf{\Omega}|^{1/2}\exp\left(-\frac{1}{2}\mathrm{tr}\left((\mathbf{x}_i-\mathbf{X}\mathbf{w}_i)(\mathbf{x}_i-\mathbf{X}\mathbf{w}_i)^T\mathbf{\Omega}\right)\right) \\ & \times |\mathbf{\Omega}|^{(\nu-T-1)/2}\exp\left(-\frac{1}{2}\mathrm{tr}\left(\mathbf{\Sigma}^{-1}\mathbf{\Omega}\right)\right), \end{aligned} \tag{17}$$

where $\mathbf{w}_i$ denotes the ith column of $\mathbf{W}$ and the terms only relevant to the optimization of w_{ij} is omitted. From Equation (17), we see, Wishart prior is conjugate to the precision matrix $\mathbf{\Omega}$ of the multi-variate Gaussian distribution.

In conclusion, we reform Equation (17) and obtain the variational posterior $q(\mathbf{\Omega})$ *via* approximation inference (38) that follows:

$$\begin{aligned} & q(\mathbf{\Omega}) \propto |\mathbf{\Omega}|^{(N+\nu-T-1)/2}\exp \\ & \times\left(-\frac{1}{2}\mathrm{tr}\left(\left((\mathbf{X}-\mathbf{X}\mathbf{W})(\mathbf{X}-\mathbf{X}\mathbf{W})^T+\mathbf{\Sigma}^{-1}\right)\mathbf{\Omega}\right)\right). \end{aligned} \tag{18}$$

Then, the optimal solution of $\mathbf{\Omega}$ can be computed analytically as follows:

$$\mathbf{\Omega} = (N+\nu)\left((\mathbf{X}-\mathbf{XW})(\mathbf{X}-\mathbf{XW})^T + \mathbf{\Sigma}^{-1}\right)^{-1}, \quad (19)$$

which is equivalent to the optimal solution of the following optimization problem,

$$\min_{\mathbf{\Omega}} \quad \mathrm{tr}\left[\mathbf{\Omega}\left((\mathbf{X}-\mathbf{XW})(\mathbf{X}-\mathbf{XW})^T + \mathbf{\Sigma}^{-1}\right)\right] - (N+\nu)\ln|\mathbf{\Omega}|. \quad (20)$$

According to the form of the optimal solution, the structure of $\mathbf{\Omega}$ is partially governed by the structure of the prior matrix $\mathbf{\Sigma}$; thus it is natural to embed prior structure *via* $\mathbf{\Sigma}$ implicitly. Without the loss of generality, in this article, we consider $\mathbf{\Sigma} = \alpha^{-1}\mathbf{I}$ as an infinitely broad prior, so that the structure of $\mathbf{\Omega}$ is almost learned from the data entirely.

To sum up, the overall objective function joining all constraints in Equations (10) and (20), the SRAND with Wishart prior constraint can be summarized as follows:

$$\begin{aligned} \min_{\mathbf{W},\mathbf{\Omega}} \quad & \mathrm{tr}\left[\mathbf{\Omega}\left((\mathbf{X}-\mathbf{XW})(\mathbf{X}-\mathbf{XW})^T + \mathbf{\Sigma}^{-1}\right)\right] \\ & - (N+\nu)\ln|\mathbf{\Omega}| + \lambda_1\|\mathbf{W}\|_1 \\ \text{s.t.} \quad & w_{ii} = 0, \forall i = 1, \cdots, N. \end{aligned} \quad (21)$$

Consequently, the procedure of SRAND with Wishart prior is listed in Algorithm 2.

Input: Data matrix $\mathbf{X}$, parameters λ_1, ν, $\mathbf{\Sigma}$
Output: Constructed FBN $\mathbf{W}$
  Initialize $\mathbf{\Omega} = \mathbf{I}$
  while not converge **do**
    Update $\mathbf{W}$ by solving the problem in Equation (10)
    Update $\mathbf{\Omega}$ by Equation (19)
  end while

Algorithm 2. SRAND with Wishart prior.

3. Experiments and results

3.1. Experimental setting

3.1.1. FBN construction

In this section, we estimate FBNs on ABIDE database using different methods, including SR and the proposed SRAND with Wishart prior. In addition, we also conducted experiments using PC and PCAND (i.e., PC modified *via* adaptive noise depression) with Wishart prior. In this study, we fix the Wishart distribution with equal diagonal entries as the prior to gain an infinitely broad prior and to evaluate the influence of the same prior on different comparison methods. The comparison between different priors on the same method is followed in the subsequent section. In general, each SR-based method contains one or more hyperparameters for regularization, which may significantly influence the network structure and then the ultimate classification results (39). Therefore, for each regularized parameter, we build multiple FBNs on different parameter values in the candidate range $[0.05, 0.1, \cdots, 0.95, 1]$ and then search the optimal parameter value *via* a separate parameter selection procedure.

3.1.2. Feature selection and classification

After obtaining the estimated FBNs, we subsequently utilize them to identify participants with ASD from NCs. In our experiments, the upper triangular edge weights of the FBNs are selected as the input features since the FBN matrix is symmetric. In particular, 90 nodes (i.e., number of ROIs) produce $90 \times (90-1)/2 = 4,005$ dimensions of the feature. Compared with the small sample size of 92, it is still too high to ensure the good generalization ability of the classifier, which obviously affects the final classification accuracy. To address this problem, we adopt the simplest t-test with $p = 0.01$ as feature selection method. After selection of the most-relevant features, we use the most popular support vector machine (SVM) (linear kernel with default parameter $C = 1$) as our classifier for disease identification (40).

Furthermore, we test the involved FBN estimation methods by the leave-one-out cross-validation (LOOCV), in which experiments repeat for K times (i.e., the total number of subjects) for each subject as the testing set, while the rest subjects are used as the training set to select features and train classifier. Moreover, in order to determine the optimal value of the regularization parameter, an inner LOOCV is conducted on the training data *via* a grid search on the candidate range of parameters, which is based on the metric of classification accuracy.

The overall detailed pipeline of our experiments is shown in Figure 1.

3.2. Results

3.2.1. FBN visualization

In order to compare the results of different FBN estimation methods intuitively, we first take one subject from ABIDE dataset as an example to visualize the adjacency matrices of the FBNs estimated by the four comparison methods in Figure 1. Since the FBN is generally dense for PC-based methods, in order to show the changes more clearly, we enlarge part of the FBNs (indicated by the black box in Figure 2) and show the enlarged part (the second row in Figure 2).

It can be observed that the PC-based FBN is significantly different from those estimated by SR-based methods since

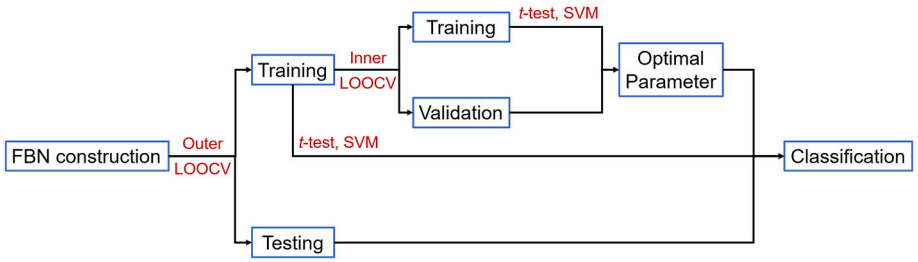


FIGURE 1
Experiment procedure based on the estimated FBNs.

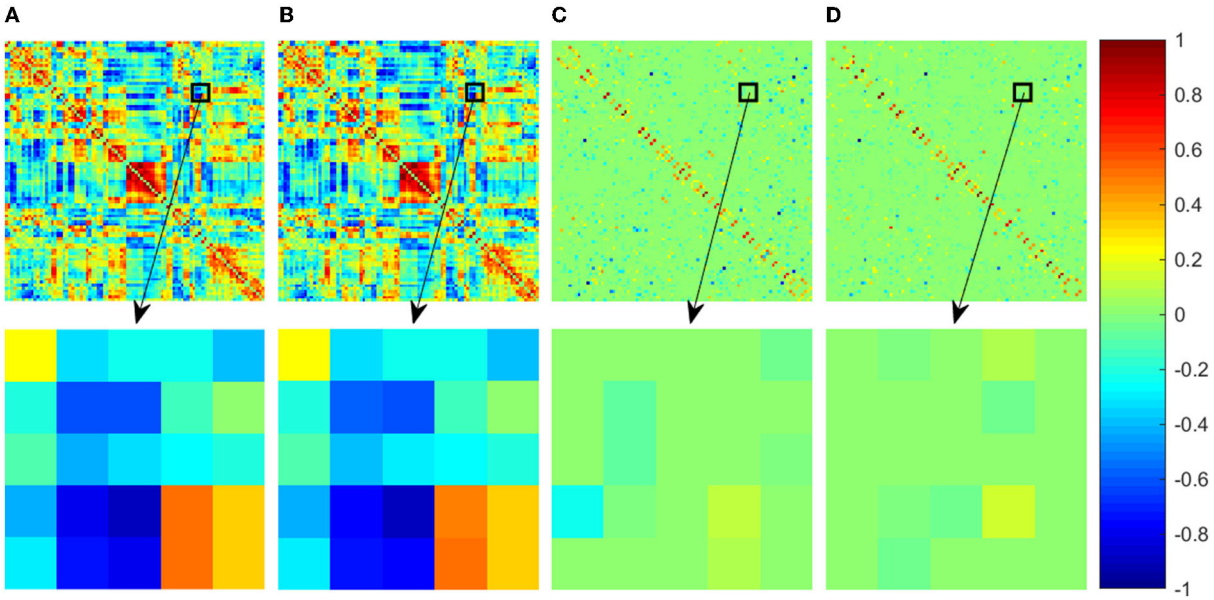


FIGURE 2
Adjacency matrices of the FBNs estimated by four comparison methods. The first row shows the whole adjacency matrix and the second row shows the enlarged part corresponding to the black box in the first row. For a convenient comparison between the visualized results, the elements in the adjacency matrices have been normalized into the interval [−1, 1]. **(A)** PC, **(B)** PCAND, **(C)** SR, and **(D)** SRAND.

they use different data-fitting terms to capture full correlation and partial correlation between ROIs, respectively. In contrast, the FBNs constructed based on the same data-fidelity term share similar topological structure, for example, PCAND is similar to that of PC and SRAND is similar to that of SR. Moreover, compared with PC and SR, the FBN estimated by PCAND and SRAND can further weaken or remove the noisy or weak connections and produce an even clearer topological structure. Specifically, although there is no sparsity constraint on estimating FBN by PCAND, it can weaken some of the dense connections constructed by the original PC (as shown by the brighter color blocks that indicate lower connection weights). As for SRAND, it can further remove the weak connections on the basis of the sparsity-contrained SR.

TABLE 2 **Classification performance corresponding to different FBN estimation methods on ABIDE dataset.**

Method	Accuracy	Sensitivity	Specificity
PC	0.6848	0.7333	0.6383
SR	0.5435	0.6000	0.4894
PCAND	**0.7174**	**0.7333**	0.7021
SRAND	0.6739	0.5556	**0.7872**

The bold values mean the best results corresponding to the performance metrics.

3.2.2. ASD identification

In this study, we adopt three quantitative metrics, including accuracy (ACC), sensitivity or true positive rate (SEN) and specificity or true negative rate (SPE) to evaluate

the classfication performance of different methods. The mathematical definitions of the first three measures are given as follows:

$$\begin{aligned} ACC &= \frac{TP + TN}{TP + TN + FP + FN} \\ SEN &= \frac{TP}{TP + FN} \\ SPE &= \frac{TN}{TN + FP}. \end{aligned} \tag{22}$$

where TP, TN, FP, and FN indicate true positive, true negative, false positive, and false negative, respectively. It should be noted that in this study, we treat the subjects with ASD as the positive class, while the NCs as the negative class.

The ASD vs. NC classification results on ABIDE dataset are reported in Table 2. As can be seen, the PC-based methods perform better than SR-based ones in our experiment. A similar problem has also been revealed in other studies, which can be blamed on the ill-posed computation of inverse covariance matrix with high feature dimension (16). Nevertheless, both modified methods enjoy better performance than the original ones, increasing classification accuracy by approximately 3.26 and 13.04%, respectively. The aforementioned results can be attributed to the adaptive noise depression module that can further filter out the noisy or weak connections in the estimated FBN and provide clearer connections highly related to neural disorders.

4. Discussion

4.1. Sensitivity to network model parameters

In general, the trade-off parameters in the FBN estimation methods play an important role to influence the ultimate classification performance (17, 41). To investigate the sensitivity of the proposed method to different parameter values, we repeat ASD classification experiments based on different parameter combinations and compute the classification accuracy *via* LOOCV on all of the subjects. In addition to SRAND with Wishart prior, in this experiment, we also include the version with sparsity prior as compared to evaluate the influence of different priors. For convenient, we denote the SRAND with Wishart prior and sparsity prior as SRAND-Wishart, and SRAND-L1, separately. In order to compare the sensitivity of the three methods, SR, SRAND-Wishart and SRAND-L1 simultaneously and conveniently, we fix the second parameter of SRAND-L1 as $\lambda_2 = 0.1$, so that the three methods have the same comparable parameter. Such operation is relatively fair for SR and SRAND-Wishart. Indeed, the fixed second parameter limits the freedom of the model and may keep SRAND-L1 away

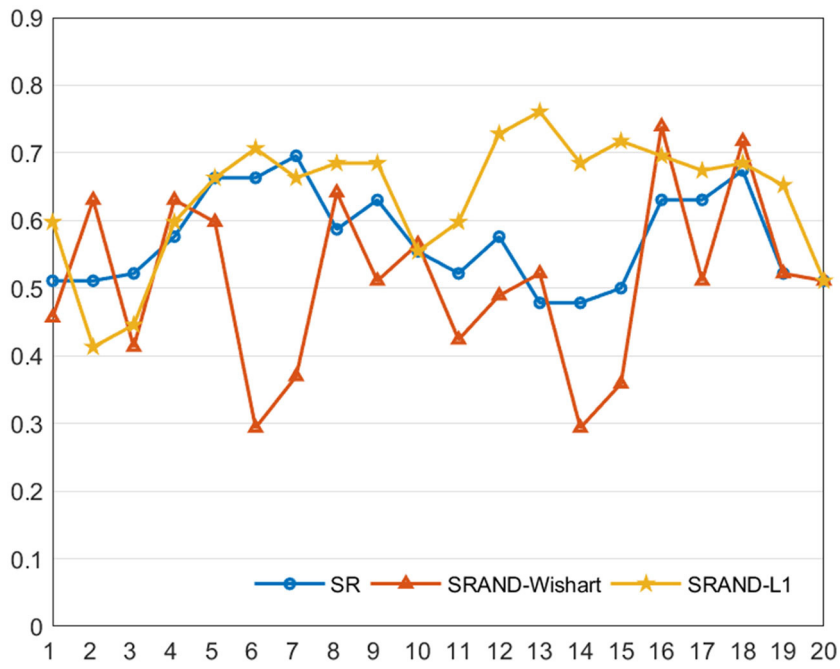


FIGURE 3
Classification accuracy of the FBN estimated by three comparison methods on 20 regularized parameters.

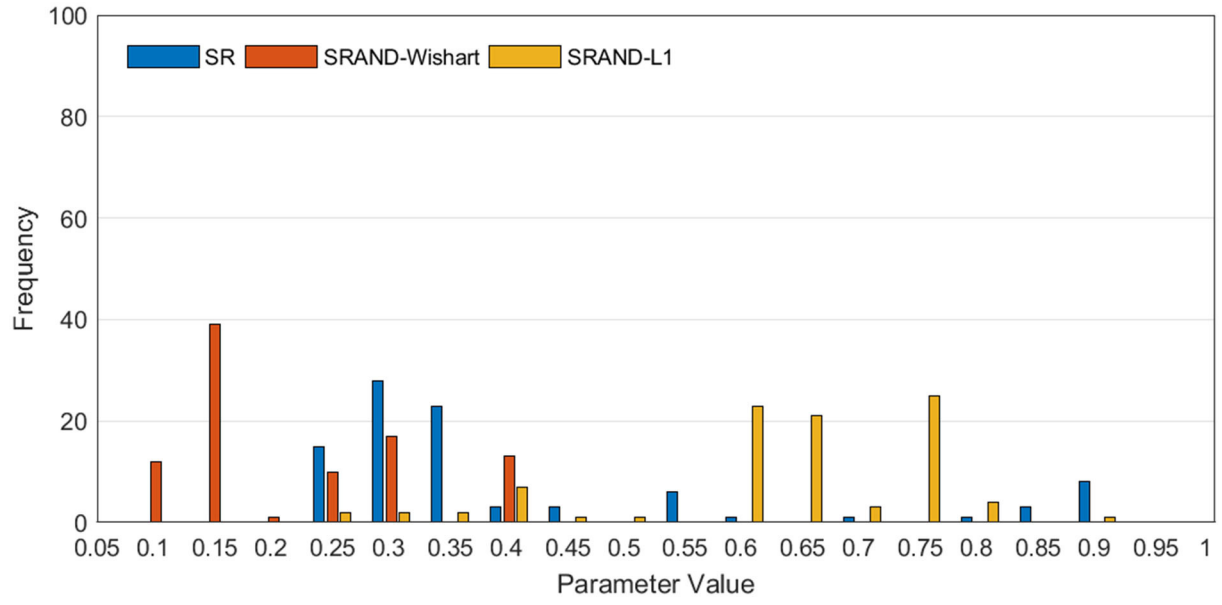


FIGURE 4
Frequency of the selected optimal values of parameter λ in the inner loops, where the horizontal axis represents the candidate parameter values, and the vertical axis represents the frequency of the parameter value being selected *via* parameter selection.

from the potential higher accuracy. The results are reported in Figure 3.

We note that most of the methods are sensitive to the parameters. Compared with the traditional SR-based methods, SRAND-Wishart is highly affected by varied parameters and even gains lower accuracy than the traditional ones. Such results are reasonable since Wishart prior actually imposes no structure constraint on the noise distribution, and the computation of $\mathbf{\Omega}$ is determined by the fMRI data. Thus, computing $\mathbf{\Omega}$ is generally ill-posed and sensitive to the parameter due to the limited sample size. In contrast, SRAND-L1 is less likely affected by parameters and thus achieves more stable results. In addition, SRAND-L1 achieves an improvement of the classification performance on most of the parameter values. Such results can be attributed to the well chosen prior that fits the noise structure. Therefore, we believe that appropriate adaptive noise depression module could eliminate bad effects on the FBN estimation and benefit the ultimate disease diagnosis.

In addition to the comparison between classification accuracies on different parameters, we also present the value distribution of the optimal parameters selected in the inner loops, as shown in Figure 4. We can find that the optimal parameters of all SR-based methods concentrate around some fixed λ, that is, $\lambda = 0.3$ for SR, $\lambda = 0.15$ for SRAND-Wishart, and $\lambda = 0.65$ for SRAND-L1. Such concentricity also implies the sensitivity of all SR-based methods that they prefer some fixed parameter value than the scattered ones.

4.2. Discriminative features

In this subsection, we use the estimated FBN of PCAND as an example to explore which features contribute the most to ASD identification in our experiments. Specifically, we apply *t*-test with a *p*-value of 0.001 to select discriminative features based on the FBN constructed by PCAND. The top 66 most discriminative connections are visualized based on the first 90 ROIs of AAL template (32) in Figure 5, where the thickness of the arc shows the discriminative power. From Figure 5, we can find that the most discriminative features focus on the brain regions, including frontal, parahippocampus and pallidum, and so on. The frontal lobe is key to communication and cognitive function and is known to be implicated in ASD (42). Parahippocampus has been shown to have hypoactivation in scene recognition, in line with the notion of peaks and valleys of neural recruitment in individuals with ASD (43). Pallidum enlargement has been found in ASD compared with NC and may be a possible related factor in stereotypic behavior and social bonding (44). Similar findings have also been presented in previous studies (45–47).

5. Conclusion

Sparse representation is one of the most commonly used schemes for estimating FBNs due to its simplicity and relatively clearer network connections. Nevertheless, the SR scheme is

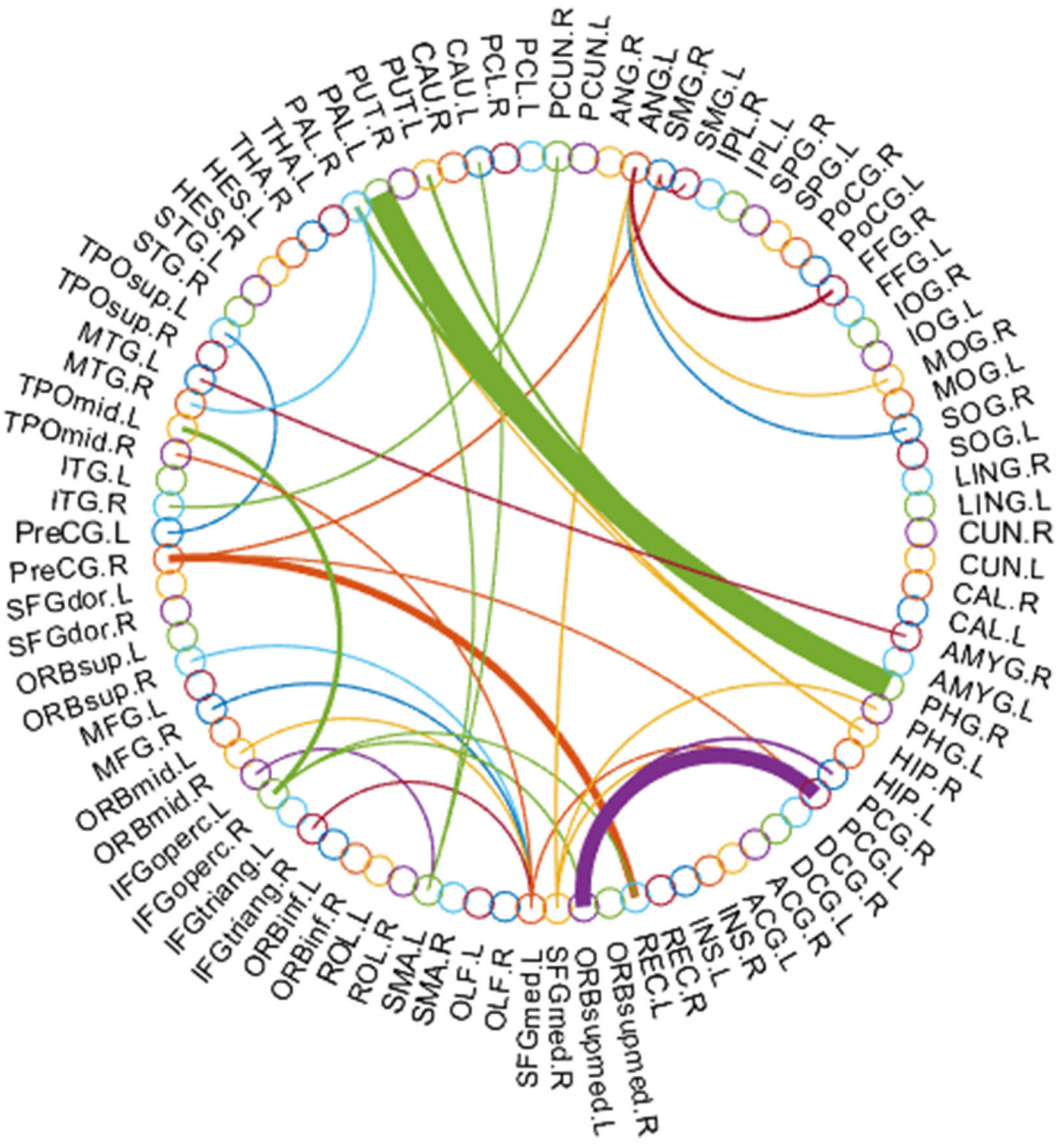


FIGURE 5
Most discriminative connections between ASD and NC for the first 90 ROIs of the AAL template.

still trapped in noisy or weak connections due to the noise introduced *via* data acquisition. In this study, we embed a noise module based on which the prior of noise pattern can be naturally incorporated in the form of regularizers. Such module has been illustrated to be plug and play that is capable of being embedded into different methods and adjusted according to different noise priors. To evaluate the effectiveness of the proposed scheme, we conduct experiments on the ABIDE database to identify subjects with ASD from normal controls. The experimental results demonstrate that the proposed method can achieve better performance than the baseline method.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

Ethical review and approval/written informed consent to participate was not required in accordance with local legislation and institutional guidelines.

Author contributions

DM and LP designed the study and drafted the manuscript. LP collected the MRI data. LP and XG analyzed and interpreted the results of the data. XG and DM revised the manuscript. All authors approved the final manuscript.

Funding

This study was partially supported by the Shanghai Municipal Commission of Health and Family Planning Science and Research Subjects (202140464) and the Scientific Research Subjects of Shanghai Universal Medical Imaging Technology Limited Company (UV2022M06, UV2021Z01, and UV2020Z02).

Conflict of interest

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.

Publisher's note

References

1. Wee CY, Yap PT, Shen D. Diagnosis of autism spectrum disorders using temporally distinct resting-state functional connectivity networks. *CNS Neurosci Therapeut*. (2016) 22:212–9. doi: 10.1111/cns.12499

2. Dadi K, Rahim M, Abraham A, Chyzhyk D, Milham M, Thirion B, et al. Benchmarking functional connectome-based predictive models for resting-state fMRI. *Neuroimage*. (2019) 192:115–34. doi: 10.1016/j.neuroimage.2019.02.062

3. Maenner MJ, Shaw KA, Bakian AV, Bilder DA, Durkin MS, Esler A, et al. Prevalence and characteristics of autism spectrum disorder among children aged 8 years autism and developmental disabilities monitoring network, 11 sites, United States, 2018. *MMWR Surveill Summ*. (2021) 70:1. doi: 10.15585/mmwr.ss7011a1

4. Guthrie W, Swineford LB, Nottke C, Wetherby AM. Early diagnosis of autism spectrum disorder: stability and change in clinical diagnosis and symptom presentation. *J Child Psychol Psychiatry*. (2013) 54:582–90. doi: 10.1111/jcpp.12008

5. O' Roak BJ, Vives L, Fu W, Egertson JD, Stanaway IB, Phelps IG, et al. Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science*. (2012) 338:1619–22. doi: 10.1126/science.1227764

6. Zhang Y, Li N, Li C, Zhang Z, Teng H, Wang Y, et al. Genetic evidence of gender difference in autism spectrum disorder supports the female-protective effect. *Transl Psychiatry*. (2020) 10:1–10. doi: 10.1038/s41398-020-0699-8

7. Wass S. Distortions and disconnections: disrupted brain connectivity in autism. *Brain Cogn*. (2011) 75:18–28. doi: 10.1016/j.bandc.2010.10.005

8. Drysdale AT, Grosenick L, Downar J, Dunlop K, Mansouri F, Meng Y, et al. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat Med*. (2017) 23:28–38. doi: 10.1038/nm.4246

9. Abraham A, Milham MP, Di Martino A, Craddock RC, Samaras D, Thirion B, et al. Deriving reproducible biomarkers from multi-site resting-state data: an Autism-based example. *Neuroimage*. (2017) 147:736–45. doi: 10.1016/j.neuroimage.2016.10.045

10. Fox MD, Raichle ME. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nat Rev Neurosci*. (2007) 8:700–11. doi: 10.1038/nrn2201

11. Starck T, Nikkinen J, Rahko J, Remes J, Hurtig T, Haapsamo H, et al. Resting state fMRI reveals a default mode dissociation between retrosplenial and medial prefrontal subnetworks in ASD despite motion scrubbing. *Front Hum Neurosci*. (2013) 7:802. doi: 10.3389/fnhum.2013.00802

12. Eslami T, Mirjalili V, Fong A, Laird AR, Saeed F. ASD-DiagNet: a hybrid learning approach for detection of autism spectrum disorder using fMRI data. *Front Neuroinform*. (2019) 13:70. doi: 10.3389/fninf.2019.00070

13. Smith SM, Miller KL, Salimi-Khorshidi G, Webster M, Beckmann CF, Nichols TE, et al. Network modelling methods for FMRI. *Neuroimage*. (2011) 54:875–91. doi: 10.1016/j.neuroimage.2010.08.063

14. Maximo JO, Cadena EJ, Kana RK. The implications of brain connectivity in the neuropsychology of autism. *Neuropsychol Rev*. (2014) 24:16–31. doi: 10.1007/s11065-014-9250-0

15. Hull JV, Dokovna LB, Jacokes ZJ, Torgerson CM, Irimia A, Van Horn JD. Resting-state functional connectivity in autism spectrum disorders: a review. *Front Psychiatry*. (2017) 7:205. doi: 10.3389/fpsyt.2016.00205

16. Li W, Wang Z, Zhang L, Qiao L, Shen D. Remodeling Pearson's correlation for functional brain network estimation and autism spectrum disorder identification. *Front Neuroinform*. (2017) 11:55. doi: 10.3389/fninf.2017.00055

17. Xue Y, Zhang L, Qiao L, Shen D. Estimating sparse functional brain networks with spatial constraints for MCI identification. *PLoS One*. (2020) 15:e0235039. doi: 10.1371/journal.pone.0235039

18. Xu X, Wang T, Li W, Li H, Xu B, Zhang M, et al. Morphological, structural, and functional networks highlight the role of the cortical-subcortical circuit in individuals with subjective cognitive decline. *Front Aging Neurosci*. (2021) 13:394. doi: 10.3389/fnagi.2021.688113

19. Khatri U, Kwon GR. Alzheimer's disease diagnosis and biomarker analysis using resting-state functional MRI functional brain network with multi-measures features and hippocampal subfield and amygdala volume of structural MRI. *Front Aging Neurosci*. (2022) 14:818871. doi: 10.3389/fnagi.2022.818871

20. Li WK, Chen YC, Xu XW, Wang X, Gao X. Human-Guided functional connectivity network estimation for chronic tinnitus identification: a modularity view. *IEEE J Biomed Health Inform*. (2022) 26:4849–58. doi: 10.1109/JBHI.2022.3190277

21. Li W, Xu X, Wang Z, Peng L, Wang P, Gao X. Multiple connection pattern combination from single-mode data for mild cognitive impairment identification. *Front Cell Dev Biol*. (2021) 9:782727. doi: 10.3389/fcell.2021.782727

22. Lee H, Lee DS, Kang H, Kim BN, Chung MK. Sparse brain network recovery under compressed sensing. *IEEE Trans Med Imaging*. (2011) 30:1154–1165. doi: 10.1109/TMI.2011.2140380

23. Huang S, Li J, Sun L, Ye J, Fleisher A, Wu T, et al. Learning brain connectivity of Alzheimer's disease by sparse inverse covariance estimation. *Neuroimage*. (2010) 50:935–49. doi: 10.1016/j.neuroimage.2009.12.120

24. Poldrack RA, Mumford JA, Nichols TE. *Handbook of functional MRI data analysis*. New York, NY: Cambridge University Press.

25. Li W, Qiao L, Zhang L, Wang Z, Shen D. Functional brain network estimation with time series self-scrubbing. *IEEE J Biomed Health Inform*. (2019) 23:2494–504. doi: 10.1109/JBHI.2019.2893880

26. Di Martino A, Yan CG, Li Q, Denio E, Castellanos FX, Alaerts K, et al. The autism brain imaging data exchange: towards a large-scale evaluation of the intrinsic brain architecture in autism. *Mol Psychiatry*. (2014) 19:659–67. doi: 10.1038/mp.2013.78

27. First MB, France A, Pincus HA. DSM-IV-TR guidebook. *American Psychiatric Publishing, Inc*.;. (2004).

28. Zhao F, Zhang X, Thung KH, Mao N, Lee SW, Shen D. Constructing multi-view high-order functional connectivity networks for diagnosis of autism spectrum disorder. *IEEE Trans Biomed Eng*. (2021) 69:1237–50. doi: 10.1109/TBME.2021.3122813

29. Satterthwaite TD, Elliott MA, Gerraty RT, Ruparel K, Loughead J, Calkins ME, et al. An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage*. (2013) 64:240–56. doi: 10.1016/j.neuroimage.2012.08.052

30. Yan CG, Cheung B, Kelly C, Colcombe S, Craddock RC, Di Martino A, et al. A comprehensive assessment of regional variation in the impact of head micromovements on functional connectomics. *Neuroimage*. (2013) 76:183–201. doi: 10.1016/j.neuroimage.2013.03.004

31. Xie Q, Zhang X, Rekik I, Chen X, Mao N, Shen D, et al. Constructing high-order functional connectivity network based on central moment features for diagnosis of autism spectrum disorder. *PeerJ*. (2021) 9:e11692. doi: 10.7717/peerj.11692

32. 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) 15:273–89. doi: 10.1006/nimg.2001.0978

33. Smith SM, Vidaurre D, Beckmann CF, Glasser MF, Jenkinson M, Miller KL, et al. Functional connectomics from resting-state fMRI. *Trends Cogn Sci*. (2013) 17:666–82. doi: 10.1016/j.tics.2013.09.016

34. Fornito A, Zalesky A, Bullmore E. *Fundamentals of Brain Network Analysis*. Cambridge, MA: Academic Press (2016).

35. Su H, Zhang L, Qiao L, Liu M. Estimating high-order brain functional networks by correlation-preserving embedding. *Med Biol Eng Comput*. (2022) 60:2813–23. doi: 10.1007/s11517-022-02628-7

36. Meinshausen N, Bühlmann P. High-dimensional graphs and variable selection with the lasso. *Ann Stat*. (2006) 34:1436–62. doi: 10.1214/009053606000000281

37. Friedman J, Hastie T, Tibshirani R. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics*. (2008) 9:432–41. doi: 10.1093/biostatistics/kxm045

38. Blei DM, Kucukelbir A, McAuliffe JD. Variational inference: a review for statisticians. *J Am Stat Assoc*. (2017) 112:859–77. doi: 10.1080/01621459.2017.1285773

39. Wee CY, Yap PT, Zhang D, Wang L, Shen D. Group-constrained sparse fMRI connectivity modeling for mild cognitive impairment identification. *Brain Struct Funct*. (2014) 219:641–56. doi: 10.1007/s00429-013-0524-8

40. Chang CC, Lin CJ. LIBSVM: a library for support vector machines. *ACM Trans Intell Syst Technol*. (2011) 2:1–27. doi: 10.1145/1961189.1961199

41. Jiang X, Zhang L, Qiao L, Shen D. Estimating functional connectivity networks via low-rank tensor approximation with applications to MCI identification. *IEEE Trans Biomed Eng*. (2019) 67:1912–20. doi: 10.1109/TBME.2019.2950712

42. Sun Y, Yao X, March ME, Meng X, Li J, Wei Z, et al. Target genes of autism risk loci in brain frontal cortex. *Front Genet*. (2019) 10:707. doi: 10.3389/fgene.2019.00707

43. Mouga S, Duarte IC, Café C, Sousa D, Duque F, Oliveira G, et al. Parahippocampal deactivation and hyperactivation of central executive, saliency and social cognition networks in autism spectrum disorder. *J Neurodev Disord*. (2022) 14:1–12. doi: 10.1186/s11689-022-09417-1

44. Turner AH, Greenspan KS, van Erp TG. Pallidum and lateral ventricle volume enlargement in autism spectrum disorder. *Psychiatry Res Neuroimaging*. (2016) 252:40–5. doi: 10.1016/j.pscychresns.2016.04.003

45. Monk CS, Peltier SJ, Wiggins JL, Weng SJ, Carrasco M, Risi S, et al. Abnormalities of intrinsic functional connectivity in autism spectrum disorders. *Neuroimage*. (2009) 47:764–72. doi: 10.1016/j.neuroimage.2009.04.069

46. Cheng Y, Chou KH, Chen IY, Fan YT, Decety J, Lin CP. Atypical development of white matter microstructure in adolescents with autism spectrum disorders. *Neuroimage*. (2010) 50:873–82. doi: 10.1016/j.neuroimage.2010.01.011

47. Van Rooij D, Anagnostou E, Arango C, Auzias G, Behrmann M, Busatto GF, et al. Cortical and subcortical brain morphometry differences between patients with autism spectrum disorder and healthy individuals across the lifespan: results from the ENIGMA ASD Working Group. *Am J Psychiatry*. (2018) 175:359–69. doi: 10.1176/appi.ajp.2017.17010100

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 25 January 2023
DOI 10.3389/fpsyt.2022.1090369

OPEN ACCESS

EDITED BY
Weikai Li,
Chongqing Jiaotong University, China

REVIEWED BY
Xize Jia,
Hangzhou Normal University, China
Lingxiao Wang,
Hangzhou Normal University, China

*CORRESPONDENCE
Yan-Ping Bao
✉ baoyp@bjmu.edu.cn
Su-Hua Chang
✉ changsh@bjmu.edu.cn
Lin Lu
✉ linlu@bjmu.edu.cn
Le Shi
✉ leshi@bjmu.edu.cn

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Computational Psychiatry, a section of the journal Frontiers in Psychiatry

RECEIVED 05 November 2022
ACCEPTED 14 December 2022
PUBLISHED 25 January 2023

CITATION
Yuan Z, Lin X, Li P, Gao Y-J, Yuan K, Yan W, Zhang Y-X, Liu L, Zhu X-M, Zhang Y-J, Bao Y-P, Chang S-H, Lu L and Shi L (2023) The neural correlation of emotion recognition ability and depressive symptoms–evidence from the HCP database.
Front. Psychiatry 13:1090369.
doi: 10.3389/fpsyt.2022.1090369

The neural correlation of emotion recognition ability and depressive symptoms–evidence from the HCP database

Ze Yuan[1,2†], Xiao Lin[2†], Peng Li[2], Yu-Jun Gao[3], Kai Yuan[2], Wei Yan[2], Yu-Xin Zhang[2,4], Lin Liu[2,4], Xi-Mei Zhu[2], Yi-Jing Zhang[2], Yan-Ping Bao[5,6*], Su-Hua Chang[2*], Lin Lu[2,4,5*] and Le Shi[2*]

[1]Savaid Medical School, University of Chinese Academy of Sciences, Beijing, China, [2]Chinese Academy of Medical Sciences Research Unit (No. 2018RU006), NHC Key Laboratory of Mental Health (Peking University), National Clinical Research Center for Mental Disorders (Peking University Sixth Hospital), Peking University Sixth Hospital, Peking University Institute of Mental Health, Peking University, Beijing, China, [3]Department of Psychiatry, Renmin Hospital of Wuhan University, Wuhan, China, [4]Peking-Tsinghua Centre for Life Sciences and PKU-IDG/McGovern Institute for Brain Research, Peking University, Beijing, China, [5]Beijing Key Laboratory of Drug Dependence, National Institute on Drug Dependence, Peking University, Beijing, China, [6]School of Public Health, Peking University, Beijing, China

Introduction: Negative bias of emotional face is the core feature of depression, but its underlying neurobiological mechanism is still unclear. The neuroimaging findings of negative emotional recognition and depressive symptoms are inconsistent.

Methods: The neural association between depressive symptoms and negative emotional bias were analyzed by measuring the associations between resting state functional connectivity (FC), brain structures, negative emotional bias, and depressive problems. Then, we performed a mediation analysis to assess the potential overlapping neuroimaging mechanisms.

Results: We found a negative correlation between depressive symptoms and emotional recognition. Secondly, the structure and function of the inferior and lateral orbitofrontal gyrus are related to depressive symptoms and emotional recognition. Thirdly, the thickness of the inferior orbitofrontal cortex and the FC between the inferior orbitofrontal gyrus and fusiform gyrus, precuneate and cingulate gyrus mediated and even predicted the interaction between emotion recognition and depressive symptoms. Finally, in response to a negative stimulus, the activation of the frontal pole and precuneus lobe associated with the inferior orbitofrontal gyrus was higher in participants with depressive symptoms.

Conclusion: The core brain regions centered on the inferior orbitofrontal cortex such as middle temporal gyrus, precuneus lobe, frontal pole, insula and cingulate gyrus are the potential neuroimaging basis for the interaction between depressive symptoms and emotional recognition.

KEYWORDS

depressive symptoms, emotional recognition ability, magnetic resonance imaging, functional connectivity, machine learning

1. Introduction

Depression is one of the most common mental illnesses and the second leading cause of disability and incapacity worldwide (1). Although neurobiological research has been carried out for more than 60 years, its pathophysiological understanding is still limited (2–4). Negative emotional recognition bias and reduced pleasure experience are two core features and symptoms of depression. Patients with depression paid more attention to negative emotions and tended to classify fuzzy/neutral faces as negative. Patients with depression showed a reduced tendency to recognize positive emotions, such as recognizing joyful faces as neutral faces, resulting in a lack of happiness and identity in their previous motivational activities (5) and motivation (6). Therefore, it is critical to study the relationship and overlapped neural mechanisms between depressive symptoms and emotional recognition bias to prevent and treat depression.

Many studies have shown that in major depression disease (MDD) patients, the dominant facial emotion recognition is destroyed (7) and it is related to the appearance of depression symptoms which are the basis and process of depression (8). Individuals with depression have also been found to have recognition bias, but the evidence of this negative bias was inconsistent (9). Previous small sample functional neuroimaging studies showed that patients with depression have abnormal emotional recognition in emotion-related brain regions (10–15), consisting of the amygdala, medial forehead, insular, anterior cingulate gyrus, nucleus accumbens, and orbitofrontal gyrus (16–18). Neuroimaging studies found that the activation of the prefrontal cortex and anterior cingulate gyrus during emotional recognition was a key feature of major depression disease (19). Resting-state fMRI also found that depressed patients had enhanced FC (functional connectivity) between the parahippocampal gyrus, temporal pole and infratemporal gyrus, and the medial orbitofrontal cortex (20, 21). Results of Granger causality analysis suggest that these areas have a solid driving effect on the medial orbitofrontal cortex (reward-related regions), so it may help to clarify that happiness is reduced in depression (22). The increased FC strength between the temporal cortex and precuneus in patients with depression may be related to the representation of self-consciousness (23), which makes patients with depression a more negative biased (22). For sMRI evidence, in a recent large-scale study of adolescent brain cognitive development datasets, depression was associated with increased volume of the orbital frontal cortex, temporal cortex, and medial frontal cortex (24). There is also some evidence of structural changes in the lateral orbitofrontal cortex in patients with depression (24, 25). Gray matter volume has increased in the posterolateral orbitofrontal cortex and the anterior cingulate cortex (26). Meta-analysis showed that the volume of frontal lobe regions, especially the anterior cingulate gyrus and orbitofrontal cortex, increased in patients with depression (27). It can be seen from the above evidence that emotion recognition bias in depressed individuals is related to orbitofrontal cortex but whether emotion recognition bias is related to depressed symptoms and the impact process is still unclear. Therefore, this study aims to clarify the relationship between depressive symptoms and emotional recognition bias and the role of the orbitofrontal cortex in the interaction between them.

Through the large sample of the international open database HCP (Human Connectome Project), this study explores the relationship between depressive symptoms, emotional recognition, the underlying overlapped neural mechanism, and the common intermediary factors. The main contents of this paper are as follows: (1) using the demographic data, psychological scale, and emotional recognition task test, we analyzed the demographic and behavioral data to explore the relationship between emotion recognition ability and depressive symptoms and whether there are common intermediary factors; (2) the brain regions and FCs related to emotional recognition ability and depressive symptoms are explored; (3) we explored the intermediary relationship between emotion recognition and depressive symptoms, and verified the hypothesis that the cortex near the orbitofrontal gyrus is more related to emotional recognition under depressive symptoms.

2. Materials and methods

2.1. Participants

To explore the relationship between emotional recognition bias and depressive symptoms, we analyzed phenotype data and imaging data from the Human Connectome Project (HCP) database (March 2017 public data release) from the Washington University-University of Minnesota (WU-Minn HCP) Consortium. We selected the participants with general demographic data, completed the emotional recognition task, ASR (Achenbach Adult Self-Report) Scale, and completed the structural state, resting state and task state MRI scanning for this analysis.

2.2. Measurement

The behavioral data analyzed in this study included the ASR depression scale (level 1) and the NIH toolbox emotional task test. ASR scale is a self-assessment tool compiled by Achenbach et al. in 1997 (28, 29). This psychological scale based on DSM-IV (Diagnostic and Statistical Manual of Mental Disorders-IV) was used to evaluate the depressive symptoms of participants in HCP. Depression syndrome scale score, age, and sex corrected depression syndrome scale score, depression scale DSM score, age, and sex corrected depression scale DSM score is also measured to evaluate the level of depressive symptoms. The higher the score, the higher the level of depressive symptoms of the participants. ASR scale is a continuous variable, so it can be used to conduct correlation analysis with the performance of emotion recognition tasks and conduct analysis on the mediating effect of FCs. The emotional NIH toolbox contains only self-reported measurements of emotional recognition functions. Therefore, we used the Pennsylvania Emotion Recognition Test to obtain behavioral measurements of emotional recognition (30–32). In this study, we used the accuracy of the emotion recognition task as the standard of emotion recognition, and we used the completion time of the emotion recognition task to assist in the prediction of depressive symptoms and intermediary effect calculation.

Besides, this study also extracted general demographic data, including gender, age, race, annual family income, years of education, left and right-handedness, body mass index (BMI), blood pressure, marijuana addiction diagnosis history, alcohol addiction diagnosis history, tobacco addiction diagnosis history and addiction drug urine test results. In addition, to explore the effect of psychological factors on the relationship between emotion recognition and depressive symptoms, we included several psychological factors, including personality factors, cognitive factors, stress level, indicators of working memory tasks and social tasks, etc. Some participants had experienced one of the nine depression diagnostic items in the DSM-IV and were counted as participants with depressive symptoms.

All participants performed MRI image acquisition by the HCP item group on a custom Siemens 3T "Connectome Skyra" machine at Washington University in St. Louis, using a standard 32-channel Siemens receiver head coil and a transmission coil customized for a smaller space.

2.3. Statistical analysis of data

2.3.1. Statistical analysis of demographic and behavioral data

According to the data in HCP, the general demographic data, the indicators of depressive symptoms, emotional recognition and MRI data were used for analysis in this study. After excluding the missing values, 995 participants were included, and 260 of 995 people had experienced depressive symptoms. With or without depressive symptoms are more suitable for predictive analysis because they are binary values. The data were analyzed using SPSS 20.0 statistical software, including correlation and intermediary analysis. The covariates included sex, race, annual family income, diagnosis history of marijuana addiction, diagnosis history of alcohol addiction, diagnosis history of tobacco addiction, urine test results of addiction to drugs, blood pressure, and anxiety syndrome score of ASR-DSM anxiety scale. Pearson partial correlation analysis was used to explore the relationship between indicators of depressive symptoms and emotional recognition ability and the psychological factors contained in the HCP database. When $p < 0.05$, it is considered that there is a statistically significant correlation. Then, we used simple intermediary analysis to analyze whether the psychological factors (mediating variable M) mediate the relationship between depressive symptoms (independent variable X) and emotional recognition (dependent variable Y) or between emotion recognition (independent variable X) and depressive symptoms (dependent variable Y). In this study, through the PROCESS plug-in of R (33), model 4 of the Bootstrap method is used for simple intermediary effect tests.

2.3.2. Statistical analysis of structural MRI data

Based on the sMRI data obtained from HCP, which has been preprocessed by using the HCP pipeline, a linear regression model was constructed to calculate whether each brain region structure was related to behavioral data (depressive symptoms or emotional recognition ability), including the cortical surface area and thickness of 68 brain regions and the volume of 44 subcortical structures (for detailed information see **Supplementary material**). Based on the covariates of correlation analysis, total gray matter volume and white matter volume were added as covariables, and multiple comparison correction was carried out using the False Discovery Rate (FDR)

method. After FDR correction, $p < 0.05$ is suggested that the brain structure is related to behavioral performance.

2.3.3. Statistical analysis of resting state fMRI and task-related fMRI

Based on fMRI data after HCP preprocessing, this study used MATLAB R2016a software and python 3.6 to deal with resting fMRI and further analyses the ALFF (amplitude of low-frequency fluctuation), fALFF (fractional amplitude of low-frequency fluctuation), ReHo (Regional Homogeneity), VMHC (voxel-mirrored homotopic connectivity), and FC. BrainNetViewer toolkit was used to visualize the results. CONN (FC toolbox) was used to analyze task-related fMRI data. The primary analysis steps include the following:

(1) Data collation: This step only includes participants with two resting fMRI data, excludes those who are missing, and then averages the two resting fMRI data of each participant.

(2) The brain region template selection: The international brain region template Shen template (34) was selected in this study.

(3) ReHo lies in describing the similarity of the time series of a given voxel to the time series of its nearest neighbors, so ReHo is used to describe consistency within a region. ALFF and fALFF are used to reveal BOLD (blood oxygen level dependence) signal strength of spontaneous regional activity and reflect the spontaneous activity of the nerve. VMHC approach aims to reflect the difference in FC between brain hemispheres. Dpabi (Data Processing and Analysis of Brain Imaging) was used to calculate the ALFF, fALFF, ReHo, and VMHC values with and without depressive symptoms and for participants with low (wrong more than once on all emotion recognition tests) and high (all right or wrong once on all emotion recognition tests) emotional recognition ability (without depressive symptoms). Two samples *T*-test corrected by age and sex was used for the statistical test, and the whole brain results were corrected by TFCE (Threshold Free Cluster Enhancement) of permutation test with the threshold of 0.05 (35, 36).

(4) Functional connectivity matrix construction: the resting fMRI data was divided into 250 brain regions according to the Shen template. Nodal signals were created by averaging the regional blood oxygen level-dependent signals of all voxels within each region. Pearson cross-correlations between all pairwise combinations of region signals were calculated to construct the FC correlation coefficient matrix for each participant, followed by Fisher' Z transformation to make the matrix approximate Gaussian distribution, and each participant constructs an FC matrix.

(5) Mediation analysis was used to examine whether FC strengths mediated the interaction between depressive symptoms and emotion recognition ability after controlling general demographic data.

(6) CONN was used to analyze task-related fMRI data with and without depressive symptoms and for participants with low and high emotional recognition ability.

3. Results

3.1. Demographic data of the participants

After deleting individuals with missing variables, 995 participants with behavior and brain structure data from HCP were used for the analysis. The participants were young adults, 22–37 years old (mean = 28.76 years, standard deviation = 3.72 years), 464 were male and 75% were white (**Table 1**). Of 995 participants, 230 had experienced one of the 9 depression diagnostic items in the DSM- IV, and these participants were counted as participants with depressive symptoms.

3.2. Emotion recognition ability was significantly related to depressive symptoms

According to the results of Pearson correlation analysis, it was suggested that there is an extensive pairwise correlation between the accuracy of emotional task completion and the indexes of depressive symptoms (**Table 2**). Among them, emotion recognition ability was negatively correlated with depressive symptoms (**Table 2**). These indexes of emotion recognition tasks contained the accuracy percentage of the completed emotion recognition task, the average accuracy percentage of the face recognition task in the emotion recognition task, and the average accuracy percentage of the face shape recognition task in the emotion recognition task. The indexes of depression scale score contained depression syndrome scale score, age and sex corrected depression syndrome scale score, depression DSM score, age and sex corrected depression DSM score. There were positive correlations between the completion time of the emotion recognition tasks and the indexes of depression scale scores.

3.3. MRI results

3.3.1. Overlapped brain regions related to emotion recognition ability and depressive symptoms

The emotional recognition ability was positively correlated with the cortical thickness of five brain regions, including the right lingual gyrus, the left lingual gyrus, and the operculum part of the inferior frontal gyrus. At the same time, the

cortical thickness of three brain regions of the anterior coracoid cingulate cortex, caudate anterior cingulate gyrus, and inferior frontal gyrus orbital part negatively correlated with emotional recognition ability. Among them, the evaluation index of emotion recognition ability was the average reaction time of each test during the completion of the emotion recognition task. The higher the value, the lower the emotion recognition ability. The depressive symptoms were negatively correlated with the cortical thickness of three brain regions, including inferior frontal gyrus orbital part, parahippocampal gyrus and orbitofrontal lobe. Therefore, the cortical thickness of the inferior frontal gyrus orbital part was correlated with depressive symptoms and emotional recognition ability (Table 3).

TABLE 1 Demographic and phenotype data characteristics (n = 995).

Variables	Mean ± SD/(%)
Gender	
Male	464 (46.63%)
Female	531 (53.37%)
Age (years)	28.76 ± 3.72
Race	
Caucasian	750 (75.38%)
Other races	245 (24.62%)
Annual household income (dollars)	
<30000	274 (27.54%)
30000–74999	424 (42.61%)
≥75000	297 (29.85%)
Years of education (years)	14.96 ± 1.76
Left and right handedness	
Left-handed	68 (6.84%)
Two-handed balance	46 (4.62%)
Right-handed	881 (88.54%)
BMI	26.60 ± 5.23
Urine test positive for addictive drugs	118 (11.85%)
DSM-4 diagnostic history of substance addiction	
Marijuana addiction	80 (8.04%)
Alcohol addiction	47 (4.72%)
Tobacco addiction	162 (16.28%)
ASR score of depression syndrome scale	5.93 ± 4.16
Age and sex correction score of ASR depression Syndrome scale	54.03 ± 8.67
ASR depression problems DSM-4 scale score	4.25 ± 2.72
Age and sex correction score of ASR depression Problems DSM-4 scale score	54.05 ± 8.48
ASR anxiety problems DSM-4 scale score	3.94 ± 2.15
Age and sex correction score of ASR anxiety Problems DSM-4 scale score	53.38 ± 8.91
The accuracy of the emotion recognition task	87.52 ± 9.87
The completion time of the emotion recognition task	756.49 ± 91.26

Some of the results are expressed in the form of mean ± SD, some in the form of numerical value and percentage, BMI, body mass index; DSM-IV, fourth edition of Diagnostic and Statistical Manual of Mental Disorders.

Furthermore, we used simple intermediary analysis to explore whether brain structure mediates the effect of depressive symptoms on emotional recognition ability. The mediating effect and proportion were tested by juxtaposing multiple intermediary analyses when controlling other mediating factors. Table 2 showed a mediating effect between the cortical thickness of the inferior frontal gyrus orbital part, four indexes of depressive symptoms, and two indexes of emotional recognition ability. It could be concluded that depressive symptoms may reduce the emotional recognition ability by affecting the thickness of the inferior frontal gyrus orbital part cortex. In contrast, the decreased emotional recognition ability might also by affecting the thickness of the inferior frontal gyrus orbital part cortex worsen the experience in life and aggravate the depressive symptoms (37–39) (Figure 1).

3.3.2. Functional brain characteristics related to emotional recognition ability and depressive symptoms

The strength of FC that plays an intermediary role in the interaction between emotion recognition ability and depressive symptoms is the strength that plays an intermediary role in the process of emotional recognition affecting depressive symptoms and in the process of depressive symptoms affecting emotional recognition ability. By analyzing the FC strength playing an intermediary role in the interaction between emotion recognition ability and depressive symptoms, we obtained 43 shared pairs of FCs. These FCs are the overlap FCs which are significantly related to depressive symptoms and emotional processing ability. Then we further explored whether the identified FCs significantly mediated this association between depressive symptoms and emotional processing ability. The specific FC is shown in Figure 2 upper panel. Figure 2 lower panel is a schematic diagram of the mediating role of FC strength in the interaction between emotion recognition and depressive symptoms.

4. Discussion

Depressive symptoms affect emotional recognition ability, and emotional recognition ability, in turn, affects depressive symptoms. Based on the international open database HCP, this study investigated the relationship and overlapped neuroimaging mechanisms between depressive symptoms and emotional recognition ability and also explored the mediating effect of the psychological factors on the interactions.

TABLE 2 Correlation analysis and mediating analysis of common brain structure between emotion recognition ability and depressive symptoms.

A/B	The average accuracy of task	Mediating#	The average RT of each test during task	Mediating#
ASR score of depression syndrome scale	−0.411***	A→B: −0.0021 B→A: −0.0024	0.152*	A→B: 0.002*
				B→A: 0.0023*
Age and sex corrected score of ASR depression syndrome scale	−0.383**	A→B: −0.0062** B→A: −0.0055*	0.133*	A→B: 0.0032**
				B→A: 0.0026*
ASR depression problems DSM-4 scale score	−0.347**	A→B: −0.0059* B→A: 0.0048*	0.129**	A→B: 0.0024
				B→A: 0.0028
Age and sex corrected score of ASR depression problems DSM-4 scale score	−0.328***	A→B: −0.0024* B→A: 0.0021*	0.106*	A→B: 0.0039*
				B→A: 0.0035*

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.
#Represents mediating analysis of inferior frontal gyrus orbital part between emotion recognition ability and depressive symptoms.

4.1. The relationship between emotion recognition ability and depressive symptoms

There is a negative correlation between emotional recognition ability and depressive symptoms. The score of depressive symptoms was not only negatively correlated with the accuracy of the emotion recognition task but also positively correlated with the reaction time of the emotion recognition task. The results of this study relatively clearly resolved the controversy in previous small sample studies on whether depressive symptoms could affect the ability of emotional recognition in images (9, 40).

TABLE 3 Brain structures associated with depressive symptoms and emotional recognition ability.

	L/R	β	*p*	*pval_fdr*
Brain structures associated with depressive symptoms				
Inferior frontal gyrus orbital part	L	3.029	0.0018	0.037
Parahippocampal gyrus	L	1.599	0.0024	0.046
Orbitofrontal lobe	L	3.187	0.0048	0.047
Brain structures associated with emotional processing ability				
Lingual gyrus	R	−72.083	0.0024	0.035
Lingual gyrus	L	−80.524	0.0034	0.044
Tegmental part of the inferior frontal gyrus	L	−79.301	0.0026	0.046
Cuneiform gyrus	R	−66.813	0.0038	0.047
Anterior coracoid cingulate gyrus	L	8.532	0.0018	0.042
Caudate anterior cingulate gyrus	L	15.614	0.0051	0.034
Paratalar gyrus	R	−64.363	0.0029	0.048
Inferior frontal gyrus orbital part	L	14.185	0.0031	0.043

L/R, left/right; β, regression coefficient; *pval_fdr*, *P*-value after FDR correction; $N = 995$.

4.2. sMRI mechanism of the relationship between emotion recognition ability and depressive symptoms

This study further explored the neural mechanism of depressive symptoms and decreased emotional recognition ability, and found that both were related to the cortical thickness of the inferior frontal gyrus orbital part. In individuals with higher levels of depressive symptoms, the inferior frontal gyrus orbital part cortex was thicker. In contrast, the thicker cortex had a lower level of emotional recognition than those with a thinner cortex. And then our results showed that the cortical thickness of the inferior frontal gyrus orbital part mediated the interaction between emotion recognition ability and depressive symptoms, which can partially verify the following hypothesis: the cortex close to the orbitofrontal gyrus is more related to emotional recognition under depressive symptoms.

4.3. fMRI mechanism of the relationship between emotion recognition ability and depressive symptoms

Our results showed that the participants who had been diagnosed with depressive disorders had stronger spontaneous activity, higher consistency and lower synchrony of spontaneous activity in the orbital part of the inferior frontal gyrus and the lateral orbitofrontal gyrus; Compared with participants with high emotion recognition ability, participants with low

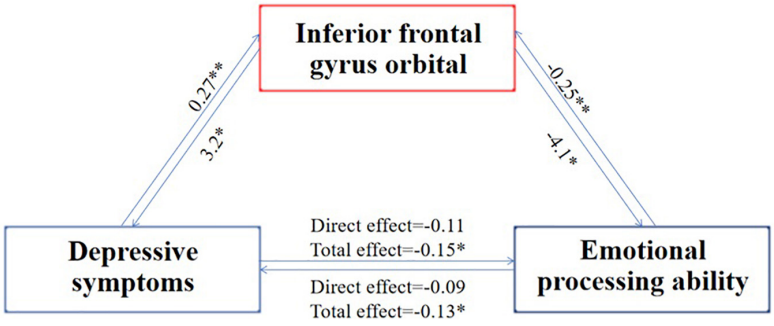


FIGURE 1
Mediating process of the common brain structure between emotion recognition ability and depressive symptoms. *N* = 995; the index of emotion recognition ability is the average reaction time of each test during the completion of the emotion recognition task, and the index of depression symptom is the score on the depression syndrome scale. The number on the arrow indicates the mediating effect coefficient, *indicates that the *p*-value is less than 0.05, and **indicates that the *p*-value is less than 0.01. The picture shows that the aggravation of depressive symptoms increases the thickness of the cortex of the inferior frontal gyrus orbital part, which in turn deteriorates the ability for emotional recognition, which in turn further increases the thickness of the cortex of the inferior frontal gyrus orbital part, which in turn aggravates the symptoms of depression.

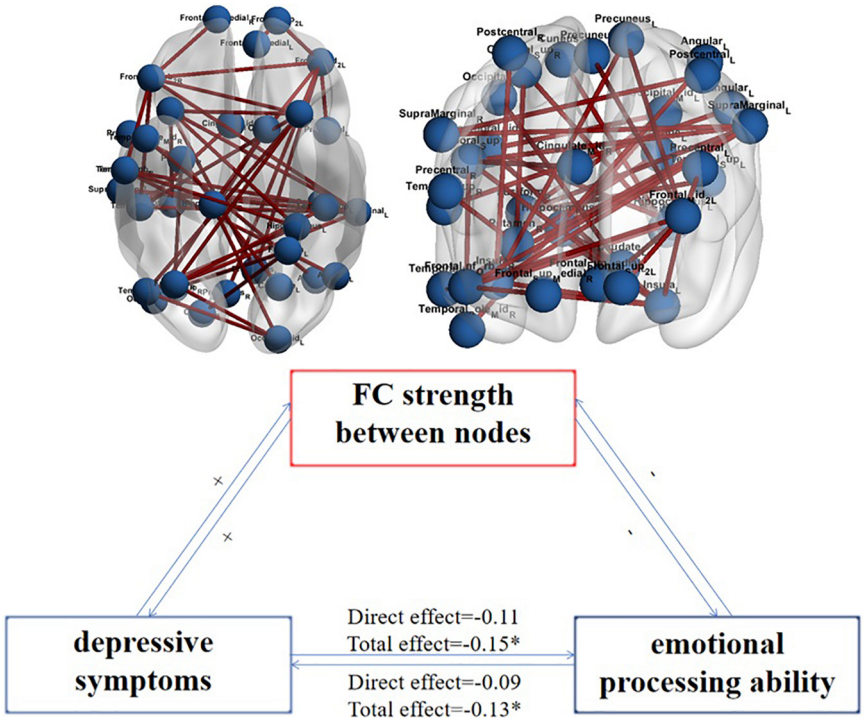


FIGURE 2
The FC strength between nodes plays an intermediary role in the interaction between emotion recognition ability and depressive symptoms and their schematic representation. *N* = 995; the upper figure shows the FCs that play an intermediary role in the interaction between emotion recognition ability and depressive symptoms; the variables on both sides are the DSM score on the depression scale and the average accuracy of the completed emotion recognition task. The lower figure shows that the aggravation of depressive symptoms increases the strength of FC between nodes (+), which makes the ability of emotional recognition worse (-), making the strength of FC between nodes further increase (+) and then more severe depressive symptoms (-). *Represents total effect is significant.

emotion recognition ability had stronger spontaneous activity, higher consistency and lower synchrony of spontaneous activity in the orbital part of the inferior frontal gyrus and the lateral orbitofrontal gyrus. And then, resting-state FC strength between the inferior frontal gyrus orbital part, fusiform gyrus, hippocampus, temporal gyrus, insula, precuneus, and cingulate gyrus mediated the interaction between emotion recognition ability and depressive symptoms. At last, the activation of the

frontal pole and precuneus associated with the inferior frontal gyrus orbital part in the emotional recognition task of patients with depression was higher than that of individuals without depression, and the activation of the anterior cingulate gyrus, insula and precuneus associated with the inferior frontal gyrus orbital part in emotional recognition tasks in patients with low completion of emotional recognition tasks were higher, and the activation of the posterior cingulate gyrus and angular gyrus associated with the inferior frontal gyrus orbital part in patients with low completion of emotional recognition tasks was lower. We did find some inconsistency of the brain regions that showed abnormal activities across different modalities. We believe that it is normal that the results of the resting state and task state are inconsistent. Their difference reflects the long-term and short-term differences in the process of emotion recognition. They already have some key brain regions to overlap so that it can be explained. Previous studies are not completely consistent, so we do believe it is acceptable.

Through the above results and previous studies (8), we can identify the critical brain regions involved in emotion recognition, including the orbitofrontal cortex, cingulate cortex, and precuneate lobe et. The orbital prefrontal cortex of humans and other primates is a crucial area of emotional recognition, which evaluates the value of stimuli and whether they are beneficial. The orbitofrontal cortex mainly projects to the anterior cingulate cortex, including its inferior commissural area, then depression is also associated with the anterior cingulate cortex (41). An rs-fMRI FC study of 654 participants shows the lateral orbitofrontal cortex has FC with inferior frontal gyrus (8). The medial orbitofrontal cortex has FC with the para-hippocampal gyrus, hippocampus, temporal cortex, fusiform gyrus, insular lobe, and cingulate gyrus. These FCs play important roles in depressive symptoms and low emotional recognition ability. A similar conclusion is obtained in our results on FC strength. In humans, activation of the medial orbitofrontal cortex is linearly correlated with the subjective (conscious) pleasure of stimulation (42, 43). These reward-related effects were found in pleasant touch, and monetary reward. A recent study of 1140 participants highlighted these ideas, showing that rewards (such as winning a prize or candy) activate the medial orbitofrontal cortex, while not winning activates the lateral orbitofrontal cortex (21). In addition, people with impaired orbitofrontal cortex may be less sensitive to rewards, which is reflected in the decrease in subjective emotion (44). It is difficult for them to recognize facial expressions and emotions related to sound, which is important for emotional and social behavior. It may be the reason for their difficulty in reversing the value of the reward (44). This is consistent with the view found in this study that the cortex near the orbitofrontal cortex is critical in the emotional recognition of patients with depression.

The previous studies and hypotheses on the orbitofrontal lobe (45, 46) were also verified in our research. Our results expand the role of the medial orbitofrontal lobe and lateral orbitofrontal cortex in depressive symptoms through intermediary analysis. The medial and lateral orbitofrontal cortex were distinguished in REHO numerical analysis. In depression, increased cerebral blood flow in areas including the lateral orbitofrontal cortex and cingulate cortex seems to be related to emotional changes because they become more normal when emotional states are relieved (47). The previous studies and hypotheses of FC in brain regions related to depressive symptoms were also verified in our study (20, 23). Our FC results found that the core brain regions such as the inferior orbitofrontal cortex, middle temporal gyrus, precuneate lobe, and cingulate gyrus are the potential neuroimaging basis for the interaction between depressive symptoms and emotional recognition ability.

4.4. The limitation of this study

This study has the following limitations: (1) the analysis data are from the international open database HCP, which currently has cross-sectional data, but no longitudinal follow-up data, so it is impossible to explain the long-term trend and causal relationship between emotion recognition ability and depressive symptoms. In the future, longitudinal follow-up studies should be carried out to explore the neural mechanism of emotional recognition ability and depressive symptoms. (2) The imaging data analysis of this study is mainly based on the data-driven method, but there is no preset hypothesis. Both MRI data and fMRI data are analyzed one by one, so the neuroimaging results are scattered, and in-depth data analysis and mining are needed. (3) Depressive symptoms are defined as the presence of one item in nine DSM4 entries, which is different from the entry criteria for the diagnosis of depression. The number of participants also limits this, so an extensive sample database for long-term follow-up is needed.

5. Conclusion

This study explores the relationship between emotion recognition ability and depressive symptoms and its neural mechanism based on the young population connectome plan of the HCP sub-project. This study not only elucidates the interaction between the emotional recognition ability and depressive symptoms mediated by the thickness of the inferior frontal gyrus orbital part but also shows that the FC strength of the rs-fMRI of the brain area centered on the inferior frontal gyrus orbital part, such as fusiform gyrus, insular, cingulate gyrus, middle temporal gyrus, precuneate lobe, and hippocampus mediates the interaction between emotion recognition and depressive symptoms. In addition, the predictive analysis of the brain regions and FC of the differences proved that the differences found were predictive.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: http://www.humanconnectomeproject.org.

Ethics statement

Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

Author contributions

All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved it for publication.

Funding

This study was supported by the National Key Research and Development Program of China (Nos. 2020YFC2003600 and 2021YFC0863700), the National Natural Science Foundation of China (Nos. 81761128036, 81821092, and 31900805), Young Elite Scientists Sponsorship Program by CAST (No. 2019 QNRC001), and PKU-Baidu Fund (No. 2020BD011).

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2022.1090369/full#supplementary-material

References

1. Ferrari A, Charlson F, Norman R, Patten S, Freedman G, Murray C, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. *PLoS Med.* (2013) 10:e1001547. doi: 10.1371/journal.pmed.1001547

2. Arroll B, Elley C, Fishman T, Goodyear-Smith F, Kenealy T, Blashki G, et al. Antidepressants versus placebo for depression in primary care. *Cochrane Datab Syst Rev.* (2009) 3:Cd007954. doi: 10.1002/14651858.Cd007954

3. Mojtabai R. Clinician-identified depression in community settings: concordance with structured-interview diagnoses. *Psychother Psychosomat.* (2013) 82:161–9. doi: 10.1159/000345968

4. Moncrieff J, Wessely S, Hardy R. Active placebos versus antidepressants for depression. *Cochrane Datab Syst Rev.* (2004) 2004:Cd003012. doi: 10.1002/14651858.CD003012.pub2

5. Fawcett J, Clark D, Scheftner W, Gibbons R. Assessing anhedonia in psychiatric patients. *Arch General Psychiatry.* (1983) 40:79–84.

6. Schlaepfer T, Cohen M, Frick C, Kosel M, Brodesser D, Axmacher N, et al. Deep brain stimulation to reward circuitry alleviates anhedonia in refractory major depression. *Neuropsychopharmacology.* (2008) 33:368–77. doi: 10.1038/sj.npp.1301408

7. Csukly G, Czobor P, Simon L, Takács B. Basic emotions and psychological distress: association between recognition of facial expressions and symptom checklist-90 subscales. *Comprehensive Psychiatry.* (2008) 49:177–83. doi: 10.1016/j.comppsych.2007.09.001

8. Du J, Rolls E, Cheng W, Li Y, Gong W, Qiu J, et al. Functional connectivity of the orbitofrontal cortex, anterior cingulate cortex, and inferior frontal gyrus in humans. *Cortex.* (2020) 123:185–99.

9. Mogg K, Bradley B. Attentional bias in generalized anxiety disorder versus depressive disorder. *Cogn Ther Res.* (2005) 29:29–45. doi: 10.1007/s10608-005-1646-y

10. Phillips M, Drevets W, Rauch S, Lane R. Neurobiology of emotion perception II: implications for major psychiatric disorders. *Biol Psychiatry.* (2003) 54:515–28. doi: 10.1016/s0006-322300171-9

11. Mayberg H. Defining the neural circuitry of depression: toward a new nosology with therapeutic implications. *Biol Psychiatry.* (2007) 61:729–30. doi: 10.1016/j.biopsych.2007.01.013

12. Phillips M, Ladouceur C, Drevets WCA. Neural model of voluntary and automatic emotion regulation: implications for understanding the pathophysiology and neurodevelopment of bipolar disorder. *Mol Psychiatry.* (2008) 13:833–57. doi: 10.1038/mp.2008.65

13. Stuhrmann A, Suslow T, Dannlowski U. Facial emotion processing in major depression: a systematic review of neuroimaging findings. *Biol Mood Anxiety Dis.* (2011) 1:10. doi: 10.1186/2045-5380-1-10

14. Drevets W, Price J, Furey M. Brain structural and functional abnormalities in mood disorders: implications for neurocircuitry models of depression. *Brain Struct Funct.* (2008) 213:93–118. doi: 10.1007/s00429-008-0189-x

15. Elliott R, Rubinsztein J, Sahakian B, Dolan R. The neural basis of mood-congruent processing biases in depression. *Arch General Psychiatry.* (2002) 59:597–604. doi: 10.1001/archpsyc.59.7.597

16. Phillips M, Drevets W, Rauch S, Lane R. Neurobiology of emotion perception I: the neural basis of normal emotion perception. *Biol Psychiatry.* (2003) 54:504–14.

17. Pessoa L, Ungerleider L. Neuroimaging studies of attention and the processing of emotion-laden stimuli. *Prog Brain Res.* (2004) 144:171–82.

18. Kober H, Barrett L, Joseph J, Bliss-Moreau E, Lindquist K, Wager T. Functional grouping and cortical-subcortical interactions in emotion: a meta-analysis of neuroimaging studies. *Neuroimage.* (2008) 42:998–1031. doi: 10.1016/j.neuroimage.2008.03.059

19. Frodl T, Scheuerecker J, Albrecht J, Kleemann A, Müller-Schunk S, Koutsouleris N, et al. Neuronal correlates of emotional processing in patients with major depression. *World J Biol Psychiatry Offi J World Federat Soc Biol Psychiatry.* (2009) 10:202–8. doi: 10.1080/15622970701624603

20. Rolls E, Cheng W, Gilson M, Qiu J, Hu Z, Ruan H, et al. Effective connectivity in depression. *Biol Psychiatry Cogn Neurosci Neur.* (2018) 3:187–97. doi: 10.1016/j.bpsc.2017.10.004

21. Xie C, Jia T, Rolls E, Robbins T, Sahakian B, Zhang J, et al. Reward versus nonreward sensitivity of the medial versus lateral orbitofrontal cortex relates to the severity of depressive symptoms. *Biol Psychiatry Cogn Neurosci Neur.* (2021) 6:259–69. doi: 10.1016/j.bpsc.2020.08.017

22. Rolls ETA. Non-reward attractor theory of depression. *Neurosci Biobehav Rev.* (2016) 68:47–58.

23. Freton M, Lemogne C, Bergouignan L, Delaveau P, Lehéricy S, Fossati P. The eye of the self: precuneus volume and visual perspective during autobiographical memory retrieval. *Brain Struct Funct.* (2014) 219:959–68. doi: 10.1007/s00429-013-0546-2

24. Cheng W, Rolls E, Gong W, Du J, Zhang J, Zhang X, et al. Sleep duration, brain structure, and psychiatric and cognitive problems in children. *Mol Psychiatry.* (2021) 26:3992–4003.

25. Ma Y. Neuropsychological mechanism underlying antidepressant effect: a systematic meta-analysis. *Mol Psychiatry.* (2015) 20:311–9. doi: 10.1038/mp.2014.24

26. Grieve S, Korgaonkar M, Koslow S, Gordon E, Williams L. Widespread reductions in gray matter volume in depression. *Neur Clin.* (2013) 3:332–9.

27. Lorenzetti V, Allen N, Fornito A, Yücel M. Structural brain abnormalities in major depressive disorder: a selective review of recent mri studies. *J Affect Dis.* (2009) 117:1–17.

28. Achenbach T, Rescorla L. *Manual for the aseba adult forms & profiles*. Burlington, VT: University of Vermont, Research Center for Children, Youth (2003).

29. Achenbach T. *Manual for the Young Adult Self-Report and Young Adult Behavior Checklist*. Burlington, VT: University of Vermont, Department of Psychiatry (1997).

30. Kalkstein S, Hurford I, Gur R. Neurocognition in schizophrenia. *Curr Top Behav Neurosci.* (2010) 4:373–90. doi: 10.1007/7854_2010_42

31. Gur R, McGrath C, Chan R, Schroeder L, Turner T, Turetsky B, et al. An fmri study of facial emotion processing in patients with schizophrenia. *Am J Psychiatry.* (2002) 159:1992–9. doi: 10.1176/appi.ajp.159.12.1992

32. Gur R, Schroeder L, Turner T, McGrath C, Chan R, Turetsky B, et al. Brain activation during facial emotion processing. *Neuroimage.* (2002) 16:651–62. doi: 10.1006/nimg.2002.1097

33. Hayes A. *Introduction to mediation, moderation, and conditional process analysis: a regression-based approach*. New York: Guilford Press (2017).

34. Shen X, Tokoglu F, Papademetris X, Constable R. Groupwise whole-brain parcellation from resting-state fmri data for network node identification. *Neuroimage.* (2013) 82:403–15. doi: 10.1016/j.neuroimage.2013.05.081

35. Chen X, Lu B, Yan C. Reproducibility of R-fmri metrics on the impact of different strategies for multiple comparison correction and sample sizes. *Hum Brain Mapp.* (2018) 39:300–18. doi: 10.1002/hbm.23843

36. Winkler A, Ridgway G, Douaud G, Nichols T, Smith S. Faster permutation inference in brain imaging. *Neuroimage.* (2016) 141:502–16. doi: 10.1016/j.neuroimage.2016.05.068

37. Gur R, Erwin R, Gur R, Zwil A, Heimberg C, Kraemer H. Facial emotion discrimination: II. *Behav Find Dep Psychiatry Res.* (1992) 42:241–51. doi: 10.1016/0165-178190116-k

38. Bouhuys A, Geerts E, Gordijn M. Depressed patients' perceptions of facial emotions in depressed and remitted states are associated with relapse: a longitudinal study. *J Nervous Mental Dis.* (1999) 187:595–602. doi: 10.1097/00005053-199910000-00002

39. Surguladze S, Young A, Senior C, Brebion G, Travis M, Phillips M. Recognition accuracy and response bias to happy and sad facial expressions in patients with major depression. *Neuropsychology.* (2004) 18:212–8. doi: 10.1037/0894-4105.18.2.212

40. Van Vleet T, Stark-Inbar A, Merzenich M, Jordan J, Wallace D, Lee M, et al. Biases in processing of mood-congruent facial expressions in depression. *Psychiatry Res.* (2019) 275:143–8. doi: 10.1016/j.psychres.2019.02.076

41. Rolls E. The neuroscience of emotional disorders. *Handbook Clin Neurol.* (2021) 183:1–26. doi: 10.1016/b978-0-12-822290-4.00002-5

42. Rolls E. The orbitofrontal cortex. *Philos Trans R Soc London Seri B Biol Sci.* (1996) 351:1433–44.

43. Rolls E, Grabenhorst F. The orbitofrontal cortex and beyond: from affect to decision-making. *Prog Neurobio.* (2008) 86:216–44. doi: 10.1016/j.pneurobio.2008.09.001

44. Fellows L. Orbitofrontal contributions to value-based decision making: evidence from humans with frontal lobe damage. *Ann NY Acad Sci.* (2011) 1239:51–8.

45. Iversen S, Mishkin M. Perseverative interference in monkeys following selective lesions of the inferior prefrontal convexity. *Exp Brain Res.* (1970) 11:376–86. doi: 10.1007/bf00237911

46. Murray E, Rudebeck P. Specializations for reward-guided decision-making in the primate ventral prefrontal cortex. *Nat Rev Neurosci.* (2018) 19:404–17. doi: 10.1038/s41583-018-0013-4

47. Drevets W. Orbitofrontal cortex function and structure in depression. *Ann NY Acad Sci.* (2007) 1121:499–527.

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 14 February 2023
DOI 10.3389/fpsyt.2023.1104190

OPEN ACCESS

EDITED BY
Xiaozheng Liu,
The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, China

REVIEWED BY
Suhua Chang,
Peking University Sixth Hospital, China
Yi Zhang,
Zhejiang University, China

*CORRESPONDENCE
Gang Wang
gangwangdoc@ccmu.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Computational Psychiatry, a section of the journal Frontiers in Psychiatry

RECEIVED 21 November 2022
ACCEPTED 31 January 2023
PUBLISHED 14 February 2023

CITATION
Li N, Feng L, Hu J, Jiang L, Wang J, Han J, Gan L, He Z and Wang G (2023) Using deeply time-series semantics to assess depressive symptoms based on clinical interview speech. *Front. Psychiatry* 14:1104190. doi: 10.3389/fpsyt.2023.1104190

Using deeply time-series semantics to assess depressive symptoms based on clinical interview speech

Nanxi Li[1†], Lei Feng[1,2†], Jiaxue Hu[3], Lei Jiang[3], Jing Wang[3], Jiali Han[1], Lu Gan[3], Zhiyang He[3] and Gang Wang[1,2*]

[1]Beijing Key Laboratory of Mental Disorders, National Clinical Research Center for Mental Disorders and National Center for Mental Disorders, Beijing Anding Hospital, Capital Medical University, Beijing, China, [2]Advanced Innovation Center for Human Brain Protection, Capital Medical University, Beijing, China, [3]Anhui iFLYTEK Health Co., Ltd., Hefei, China

Introduction: Depression is an affective disorder that contributes to a significant global burden of disease. Measurement-Based Care (MBC) is advocated during the full course management, with symptom assessment being an important component. Rating scales are widely used as convenient and powerful assessment tool, but they are influenced by the subjectivity and consistency of the raters. The assessment of depressive symptoms is usually conducted with a clear purpose and restricted content, such as clinical interviews based on the Hamilton Depression Rating Scale (HAMD), so that the results are easy to obtain and quantify. Artificial Intelligence (AI) techniques are used due to their objective, stable and consistent performance, and are suitable for assessing depressive symptoms. Therefore, this study applied Deep Learning (DL)-based Natural Language Processing (NLP) techniques to assess depressive symptoms during clinical interviews; thus, we proposed an algorithm model, explored the feasibility of the techniques, and evaluated their performance.

Methods: The study included 329 patients with Major Depressive Episode. Clinical interviews based on the HAMD-17 were conducted by trained psychiatrists, whose speech was simultaneously recorded. A total of 387 audio recordings were included in the final analysis. A deeply time-series semantics model for the assessment of depressive symptoms based on multi-granularity and multi-task joint training (MGMT) is proposed.

Results: The performance of MGMT is acceptable for assessing depressive symptoms with an F1 score (a metric of model performance, the harmonic mean of precision and recall) of 0.719 in classifying the four-level severity of depression and an F1 score of 0.890 in identifying the presence of depressive symptoms.

Disscussion: This study demonstrates the feasibility of the DL and the NLP techniques applied to the clinical interview and the assessment of depressive symptoms. However, there are limitations to this study, including the lack of adequate samples, and the fact that using speech content alone to assess depressive symptoms loses the information gained through observation. A multi-dimensional model combing semantics with speech voice, facial expression, and other valuable information, as well as taking into account personalized information, is a possible direction in the future.

KEYWORDS

depression, mood disorder, psychiatric assessment, semantic, time-series, natural language processing

Introduction

Depression is a common mental disorder characterized by a persistently depressed mood, a loss of pleasure or interest in activities, and other associated symptoms. The World Health Organization (WHO) reports that approximately 5% of adults worldwide suffer from depression. Depression is a major contributor to the global burden of disease (1). Appropriate assessment plays a key role in clinical practice with Measurement-Based Care (MBC) being recommended for depression management in several clinical practice guidelines (2). Symptoms assessment is one of the most important dimensions with a number of scales available in depression evaluation.

The Hamilton Depression Rating Scale (HAMD) is the most commonly used assessment tool to probe the presence of depressive and associated symptoms, and is considered the "gold standard" of depression measurement (3), and has been used to establish the criteria for the severity level of depressive symptoms (4). Numerous versions of the HAMD exist, and a 17-item version of the Hamilton Depression Rating Scale (HAMD-17) is the most classic and widely used version (5). Each item of the HAMD-17 examines a subsymptom of depression, and some of the items are formed into a factor structure. The psychopathology and symptom clusters can be specifically characterized by factor analysis. The HAMD is not only used for depression, but can also be used for a variety of diseases such as bipolar disorder (BD), neurological disorders and other medical conditions with depressive symptoms. Therefore, a clinical interview incorporating HAMD-17 is appropriate for broad questioning and assessing depressive symptoms.

As the HAMD-17 is a hetero-rated scale, it requires a trained rater, with specific expertise, sufficient knowledge of the scale, and reliable accuracy. The scale was originally designed to be completed after an unstructured clinical interview. Although semi-structured interview guides are available (6) that record only the score and not the interview process, there is a risk of bias in obtaining accurate scores based on the unreviewable interview. In addition, due to medical resource constraints, there is a need for more professionals to conduct regularly high quality assessments in a real-world clinical setting.

Natural language processing (NLP) is a branch of AI that focuses specifically on understanding, interpreting, and manipulating large amounts of human language and speech data. Since the 1980s, NLP has combined computational linguistics with statistical machine learning and Deep Learning models in order to take unstructured, free-form data and produce structured, quantitative linguistic outputs. With the growth of available public data, NLP technology based on time-series learning has grown significantly in recent years (7), particularly in medicine, where more and more research is demonstrating the value of Deep Learning-based NLP (8). The use of Deep Learning-based NLP in medicine is particularly useful in prediction and reverse distillation based on regular medical records for risk assessment (9), thus, using time-series semantic information by simulating clinical decision-making for risk forecasting (10). Resent research has shown that NLP has the ability to perform highly repetitive manual tasks consistently and to integrate and compute knowledge efficiently compared to human beings, opening up more opportunities for the use of NLP in practical applications. The assessment of specific depressive symptoms is a suitable application for NLP. The process of assessing depressive symptoms, because of its clear purpose and specific content, depends on specific expertise and information integration based on the interview. In addition, HAMD-17 has provided a framework for interviewing and evaluating depressive symptoms, as well as normative classification criteria for the severity level of depressive symptoms, which meets the need for the application of Deep Learning techniques.

Numerous studies have focused on and attempted to apply NLP technology approaches to detect and evaluate depression; however, they have mostly extracted data in the form of written text (11), which differs significantly from oral text. The data used for NLP has been extracted from electronic health records (12) and social media (13). The text is either processed by a doctor or without any professional processing, and the semantic density of the accessible information is sparse compared to specific interviews about depressive symptoms. Therefore, it is valuable to apply NLP techniques directly to the interview text for the assessment of depressive symptoms in order to build a framework of depressive semantic model, thus providing the opportunity to bring AI technology into psychiatric clinical practice in the future.

The aims of this study include: (1) to construct a task-oriented algorithmic model using text from specific clinical interviews, and to validate the feasibility of Deep Learning-based NLP for depressive symptoms assessment. It should be noted that the model construction and the core algorithm are not the entire purpose of this study, but rather a methodological approach; therefore, its introduction is presented in the Data Analytic Strategy section, (2) to validate the proposed time-series semantic algorithm model, and to measure the performance of classifying the depressive symptoms severity level.

Materials and methods

Data for this study were derived from two clinical research projects, one is about emotional recognition among patients with depression, and the other is regarding identifying unipolar and bipolar depression using speech. The Ethical Committee of Beijing Anding Hospital has approved both projects. Each participant was asked to sign a written informed consent before data collection.

Participants

In this study, 329 participants with Major Depressive Episode (MDE) were recruited at the Beijing Anding Hospital inpatient or outpatient departments from September 2020 to July 2022. At the time of enrolment, 233 participants were diagnosed with Major Depressive Disorder (MDD), diagnosed by experienced psychiatrists according to the ICD-10 for inpatients and the Mini International Neuropsychiatric Interview (MINI, version 7.0.2) for outpatients. In addition,, 96 participants were diagnosed with bipolar disorder having concurrent depressive episodes (BDD) using the MINI. All participants met the inclusion criteria, which included: (a) age between 18 and 65 years, (b) speaking Chinese without obvious dialect, (c) educational level of primary school or above, (d) understanding and cooperating with the research content. Exclusion criteria included: (a) a diagnosis of schizophrenia, schizoaffective disorder or other mental disorders, (b) a history of organic brain disease. All 329 participants had a mean age

of 34.1 (SD: 12.4) years, ranging from 18 to 64 years, and 66.0% (n = 217) of the participants were female. The participants' current MDE lasted 28.8 (SD: 47.9) weeks, with a mean HAMD-17 score of 20.2 (SD:5.72).

Procedure and measures

Each participant was asked to complete a face-to-face clinical interview using the HAMD-17 with an audio recording. All interviewers were trained and scorer reliability was maintained to ensure the quality of the interview. A number of standard phrases were developed in the interview outline which were used to locate the interview content and facilitate text processing. All interviews were conducted in special test rooms with no noticeable background noise. The recording device was either an audio recorder (brand and model: iFLYTEK SR502) or a smartphone (brand and model: honor 9X). The recording device was placed approximately 50 cm away from the participant.

Basic demographic information and a brief medical history were collected before the audio recording. The HAMD-17 was used in this study. In this version, each item is scored from 0 to 2 or from 0 to 4, and the total score ranges from 0 to 52. We defined the cut-off points, and the severity levels of the depressive symptoms as follows: >24 = severe depression, 18–24 = moderate depression, 8–17 = mild depression, <8 = euthymia. A total of 387 audio recordings were collected during the study, as 58 of the 233 participants with MDD received the same secondary clinical interview 4 weeks after the initial interview. Finally, according to the HAMD-17 total score, 46 audio recordings were classified as euthymia, 102 were classified as mild depression, 160 were classified as moderate depression, and 79 were classified as severe depression.

Data preprocessing

The initial form of data collected was an audio recording of the clinical interview between the doctor and the patient. Considering the content composition of the audio, the data was pre-processed in four steps: speaker diarization, role identification, speech recognition, and item-centered classification. The final data output is presented as a structured Chinese text of the doctor-patient dialogue (see **Supplementary Appendix 1** for more details). The speaker diarization and speech recognition technology are supported by the iFLYTEK open platform.[1] Role identification was used to distinguish between doctor and patient through a rule-based approach. Two main rules are used: (1) after building a library of doctor question sentences, the edit distance is calculated from the input data to determine whether the role is a doctor, (2) a keyword database of question sentences was summarized and constructed, the identities of the doctor and the patient are determined by calculating the frequency of keywords throughout the conversation. An item-centered classification scheme, based on the temporal analysis of bidirectional long and short-term memory (BiLSTM) (14), was used to cut the text and extract the content related to the corresponding HAMD-17 item using a pair of question and answer (Q&A) sentences as the input of each temporal step as well as the corresponding item names as the output. The topics measured by each item of HAMD-17 are defined and described in the model as *scene*.

1 https://www.xfyun.cn/

Dataset settings

The dataset consisting of 387 audio recordings was randomly grouped into training set and test set in a ratio of 7:3, with no significant differences in the overall distribution of the HAMD-17 total score. Thus, 114 audio recordings were put into the test set, and the corresponding severity level of depressive symptoms was distributed as 23 euthymia, 50 mild depression, 30 moderate depression, and 11 severe depression. Additionally, 273 audio recordings were put into the training set, and the distribution of the corresponding severity level of depressive symptoms is 23 euthymia, 52 mild depression, 130 moderate depression, and 68 severe depression. The detailed distribution is shown in **Table 1**.

Model algorithms

Applying the level-classification measurement of depressive symptoms, a time-series semantics model based on multi-granularity and multi-task joint training (MGMT) was customized in this study. A brief introduction of MGMT is provided as the model and is being proposed for the first time.

General framework

The clinical interview speech is in the form of a doctor-patient dialogue which has a certain temporal development pattern. A quantitative scheme based on the time-series is used in this algorithm. For any given audio dialogue, after data pre-processing, the entire text of one scene is represented by X, $X = \{x_i^j: 0 < j < L, 0 < i = 17\}$, in which i is the subscript of the scene, j is the text subscript of each scene, and L refers to the maximum text length. The text of each scene is encoded using Bidirectional Encoder Representation from Transformers (BERT) (15), [CLS] output, a token with no obvious semantic information, is taken as the text embedding of each scene, H is referred to as scene coding, and $H = \{h_i, 0 < = i < = 17, h_i \in R^d\}$, the d is 256 dimensions in this scheme. The scene granularity score is represented as S, where $S = \{s_i, 0 < i < = 17, s_i \in R^1\}$. The formulas are shown as follows:

$$H = \text{Pool (BERT(X))}$$

$$S = \text{Softmax (H)}$$

While the H and S are calculated in the set of independent scene, BiLSTM is introduced to obtain information from the holistic dialogue and to concatenate scenes according to the time-series. The equations are present below. $H^F \in R^d$ and $H^b \in R^d$ are output vectors of the forward and backward LSTMs, respectively.

$$H^F = \overrightarrow{LSTM}(H)$$

$$H^b = \overleftarrow{LSTM}(H)$$

$$H = [H^F, H^b]$$

After performing Self-Attention, Multilayer Perceptron (MLP), and Softmax, combined with the obtained information

TABLE 1 Distribution of the number of audio data in training set and test set.

Depression severity level	Number of audio		
	Sum of total (387)	Training set (273)	Test set (114)
Euthymia	46	23	23
Mild depression	102	52	50
Moderate depression	160	130	30
Severe depression	79	68	11

from entire dialogue the depressive severity level and its probability from a holistic perspective are presented as P^e, and $P^e = \{p_i^e, 0 < = i < 4\}$.

The final level is produced by the Decision Level Fusion competition. Equations are expressed as follows:

$$S\_e2e\_level = Softmax(MLP(Self - Attention(H)))$$

$$S\_scene\_level = HAMD\ score(S)$$

$$Level = Decision\ Level\ Fusion(s\_e2e\_level, S\_scene\_level)$$

Three strategies were used in the Decision Level Fusion module: (a) taking the S_e2e_level as the final level, (b) taking the S_scene_level as the final level, and (c) weighting the scene levels to the corresponding total score, reordering the levels based on corresponding probabilities, and taking the level with the highest probability as the final level. The general framework is illustrated in Figure 1.

Multi-task joint optimization

The optimization of MGMT is conducted using multi-task joint training with scene and HAMD scores. Due to dialogue content varying in complexity and the scene scores are unevenly distributed. The Focal Loss (16) and the GHM Loss (17) were used as the joint loss function for optimization. The idea of the Focal Loss is to reduce the weight of easily distinguishable samples (i.e., samples with high confidence) and to increase the weight of hard to distinguish samples, forcing the model to pay more attention to these hard to distinguish samples. Furthermore, considering the existence of many indistinguishable, mislabeled, and confused samples due to the lack of information the stability and optimization direction of the model will be affected. The GHM Loss is introduced to balance attention to the indistinguishable samples. Therefore, the final loss is the weighted and balanced result of the Focal Loss and GHM Loss, which is represented as:

$$loss = L_{FL} + \gamma * L_{GHM}$$

In the above formulae, γ is hyperparameter configured to balance the weight of L_{GHM}, which is specified by the distribution of the data set. Finally, the loss function of MGMT is present as the following equation where α and β are hyperparameters configured to balance the weight of the scene and the weight of the End-to-End separately.

$$Loss = \alpha * loss_{scene} + \beta * loss_{e2e}$$

Considering that in an actual clinical setting, there may be some scenes that are incomplete or insufficient to provide valid information; therefore, each scene is masked with a probability of 5%. The score of the masked scene is set to 0 and the entire data is used as a new sample for training. This operation is intended to simulate the real interview process and to enhance the generalization ability of the model.

Model training and performance evaluation

A BERT-small model trained each model with 6 transformer block layers, with each block having a hidden size of 256 and 12 multi-head attentions. The models were trained on a Tesla V100 32G GPU with a training epoch of 100 and a batch size of 16. An Adaptive Moment Estimation (Adam) optimizer with an initial learning rate of 0.001 and a warm-up learning rate decay strategy was used during model training. A single layer BiLSTM was used for temporal aggregation training with a hidden size of 256. The hyperparameters α,β, andγ were set to 0.8, 0.2, and 0.2, and the maximum length (L) of each scene was set to 512.

Model performance was measured using an F1 score defined as the following formulae. True Positive (TP): judged to be a positive sample when in fact it is a positive sample. False Positive (FP): judged to be a positive sample when in fact it is a negative sample. False Negative (FN): judged to be a negative sample when in fact it is a positive sample. The F1 score can be interpreted as a weighted average of precision and recall, with values taken at an interval range of 0–1.

$$F1 = \frac{2TP}{2TP+FP+FN}$$

Results

Formatted text description

The mean number of interactive rounds of the total data were 85.5 (SD: 24.5), the mean number of words spoken by each participant was 1,100 (SD: 599) counts, and the mean length of the audio was 681 (SD: 210) seconds. As we used Automatic Speech Recognition (ASR) technology, the quality of the formatted text is measured using word correctness (Corr) and word accuracy (Acc). Corr = H/N, where H represents the number of correctly recognized words and N is the total number of recognized words. Acc = (H-I)/N, H and N have the same meaning as before, I represents the number of inserted non-existent words. The formatted text used in this study is of good quality, with a Corr of 94.60% and an Acc of 93.10%.

Classification result

Three types of classification were used to assess the performance of MGMT in evaluating depressive symptoms, including: (a) a four-level classification of depression severity (severe depression vs. moderate depression vs. mild depression vs. euthymia), (b) a binary classification of mild depressive and severe depressive symptoms (severe depression and moderate depression vs. mild depression and euthymia), and (c) a binary classification of identifying the presence of depressive symptoms (severe depression and moderate depression and mild depression vs. euthymia). In addition, 96 of the 387 audio recordings were collected from participants with BDD, who had a different diagnosis from those with MDD, so the original test set (ALL) was divided into a dataset including only MDD (with 92 audio recordings) and a dataset including only BDD (with 22 audio recordings) in order to verify the robust of MGMT.

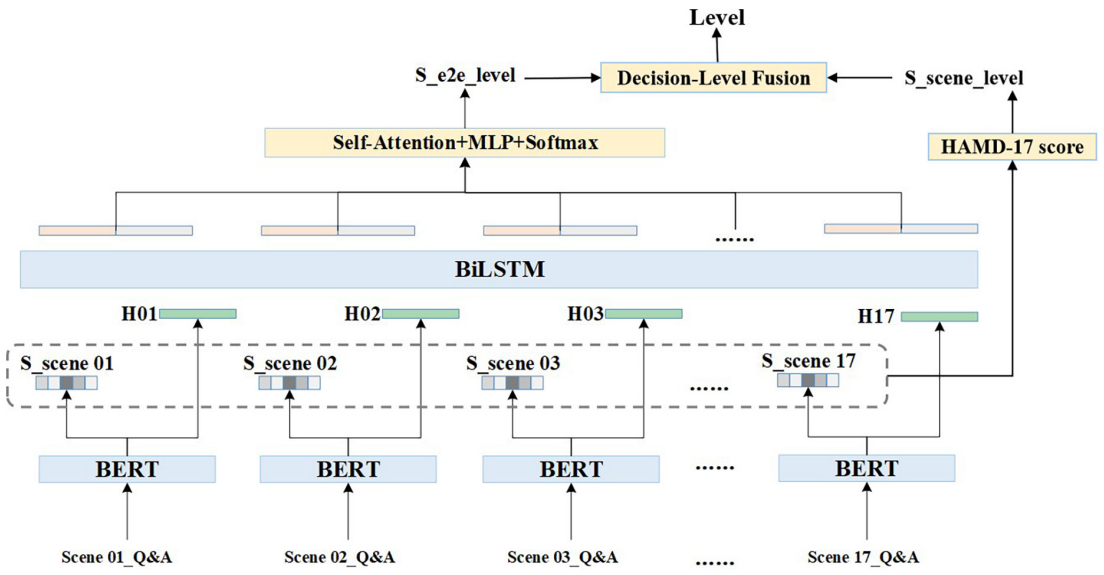


FIGURE 1
General framework of multi-granularity and multi-task joint training (MGMT).

Using the ALL training set, MGMT with ALL test set (ALL-ALL) has an F1 score of 0.719 in the classification of the four-level depression severity, 0.884 in the binary classification of mild depressive and severe depressive symptoms, and 0.890 in the binary classification of identifying the presence of depressive symptoms. For only MDD included in test set (ALL-MDD), MGMT has an F1 score of 0.706 in the four-level depression severity classification, 0.913 in the binary classification of mild and severe depression, and 0.837 in the binary classification of identifying the presence of depressive symptoms. There is no significant difference in the F1 score between ALL-ALL and ALL-MDD. Using the test set only including BDD (ALL-BDD), the best result is obtained with an F1 score of 0.772 in the classification of the four-level depression severity and an F1 score of 0.955 in both binary classification models. The results are shown in Table 2.

The specific values for the accuracy of the four-level depression severity classification are shown in Table 3. In ALL-ALL, MGMT has an accuracy of 69.57% (16 of 23) in identifying euthymia, 76.00% (38 of 50) in identifying mild depression, 73.33% (22 of 30) in identifying moderate depression, and 54.55% (6 of 11) in identifying severe depression. In ALL-MDD, the performance of MGMT is similar to that of ALL-ALL, with an accuracy of 66.67% (14 of 21) in identifying euthymia, 73.17% (29 of 41) in identifying mild depression, 70.83% (17 of 24) in identifying moderate depression, and 66.67% (4 of 6) in identifying severe depression. In ALL-BDD, MGMT is correct in identifying all of euthymia, and has an accuracy of 88.89% (8 of 9) in identifying mild depression, 83.33% (5 of 6) in identifying moderate depression, while having an accuracy of 40.00% (2 of 5) in identifying severe depression.

TABLE 2 Performance of MGMT in depressive symptoms evaluation (ALL as training set).

Classification type	F1 score		
	ALL-ALL	ALL-MDD	ALL-BDD
Four-level depression severity classification	0.719	0.706	0.772
Mild/severe depression binary classification	0.884	0.913	0.955
Euthymia/depression binary classification	0.89	0.837	0.955

TABLE 3 Performance of MGMT in four-level depression severity classification of three test sets.

Depression severity level	Accuracy (correct/total number of audio)		
	ALL-ALL	ALL-MDD	ALL-BDD
Euthymia	69.57%	66.67%	100.00% (2/2)
Mild depression	76.00%	73.17%	88.89% (8/9)
Moderate depression	73.33%	70.83%	83.33% (5/6)
Severe depression	54.55%	66.67%	40.00% (2/5)

To present the results more clearly, the confusion matrix of the four-level depression severity classification of ALL-ALL is shown in Table 4. Misclassification occurs more frequently in the proximity category.

Sensitivity analysis

To further verify the reliability of the above results, we used only a total of 199 audio recordings of MDD patients as training set data (MDD-MDD), removing those of BDD patients from the training set. The performance of MGMT is shown in Table 5. MGMT has an F1 score of 0.685 in the four-level depression severity classification with MDD-MDD, which is slightly lower than that being produced by ALL training set (ALL-MDD). MGMT has an F1 score of 0.902 in the binary classification of mild depressive and severe depressive symptoms, and 0.826 in the binary classification of identifying the presence of depressive symptoms.

Table 6 shows the performance of MGMT with MDD-MDD in the four-level depression severity classification, MGMT owns an accuracy of 61.90% in identifying euthymia (13 of 21), 73.17% in

TABLE 4 Confusion matrix of the four-level depression severity classification of ALL-ALL.

Confusion matrix		Predict			
		Euthymia	Mild depression	Moderate depression	Severe depression
True	Euthymia	16	7	0	0
	Mild depression	8	38	4	0
	Moderate depression	1	3	22	4
	Severe depression	0	1	4	6

TABLE 5 Performance of MGMT in depressive symptoms evaluation (MDD-MDD).

Classification type	F1 score
Four-level depression severity classification	0.685
Mild/severe depression binary classification	0.902
Euthymia/depression binary classification	0.826

TABLE 6 Performance of MGMT in four-level depression severity classification (MDD-MDD).

Depression severity level	Accuracy
Euthymia	61.90%
Mild depression	73.17%
Moderate depression	66.67%
Severe depression	66.67%

identifying mild depression (30 of 41), and an accuracy of 66.67% in identifying both moderate depression (16 of 24) and severe depression (4 of 6). There is a slight decrease in accuracy compared to ALL-MDD.

Discussion

The present study developed a time-series semantics model primarily based on multi-granularity and multi-task joint training. MGMT obtained information about depressive symptoms in various dimensions, performed well on the task of classifying the severity level of depressive symptoms, and demonstrated the feasibility of Deep Learning combined with NLP applied to psychiatric assessment.

Early studies of depressive speech using computational analysis were generally based on psycholinguistics, with measures characterizing lexical diversity, syntactic complexity and speech content. Word counting, at the level of vocabulary granularity, was the most direct method of representing the speech characteristics. Relevant studies have identified differences in the frequency of first-person singular pronouns, negative mood words, and positive mood words between depressed patients and healthy controls (18). These differences were also found in patients with different severity levels of depressive symptoms (19). Sentence-level analysis provides insight into cognitive-linguistic conditions through the sentence structure. Sentence structure changes are less significant in patients with depression than in those with schizophrenia or Alzheimer's disease (20). Changes in speech at the sentence-level in depressed individuals are more likely to be summarized by a reduced number of words in a sentence and a decrease in overall speech activity. While these changes have been shown to correlate with attention and psychomotor speed performance; however, they are less correlated with depressive severity (21).

According to the results in Tables 3, 6, the different performance of MGMT in distinguishing depression severity is in line with the corresponding sample size. The accuracy of the classification is relatively low with a small sample size. The small sample size is reflected not only in the audio number rated as severe depression but also in the frequency of occurrence of the extreme point of each item, especially when the variables are measured on a five-point. In addition, the confusion matrix in Table 4 shows that confusion occurs more frequently in the proximity category. The model has not learnt the key point of classification through adequate samples, it is prone to misjudgment when encountering unfamiliar or rare text. This just further confirms the importance of having a sufficient number of samples with clear distinctions for model training. Although the severity levels are conventionally and strictly divided by the HAMD-17 total score, samples on either side of the cut-off value have high similarity and low discrimination. A difference of 1 point in the total score may not make a significant difference in the evaluation of depressive symptoms, although they belong to different severity levels of depressive symptoms. The better option in this situation is to use the HAMD-17 total score directly or to redivide the depression severity sub-intervals. Moreover, the clear differentiation between samples is reflected in different scores for each item, different combinations of item scores, and slight differences in the total score. It is difficult to obtain sufficient data to build an equally distributed training data in a clinical setting, while it is critical for Deep Learning-based NLP technology.

In addition to the sample size mentioned above, there is another limitation of this study. Audio recordings contain information from both text and voice, and voice characteristics are also considered valuable in the assessment of depressive symptoms. Studies have found differences between the voices of depressed and healthy people (22), and acoustic features are correlate with the severity of depressive symptoms and their variability (23). Several acoustic features are thought to correlate with depression (24), Machine Learning and Deep Learning techniques have been widely used in the studies of voice analysis (25). For the audio recordings in this study, we also performed some voice analysis. we conducted a multi-feature decision fusion classification model, including X-vector, the extended Geneva Minimalistic Acoustic Parameter Set (eGeMAPs) (26), wav2vec 2.0, and conformer ASR decoder features. However, the accuracy of the model in classifying four-level severity of depression is only 43.86%. We also attempted to construct a semantic-voice fusion model, and the performance of this model did not improve over the text-only model conducted in this study. Therefore, we have mainly confirmed the value of semantics in depression severity classification without adding voice analysis.

Agitation and retardation are observational items based on the patient's behavior during clinical interview. Using speech text alone is impractical for observational assessment and may lead to bias in the assessment of depressive severity using the HAMD rating scale. The external performance of agitation is excessive physical activity with significant fidgeting, tension and excitement, and psychomotor retardation can be detected by speech volume,

response duration, and movement changes. Simulating actual human judgment, performance can be captured by other forms of behavioral indicators, such as speech as mentioned above, as well as, facial expressions and gestures. Datasets consisting of facial expressions and gestures are available for clinical analysis, and several features are associated with depressive symptoms, which can be used to construct depression detection models (27). Multidimensional information helps to optimize the assessment of depressive symptoms and compensates for observational information (28).

This study has established a general framework for assessing the severity of depression using clinical speech; thus, a deeply time-series semantics model has been constructed. The algorithm model has significant clinical application value because face-to-face interview speech related to the HAMD-17 assessment was selected as the corpus. These were highly correlated with depressive symptoms and closely related to the assessment in real clinical practice. MGMT takes into account the multiple granularity of information as much as possible at the scene-level. Effective and differentiated information results in an accurate score for a given scene; thus the ability to provide an accurate HAMD-17 total score can be developed. The multi-task setting of MGMT has considerable potential for use in research on sub-syndromes, specific dimensions and across diagnostic symptoms, as well as in individualized purpose-oriented studies.

Conclusion

In this study, we designed and tested an algorithmic model for depression severity classification. For clinical interview text related to depressive symptoms evaluation, considering its time-series, we proposed a Deep Learning-based NLP model based on multi-granularity and multi-task joint training, making full use of each item as well as the overall information. The test results of the proposed model demonstrate the feasibility of applying Deep Learning techniques to depressive symptoms severity assessment and have shown excellent performance. Future studies using more appropriate datasets will allow us to further improve our approach to assessing depressive symptoms. A multi-dimensional model may be a potential research direction in the future.

Data availability statement

The original contributions presented in this study are included in this article/**Supplementary material**, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethical Committee of Beijing Anding Hospital. The patients/participants provided their written informed consent to participate in this study.

Author contributions

GW and ZH conceived the presented idea. JH supervised the project and the findings of this work. NL and LF wrote the manuscript. JW and LJ developed the model and carried out the experiment. NL and JHa collected the data. LJ and LG pre-processed and annotated the dataset. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by STI2030-Major Projects (2021ZD0200600), Beijing Municipal Administration of Hospitals Incubating Program (PX2022074), and Beijing High Level Public Health Technical Talents Training Plan (xuekegugan-01-12).

Conflict of interest

JHu, LJ, JW, LG, and ZH were employed by Anhui iFLYTEK Health Co., Ltd.

The remaining 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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2023.1104190/full#supplementary-material

References

1. WHO. *World Health Organization: Depression*. Geneva: World Health Organization (2021).

2. Zhu M, Hong R, Yang T, Yang X, Wang X, Liu J, et al. The efficacy of measurement-based care for depressive disorders: systematic review and meta-analysis of randomized controlled trials. *J Clin Psychiatry*. (2021) 82:21r14034. doi: 10.4088/JCP.21r14034

3. Hamilton M. A rating scale for depression. *J Neurol Neurosurg Psychiatry*. (1960) 23:56–62. doi: 10.1136/jnnp.23.1.56

4. Worboys M. The Hamilton rating scale for depression: the making of a "gold standard" and the unmaking of a chronic illness, 1960-1980. *Chronic Illn*. (2013) 9:202–19. i: 10.1177/1742395312467658

5. Lam R, Michalaak E, Swinson R. *Assessment Scales in Depression, Mania and Anxiety: (Servier Edn)*. 1st ed. London: CRC Press (2004).

6. Williams J. A structured interview guide for the Hamilton depression rating scale. *Arch Gen Psychiatry*. (1988) 45:742–7. doi: 10.1001/archpsyc.1988.01800320058007

7. Otter D, Medina J, Kalita J. A survey of the usages of deep learning for natural language processing. *IEEE Trans Neural Netw Learn Syst*. (2021) 32:604–24. doi: 10.1109/TNNLS.2020.2979670

8. Wu S, Roberts K, Datta S, Du J, Ji Z, Si Y, et al. Deep learning in clinical natural language processing: a methodical review. *J Am Med Inform Assoc*. (2020) 27:457–70. doi: 10.1093/jamia/ocz200

9. Kodialam R, Boiarsky R, Lim J, Sai A, Dixit N, Sontag D. Deep contextual clinical prediction with reverse distillation. *Proc AAAI Conf Artif Intell*. (2021) 35:249–58. doi: 10.1609/aaai.v35i1.16099

10. Luo J, Ye M, Xiao C, Ma F. HiTANet: hierarchical time-aware attention networks for risk prediction on electronic health records. In: *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining. Association for Computing Machinery*. New York, NY (2020). p. 647–56. doi: 10.1145/3394486.3403107

11. Bernard J, Baddeley J, Rodriguez B, Burke P. Depression, language, and affect: an examination of the influence of baseline depression and affect induction on language. *J Lang Soc Psychol*. (2015) 35:317–26. doi: 10.1177/0261927X15589186

12. Panaite V, Devendorf A, Finch D, Bouayad L, Luther S, Schultz S. The value of extracting clinician-recorded affect for advancing clinical research on depression: proof-of-concept study applying natural language processing to electronic health records. *JMIR Form Res*. (2022) 6:e34436. doi: 10.2196/34436

13. Raymond C, Gregorius Satia B, Sandeep D, Fabian C. A textual-based featuring approach for depression detection using machine learning classifiers and social media texts. *Comput Biol Med*. (2021) 135:104499. doi: 10.1016/j.compbiomed.2021.104499

14. Hochreiter S, Schmidhuber J. Long short-term memory. *Neural Comput*. (1997) 9:1735–80. doi: 10.1162/neco.1997.9.8.1735

15. Jacob D, Ming-Wei C, Kenton L editors. BERT: pre-training of deep bidirectional transformers for language understanding. In: *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. (Minneapolis, MN: Association for Computational Linguistics) (2019).

16. Lin T, Goyal P, Girshick R, He K, Dollar P. Focal loss for dense object detection. *IEEE Trans Pattern Anal Mach Intell*. (2020) 42:318–27. doi: 10.1109/TPAMI.2018.2858826

17. Li B, Liu Y, Wang X editors. Gradient harmonized single-stage detector. In: *Proceedings of the AAAI Conferenceon Artificial Intelligence*. (Honolulu, HI: AAAI Press) (2019). doi: 10.1609/aaai.v33i01.33018577

18. Di Matteo D, Wang W, Fotinos K, Lokuge S, Yu J, Sternat T, et al. Smartphone-detected ambient speech and self-reported measures of anxiety and depression: exploratory observational study. *JMIR Form Res*. (2021) 5:e22723. doi: 10.2196/22723

19. Humphreys K, King L, Choi P, Gotlib I. Maternal depressive symptoms, self-focus, and caregiving behavior. *J Affect Disord*. (2018) 238:465–71. doi: 10.1016/j.jad.2018.05.072

20. Voleti R, Liss J, Berisha V. A review of automated speech and language features for assessment of cognitive and thought disorders. *IEEE J Sel Top Signal Process*. (2020) 14:282–98. doi: 10.1109/JSTSP.2019.2952087

21. Little B, Alshabrawy O, Stow D, Ferrier I, McNaney R, Jackson D, et al. Deep learning-based automated speech detection as a marker of social functioning in late-life depression. *Psychol Med*. (2021) 51:1441–50. doi: 10.1017/S0033291719003994

22. Wang J, Zhang L, Liu T, Pan W, Hu B, Zhu T. Acoustic differences between healthy and depressed people: a cross-situation study. *BMC Psychiatry*. (2019) 19:300. doi: 10.1186/s12888-019-2300-7

23. Mundt J, Vogel A, Feltner D, Lenderking W. Vocal acoustic biomarkers of depression severity and treatment response. *Biol Psychiatry*. (2012) 72:580–7. doi: 10.1016/j.biopsych.2012.03.015

24. Low D, Bentley K, Ghosh S. Automated assessment of psychiatric disorders using speech: a systematic review. *Laryngoscope Investig Otolaryngol*. (2020) 5:96–116. doi: 10.1002/lio2.354

25. Hecker P, Steckhan N, Eyben F, Schuller B, Arnrich B. Voice analysis for neurological disorder recognition-A systematic review and perspective on emerging trends. *Front Digit Health*. (2022) 4:842301. doi: 10.3389/fdgth.2022.842301

26. Eyben F, Scherer K, Schuller B, Sundberg J, Andre E, Busso C, et al. The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing. *IEEE Trans Affect Comput*. (2016) 7:190–202. doi: 10.1109/TAFFC.2015.2457417

27. Gavrilescu M, Vizireanu N. Predicting depression, anxiety, and stress levels from videos using the facial action coding system. *Sensors (Basel)*. (2019) 19:3693. doi: 10.3390/s19173693

28. Dibeklioglu H, Hammal Z, Cohn J. Dynamic multimodal measurement of depression severity using deep autoencoding. *IEEE J Biomed Health Inform*. (2018) 22:525–36. doi: 10.1109/JBHI.2017.2676878

frontiers | Frontiers in Psychiatry

TYPE Methods
PUBLISHED 02 March 2023
DOI 10.3389/fpsyt.2023.1139451

OPEN ACCESS

EDITED BY
Weikai Li,
Chongqing Jiaotong University, China

REVIEWED BY
Xiao Lin,
Peking University Sixth Hospital, China
Zhongchen Ma,
Jiangsu University, China

*CORRESPONDENCE
Li Zhang
✉ lizhang@njfu.edu.cn
Yonggui Yuan
✉ yygylh2000@sina.com
Daoqiang Zhang
✉ dqzhang@nuaa.edu.cn

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Computational Psychiatry, a section of the journal Frontiers in Psychiatry

RECEIVED 07 January 2023
ACCEPTED 06 February 2023
PUBLISHED 02 March 2023

CITATION
Zhang L, Pang M, Liu X, Hao X, Wang M, Xie C, Zhang Z, Yuan Y and Zhang D (2023) Incorporating multi-stage diagnosis status to mine associations between genetic risk variants and the multi-modality phenotype network in major depressive disorder.
Front. Psychiatry 14:1139451.
doi: 10.3389/fpsyt.2023.1139451

Incorporating multi-stage diagnosis status to mine associations between genetic risk variants and the multi-modality phenotype network in major depressive disorder

Li Zhang[1,2*†], Mengqian Pang[1†], Xiaoyun Liu[3†], Xiaoke Hao[4], Meiling Wang[2], Chunming Xie[5], Zhijun Zhang[5], Yonggui Yuan[3*] and Daoqiang Zhang[2*]

[1]College of Computer Science and Technology, Nanjing Forestry University, Nanjing, China, [2]College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China, [3]Department of Psychosomatic and Psychiatry, Zhongda Hospital, School of Medicine, Southeast University, Nanjing, China, [4]School of Artificial Intelligence, Hebei University of Technology, Tianjin, China, [5]Department of Neurology, ZhongDa Hospital, School of Medicine, Southeast University, Nanjing, China

Depression (major depressive disorder, MDD) is a common and serious medical illness. Globally, it is estimated that 5% of adults suffer from depression. Recently, imaging genetics receives growing attention and become a powerful strategy for discovering the associations between genetic variants (e.g., single-nucleotide polymorphisms, SNPs) and multi-modality brain imaging data. However, most of the existing MDD imaging genetic research studies conducted by clinicians usually utilize simple statistical analysis methods and only consider single-modality brain imaging, which are limited in the deeper discovery of the mechanistic understanding of MDD. It is therefore imperative to utilize a powerful and efficient technology to fully explore associations between genetic variants and multi-modality brain imaging. In this study, we developed a novel imaging genetic association framework to mine the multi-modality phenotype network between genetic risk variants and multi-stage diagnosis status. Specifically, the multi-modality phenotype network consists of voxel node features and connectivity edge features from structural magnetic resonance imaging (sMRI) and resting-state functional magnetic resonance imaging (rs-fMRI). Thereafter, an association model based on multi-task learning strategy was adopted to fully explore the relationship between the MDD risk SNP and the multi-modality phenotype network. The multi-stage diagnosis status was introduced to further mine the relation among the multiple modalities of different subjects. A multi-modality brain imaging data and genotype data were collected by us from two hospitals. The experimental results not only demonstrate the effectiveness of our proposed method but also identify some consistent and stable brain regions of interest (ROIs) biomarkers from the node and edge features of multi-modality phenotype network. Moreover, four new and potential risk SNPs associated with MDD were discovered.

KEYWORDS

imaging genetics, major depressive disorder, multi-modality, multi-stage diagnosis status, single-nucleotide polymorphisms

1. Introduction

Major depressive disorder (MDD) is a serious mental illness characterized by persistent sadness, lack of energy, loss of interest, sleep disturbance, and a high risk of suicide, with a high lifetime prevalence of MDD (16.2%) (1). It already affects about 350 million people worldwide and is expected to be the second most debilitating disease in the world by 2030 (2). Patients with major depression are often suicidal and have a risk of suicide that exceeds 10 times that of the general population (3). However, the etiology of MDD is still not clear. The pathogenesis varies for each individual, and an uniform pathogenesis has not been found in the literature. Currently, the diagnosis of depression mainly depends on patients clinical symptoms or the scores of Hamilton Depression Rating Scale (4). Due to human subjectivity, these methods cannot provide an objective diagnosis. Thus, effective prevention and early diagnosis are the important and urgent research topic for MDD.

With the rapid development of neuroimaging technology, magnetic resonance imaging (MRI), functional magnetic resonance imaging (fMRI), diffusion tensor imaging (DTI), and positron emission tomography (PET) are widely used for the prediction and diagnosis of mental illness and neurological disorders (5–7). Among them, resting-state fMRI (rs-fMRI) can reflect the neural activity of the brain and thus becomes the most popular brain imaging technology to discriminate MDD from healthy controls. Many studies have been made with the intent to explore the potential neuroimaging biomarkers for MDD diagnosis based on brain functional connectivity (FC). Yao et al. proposed a temporal adaptive graph conventional network(GCN), which not only took advantage of both spatial and temporal information using resting-state FC patterns and time-series but also explicitly characterized the subject-level specificity of FC patterns (8). Thereafter, Yao et al. made the full use of multiple indexes derived from rs-fMRI and designed a tensor-based multi-index representation learning framework for fMRI-based MDD prediction (9). Kong et al. first developed a spatio-temporal GCN framework to learn discriminative features from FC for automatic diagnosis and antidepressant treatment response prediction for MDD (10). Subsequently, this team successively proposed a novel multi-stage graph fusion networks framework by integrating the multiple stages of data representations, which fully considered the interactions between the white matter and the gray matter (11). Aforementioned diagnosis methods have the ability to capture dynamic FC alterations from rs-fMRI, which are considered as potential neuroimaging biomarkers to aid diagnosis and treatment for MDD. However, these methods focus on single imaging modality (e.g., rs-fMRI) and are limited in discovering the consistent biomarkers across multiple modalities.

In recent years, high-throughput genotyping technology coupled with brain neuroimaging provides a great promise to investigate the role of genetic variation on the brain structure and function and emerges as a new research field, namely, imaging genetics. The major task of this field is to measure the association between genetic variation (e.g., SNP) and neuroimaging biomarkers extracted from different imaging modalities. The obtained results may help us to deeper understand the complex pathogenesis of the diseases (12). Because Alzheimer's disease neuroimaging initiative database contains multi-modality imaging data and genotyping data, most imaging genetic studies of brain disorders focus on Alzheimer's disease. Researchers applied the multi-task learning framework to discover several common brain regions of interests (ROIs) which are associated with the well-known AD-risk SNP (APOE rs429358) and disease status by multi-modality imaging fusion technology (13). Some methods utilized canonical correlation analysis to measure the association between multiple genetic variations and neuroimaging data (14). Though numerous studies focused on imaging genetics for AD and yielded some interesting results to understand the pathogenic mechanisms of AD, few studies focused on the imaging genetic association of MDD by using machine learning technology. Only some studies from clinicians usually use simple statistical analysis methods, such as ANVOVA analysis, correlation analysis, and mediation analysis, to analyze the association between the given SNP and brain regions from neuroimaging for MDD (15, 16). Despite these methods being simple, quick and easy, and useful, the machine learning methods in the recent medical imaging studies receives unprecedented breakthroughs (5) and have the ability to automatically and fully explore the association between genetic variations and neuroimaging.

In this study, our goal was to develop a simple yet powerful model for the automatic discovery of the association between genetic risk factors and disease status by using multi-modality neuroimaging data. Thus, we designed a novel imaging genetic association framework to mine the multi-modality phenotype network between genetic risk factors and multi-stage diagnosis status. This framework can not only identify consistent ROIs from multi-modality imaging data but also search new risk SNPs associated with MDD. In detail, the proposed approach consists of two steps: (i) Constructing the multi-modality phenotype network. The network of each subject consists of voxel node features and connectivity edge features, which are extracted from sMRI and rs-fMRI, respectively. (ii) Building the imaging genetic association model. Compared with existing simple MDD imaging genetic methods, our proposed association model utilized multi-task learning to explore the relationship between the MDD-risk SNP and the multi-modality phenotype network. Thereafter, multi-stage diagnosis status are embedded into the association model by a novel regularization term to fully use the internal relation across different modalities of different subjects. All the data were collected from the Affiliated Zhongda Hospital of Southeast University and the Second Affiliated Hospital of Xinxiang Medical University. The experimental results show that our method can not only improve the performance on metrics of root mean squared error and correlation coefficient but also identify a compact set of common ROIs across two brain network features, which are closed related to MDD genetic risk SNP TPH1 rs1799913. Moreover, some new and potential risk SNPs associated with MDD are discovered.

The contributions of this study are listed as follows:

1. An imaging genetic association model is proposed to fully explore the relationship between a given MDD-risk SNP TPH1 rs1799913 and the multi-modality phenotype network, which is constructed by voxel node features and connectivity edge features from sMRI and rs-fMRI, respectively.
2. The multistage diagnosis status is introduced into the associated model by a novel regularization term, which brings the ability

TABLE 1 Demographic statics of subjects.

Hospital	ZhongDa			XinXiang		
subjects	HC	MD	SD	HC	MD	SD
Number	26	34	11	38	44	18
Gender(M/F)	10/16	15/19	4/7	21/17	26/18	7/11
Age	36.69 ± 13.88	11.09 ± 4.14	10.09 ± 5.99	10.74 ± 4.67	9.89 ± 4.25	9.44 ± 4.19
HAM-D24	1.27 ± 2.16	27.65 ± 4.68	38.91 ± 2.26	1.13 ± 1.85	29.91 ± 3.66	39.22 ± 3.95

HC, healthy control; MD, moderate depression; SD, severe depression; M/F, male/female; HAM-D24, Hamilton Depression Scale-24.

to fully use the relationship across the multiple modalities of different subjects.

3. From our sequencing SNP genotype data, some new and potential risk SNPs associated with MDD is discovered, which can help researchers to further investigate the pathogenesis of MDD.

The rest of the article is organized as follows: Section 2 shows the demographic statistics of participants and pre-processing of multi-modality data. Section 3 presents our proposed method to mine the multi-modality phenotype network between genetic risk factors and multi-stage diagnosis status. Presentation and analysis of experimental results are shown in Section 4. Finally, discussion and conclusion are given in Section 5 and 6, respectively.

2. Data source and preprocessing

2.1. Participants

This study-utilized datasets were obtained from two hospitals, namely, the Affiliated Zhongda Hospital of Southeast University and the Second Affiliated Hospital of Xinxiang Medical University. The patients were recruited through the inpatient and outpatient departments of psychiatry in the mentioned two hospitals, while the healthy controls (HCs) were recruited from media advertising and community posting. The whole research procedures adhered to the Declaration of Helsinki. All patients signed the informed consent document and met an identical inclusion criteria as follows: (1) They met the criteria listed in Diagnostic and Statistical Manual of Mental Disorder (Fourth Edition); (2) they were in the first depressive episode, and the age of onset was over 18 years old; (3) they got a Hamilton Depression Scale-24 (HAM-D24) scores $\geq$ 20; (4) the absence of other major psychiatric illness history; (5) the absence of primary neurodegenerative disorders, including dementia or stroke; (6) no substance abuse or dependence (drug, caffeine, nicotine, alcohol, or others), head trauma, or loss of consciousness; (7) no cardiac or pulmonary disease which could influence the MRI scan. HC subjects met the (4) – (7) rules of the above inclusion criteria and were required to get a HAM-D24 score $\leq$ 8.

After removing poor quality images due to head motion or ghost intensity, this study contained 26 HCs and 45 patients with MDD from the Affiliated ZhongDa Hospital of Southeast University, and 38 HCs and 62 patients with MDD from the Second Affiliated Hospital of XinXiang Medical University. Thereafter, we utilized HAM-D24 scores to assess depression severity. A total of 107 patients with MDD (HAM-D24 scores $\geq$ 20) can be further divided into two subgroups: the HAM-D24 score of 20–34 points was defined as moderate depression (MD) and a HAM-D24 score of $\geq$ 35 was defined as severe depression (SD) (17). The detailed demographic of subjects are shown in Table 1.

2.2. Magnetic resonance imaging (MRI) data acquisition and preprocessing

All participants underwent MRI scans at the baseline. MRI data were acquired using a 3.0 T Siemens scanner (Siemens, Erlangen, Germany) with a 12-channel head coil. The head of all subjects was immobilized with pads to minimize head movements. High-resolution 3D T1-weighted scan using a magnetization-prepared fast gradient echo (MPRAGE) sequence was performed according to the following parameters: repetition time (TR) = 1,900 ms, echo time (TE) = 2.48 ms, flip angle (FA) = 9°, acquisition matrix = 256 × 256, field of view (FOV) = 250 × 250 mm^2, thickness = 1.0 mm, gap = 0, time = 4 min 18 s, and volume = 176. rs-fMRI parameters were as follows: TR = 2,000, msTE = 25 ms, FA = 90°, acquisition matrix = 64 × 64, FOV = 240 × 240 ^{2}, thickness = 3.0 mm, gap = 0 mm, axial slice = 36, volume = 240; in-plane resolution parallel to the anterior-posterior conjunction = 3.75 × 3.75 ^{2}, and acquisition time = 8 min. During the scan, subjects were asked to lie on their backs with their hands naturally resting on their sides. All subjects' heads were held in place with pads to minimize head movement. Ear plugs were used to reduce the noise of the scanner. Subjects were asked to relax their bodies, open their eyes, stay awake, and not think about anything specific to avoid falling asleep. Images were checked immediately after scanning to ensure quality, and scans were repeated if necessary.

For quality control, all image data were examined by two experienced radiologists. The rs-fMRI images were preprocessed using the Resting State Functional Data Processing Assistant (DPARSF 2.3 Advanced) MRI toolkit, which combines the Resting State Functional MRI Toolkit (REST, http://www.restfmri.net) and the Statistical Parametric Mapping Package (SPM, https://www.fil.ion.ucl.ac.uk/spm/) programs (18). The first 10 time points were excluded to ensure stable longitudinal magnetization and to accommodate the inherent scanner noise. The remaining 230 images were processed sequentially according to the following steps: (1) correction for time differences and head motion using the 36th slice as the slice time of the reference slice (participants with maximum head motion displacement greater than 1.5 mm

in any direction [x, y, or z] or angular motion greater than 1.5° were excluded from the analysis); (2) T1 co-alignment with the functional image and subsequent reorientation; and (3) for spatial normalization, T1-weighted anatomical images were segmented into the white matter, the gray matter, and the cerebrospinal fluid and subsequently normalized to the Montreal Neurological Institute (MNI) space using transform parameters estimated by a uniform segmentation algorithm. These transform parameters were applied to the functional images, and the images were resampled with 3 mm isotropic voxels; (4) spatial smoothing was performed with a 4 mm full-width at half-peak (FWHM) isotropic Gaussian kernel; linear trends within each voxel time series were removed; interference signals (white matter, cerebrospinal fluid signals, and head motion parameters were calculated using rigid body six correction) and spiked regression volumes were regressed; and finally, a temporal bandpass (0.01–0.08 Hz) to minimize low-frequency drift and filter out high-frequency noise was performed.

2.3. SNP genotype data sequencing and processing

DNA genotyping was performed by Tianhao Biotechnology (Shanghai, China), and the standard protocol was employed to extract DNA from blood. The pre-designed Illumina next sequencing and array technology (Illumina Inc., San Diego, CA, USA) were utilized to determine the SNP genotypes in genes. Thereafter, we applied PLINK (v1.9) software to calculate the Hardy–Weinberg equilibrium (HWE) test, linkage disequilibrium statistics, and allele and genotype frequencies (19). After removing missing or incorrect values, the retained 5897 SNPs were used in this study.

Genetic risk variants can help researchers understand the relevant diseases of biological mechanism and provide an effective hypothesis for drug design. In this study, we focused on fully to explore the relationship between a given risk SNP and quantitative traits of the brain structure and functional level. Some researchers have implicated a large number of gene-related depression, including CACNA1E, BDNF, CRHR1, GSK3β, TPH1, and so on, see those in the systematic review (20). However, different from Alzheimer's disease with a well-known risk SNP APOE rs429358, there were no consistent gene hypotheses for the pathogenesis of MDD. SNPedia (http://www.SNPedia.com) is a wiki resource of the functional consequences of human genetic variation as published in peer-reviewed studies. We searched the related genetic-risk SNP associated with MDD in the SNPedia database. Among the retained 5897 SNPs, only TPH1 rs1799913 is successfully matched. Tryptophan hydroxylase (TPH) is the rate-limiting enzyme in the biosynthesis of serotonin (5-HT) and haplotype analysis indicates that TPH-1 associates with MD. In literature (21), six SNPs were found at linkage disequilibrium in both patients and control subjects, but only one SNP (rs1799913) significantly associated with MD by single marker association analysis. This SNP, also known in the literature as SNP A779C, has been associated with cerebrospinal fluid (CSF) 5-hydroxyindoleacetic acid (5-HIAA) concentrations, and several reports have shown its association with suicidal behavior (22, 23).

Thus, TPH1 rs1799913 was used in the following experiments to verify our proposed model. TPH1 rs1799913 value was coded in an additive fashion as 0, 1, and 2, where the alleles were divided into major and minor alleles by genotype frequency, and the major allele was coded as "0," the minor allele as "2," and the heterozygote of the remaining alleles as "1" (19).

3. Method

Recently, most of the neuroimaging genomic studies using machine learning technology focus on Alzheimer's disease, and few studies focus the pathogenesis of MDD. These methods usually research the association between genetic variants and structural imaging but ignore the functional connectivity information between various brain regions. As it is well-known, the human brain is a highly complex network, which contains higher-order connectivity information to help understanding the pathogenesis of diseases. Therefore, our model utilized the brain rs-fMRI and sMRI of each subjects as the input. The overview of our proposed model is shown as Figure 1. First, the network voxel node features and connectivity edge features were extracted from sMRI and rs-fMRI scans based on Automated Anatomical Labeling (AAL), respectively. After that, a multi-modality association model-introduced multi-stage diagnosis status of subjects was proposed to fully explore the relationship between two brain features and a given risk MDD SNP (TPH1 rs1799913) in high confidence.

3.1. Building the multi-modality phenotype network

In this study, each subject builds a multi-modality brain phenotype network. The node and edge features of brain network were extracted from sMRI and rs-fMRI, respectively.

After the preprocessing sMRI data, the voxel-based morphometry (VBM) of each subject were obtained from the normalized gray matter density maps, which were created in the MNI space as $2 \times 2 \times 2mm^3$ voxels. Thereafter, we aligned VBM to each participant's same visit scan and further extracted 116 ROI level measurements of mean gray matter densities based on the AAL template. In our proposed method, each ROI is modeled as a single node of multi-modality brain network. Thus, each subject could produce one set of node features.

For rs-fMRI data, a functional connectivity network usually is constructed for representing each subject, with each node denoting a pre-defined brain ROI and each edge representing the pairwise functional connection between ROIs. In this study, we extracted the mean time series of each ROI based on the AAL template and normalized them with zero mean and unit variance. After that, functional connectivity networks were generated by using the Pearson correlation coefficient and hence could capture the correlation between the BOLD signals of paired ROIs.

Due to the psychiatric disorders with abnormal topological properties of brain networks, many graph theory-based methods have played an important role in the human brain disorder

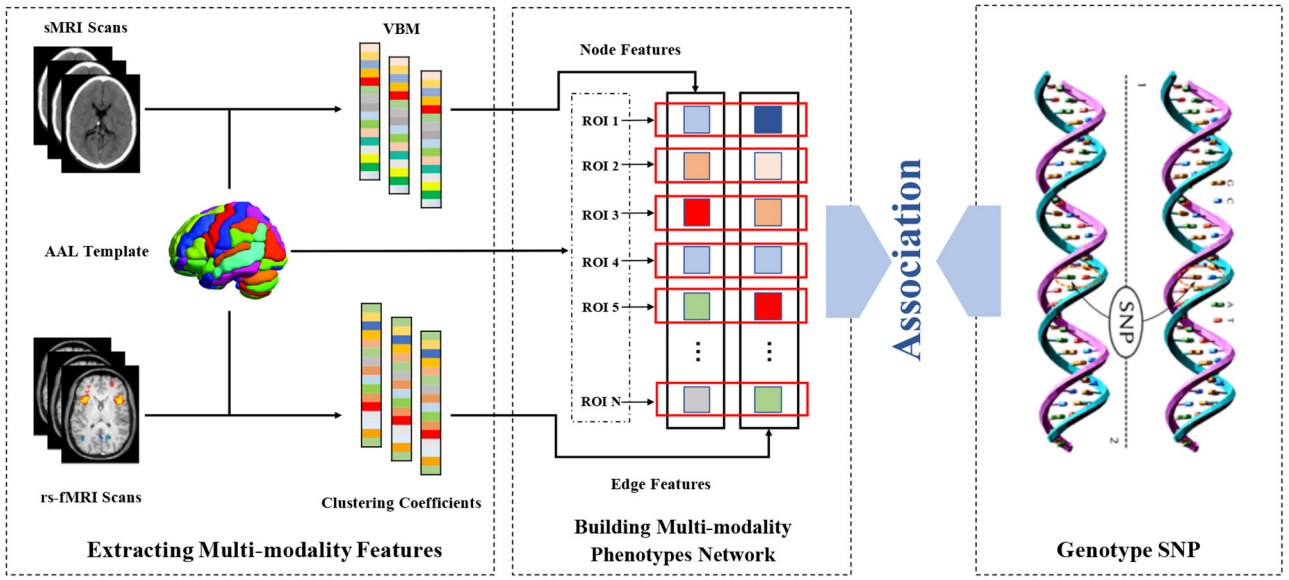


FIGURE 1
The overview of our proposed model.

analysis. Among them, clustering coefficient (CC) is one of the most popular methods and reflects local clustering properties of the brain network (24). Therefore, after constructing the functional connectivity network, the clustering coefficients can be calculated by

$$CC^W(i) = \frac{2}{k_i(k_i-1)} \sum_{j,h} (\overline{w_{ij}} \cdot \overline{w_{ih}} \cdot \overline{w_{jh}}), \quad (1)$$

where w_{ij}, w_{ih}, and w_{jh} are the connection weights between node i and j, between node i and h, and between node j and h, respectively. Scaled for weights $\overline{w_{ij}}$, $\overline{w_{ih}}$, and $\overline{w_{jh}}$, that is, $\overline{w_{ij}} \longleftarrow w_{ij}/max(w)$, $\overline{w_{ih}} \longleftarrow w_{ih}/max(w)$, $\overline{w_{jh}} \longleftarrow w_{jh}/max(w)$, $max(w)$ denotes the maximum connection weight in the brain network. The number of edges connected to node i is denoted by k_i. Finally, we extracted a set of clustering coefficients from the functional connectivity networks as the connectivity edge furthers of the multi-modality phenotype network. Thus, each subject could produce one set of edge features.

3.2. Associations between genotype and multi-modality phenotype network

After building the multi-modality phenotype network of each subject, pathological alterations are considered to be abnormal changes in phenotype networks. Since the network node and edge features consist of VBM from sMRI and clustering coefficients from rs-fMRI, abnormal changes are closely related to associated ROIs and significant connectivity edges. In this study, we suppose that there are N subjects, with each one represented by a multi-modality phenotype network. Given M modalities of phenotypes $X^m = [X_1^m, ..., X_n^m, ..., X_N^m]^T \in R^{N\times d}$ as the input and the corresponding response value() $y = [y_1, ..., y_n, ..., y_N]^T \in R^N$ as the output, where d is the number of node and edge features dimensionality. Let $w^m \in R^d$ denote the linear discriminant function corresponding to the m-th modality. Then the multi-modality network phenotype association model can be formulated as:

$$\min_W \frac{1}{2} \sum_{m=1}^{M} \left\| y - X^m w^m \right\|_2^2 + \lambda \left\| W \right\|_{2,1}, \quad (2)$$

where $W = [w^1, w^2, ..., w^M] \in R^{d\times M}$ is the weight matrix, and each row w_j represents the vector of coefficients assigned to the j-th features across multiple modalities. It is worth noting that Equation (2) adds the $L_{2,1}$-norm regularization term, $\left\| W \right\|_{2,1} = \sum_{j=1}^{d} \left\| w_j \right\|_2$, which is a "group-sparsity" regularizer and penalizes all coefficients in the same row of matrix W for joint feature selection. This means that our proposed modal can force only a small number of features to be selected across multiple modalities. The regularization parameter λ is used to balance the contributions of two terms in Equation (2). The larger the value of λ, the fewer the features are selected.

3.3. Introduction multi-stage diagnosis status into the association model

One disadvantage of the aforementioned multi-modality association model is that it only considers the relationship between the multiple modalities of the same subject and ignores the relationship between imaging phenotypes and diagnosis status among subjects. In order to overcome this limitation, we made full use of the multi-stage diagnosis status of each subjects, i.e., HC, SD, and MD, and then introduced a novel regularization term that can embed the multi-stage diagnosis status

TABLE 2 Detailed descriptions of various comparisons.

Comparison	Modality	Diagnosis	Description
SM	Node	No	Using Lasso (Least absolute shrinkage and selection operator) to detect a sparse significant subset from node or edge features.
SM	Edge	No	
MSD-SM	Node	Yes	
MSD-SM	Edge	Yes	
CM	-	No	Concatenating node and edge features, and then using Lasso to detect a sparse significant subset from combined features.
MSD-CM	-	Yes	
MM	Node	No	Detecting a sparse subset of common ROIs from node and edge features.
MM	Edge	No	
MSD-MM	Node	Yes	
MSD-MM	Edge	Yes	

of MDD:

$$\sum_{i,j}^{N} \left\| (w^m)^T x_i^m - (w^m)^T x_j^m \right\|_2^2 S_{ij}^m = 2(w^m)^T (X^m)^T L^m X^m w^m, \quad (3)$$

where $L^m = D^m - S^m$ denotes a Laplacian matrix for the m-th modality, where D^m is the diagonal matrix and each element is defined by $D_{ij}^m = \sum_{j=1}^{N} S_{ij}^m$. In this study, $S = [S_{ij}^m] \in R^{n \times n \times m}$ represents a similarity matrix which measures the similarity between each pair of subjects on the m-th modality and can be defined as

$$S_{ij}^m = \begin{cases} 1, & \text{if } x_i^m \text{ and } x_j^m \text{ are from the same class in } m\text{-th modality} \\ 0, & \text{otherwise.} \end{cases} \quad (4)$$

The similarity between subjects within the same class and modality can be defined as 1; otherwise, it is 0. The purpose of Equation (3) is to enforce these subjects from the same class and modality to be close to each other in the label space. When x_i^m and x_j^m are from the same class in the m-th modality, the distance between $(w^m)^T x_i^m$ and $(w^m)^T x_j^m$ should be as small as possible in the label space.

In this study, we introduced the multi-stage diagnosis status into the multimodality association model by incorporating the regularizer (3) into Equation (2). The objective function of our proposed association model (MSD-MM) can be formulated as follows:

$$\min_W \frac{1}{2} \sum_{m=1}^{M} \left\| y - X^m w^m \right\|_2^2 + \lambda_1 \left\| W \right\|_{2,1} + \lambda_2 \sum_{m=1}^{M} (w^m)^T (X^m)^T L^m X^m w^m, \quad (5)$$

where parameter λ_1 and λ_2 are used to control two regularization terms, respectively. Their values can be determined by inner cross-validation on training data. From the objective function Equation (6), the MSD-MM model can not only jointly select a sparse subset of common features from multi-modality data but also fully use the prior diagnosis information among subjects. To efficiently solve the objective function in Equation (6), we used the Nesterov's accelerated proximal gradient optimization algorithm (25).

4. Experimental results and analysis

4.1. Experimental settings

In our experiments, we adopted two evaluation metrics, i.e., root mean squared error (RMSE) and correlation coefficient (CC), which are wildly used to measure performance regression and association analysis between the predicted and actual response values, respectively.

The five-fold cross validation strategy is implemented to validate the effectiveness of our proposed method. For the parameters λ_1 and λ_2 of regularization in Equation (5), we tuned them from $10^{-5}, 3 \times 10^{-5}, 10^{-4}, 3 \times 10^{-4}, ..., 3$ and determined their values by the nested five-fold cross validation on the training dataset.

In this study, we compared the single-modality (SM) method, concatenate-modality (CM) method, and multimodality (MM) method with/without multi-stage diagnosis status. In Table 2, SM, CM, and MM are conventional methods without diagnosis status. Since one contribution of this study is the introduction of diagnosis status, which provide more prior knowledge, MSD-SM and MSD-CM are improved SM and CM methods incorporating multi-stage diagnosis status, respectively. MSD-MM is our proposed method, which simultaneously considers the multi-modality images and multi-stage diagnosis status. The detailed description of various comparisons are shown in Table 2.

4.2. Association between risk SNP and the multi-modality network phenotype

We compared our proposed MSD-MM method with conventional methods without diagnosis status (including SM, CM, and MM) and improved methods with diagnosis status (including MSD-SM and MSD-CM). In order to eliminate the bias of random division, we performed five times independent and non-repetitive five-fold cross validation. Thereafter, the average results of RSEM and CC on the training and testing data on node and edge modalities were calculated, respectively, as shown in Table 3.

TABLE 3 Comparison of regression performance on risk SNP TPH1 rs1799913 by different methods.

Method		RMSE (mean ± std)		CC (mean ± std)	
		Train	Test	Train	Test
SM	Node	1.3107 ± 0.1007	1.3644 ± 0.1654	0.0465 ± 0.0440	0.0121 ± 0.0045
SM	Edge	1.3454 ± 0.2189	1.5456 ± 0.1566	0.0983 ± 0.0564	0.0488 ± 0.0027
MSD-SM	Node	0.7030 ± 0.0564	1.0114 ± 0.1291	0.4340 ± 0.0617	0.2164 ± 0.0945
MSD-SM	Edge	1.1610 ± 0.0790	1.3726 ± 0.2136	0.3029 ± 0.0594	0.2196 ± 0.1448
CM	-	1.3988 ± 0.0949	1.4245 ± 0.1314	0.0984 ± 0.0654	0.0745 ± 0.0915
MSD-CM	-	0.8736 ± 0.0654	1.3034 ± 0.1195	0.4021 ± 0.0390	0.2254 ± 0.0966
MM	Node	0.6144 ± 0.1647	1.1348 ± 0.1649	0.5416 ± 0.0412	0.1654 ± 0.0925
MM	Edge	0.8117 ± 0.1054	1.2654 ± 0.1267	0.4645 ± 0.0561	0.1714 ± 0.0729
MSD-MM	Node	0.6887 ± 0.0653	0.9685 ± 0.1308	0.4971 ± 0.0712	**0.2432±0.0799**
MSD-MM	Edge	0.8310 ± 0.1410	1.1892 ± 0.2271	0.3625 ± 0.0254	**0.2697±0.0910**

This means that two results of the bold vales are significantly better than the above results.

As shown in Table 3, MSD-SM receives the RMSE values of 1.0114 and 1.3726 and the CC values of 0.2164 and 0.2196 on the node and edge features, respectively, achieving better performance than the conventional SM method. MSD-CM obtains the RMSE value of 1.3034 and the CC value of 0.2254, which are better than those of the CM method. MSD-MM yields the best RMSE values of 0.9685 and 1.1892 and the CC values of 0.2432 and 0.2697 on the two different features. These results indicate three findings: 1) Compared with the SM-type method, the MM-type method can jointly select node and edge features and significantly improve the performance of regression and association analysis; 2) after introducing multistage diagnosis information, the proposed MSD-type method consistently outperform their conventional methods in both RMSE and CC performance measurements; 3) compared with voxel-based morphometry node features, the functional connectivity edge features between different brain regions provide more insights for the mechanistic understanding of MDD; moreover, CM-type and MM-type methods both utilize node and edge features, but they apply distinct strategies to combine two features and receive different performance. The CM-type method directly concatenate node and edge features. This way may be lost when the relationship information of two modalities and bring more noise in widespread feature space, while the MM-type method uses the multi-task strategy ($L_{2,1}$-norm constrain) to jointly select node and edge features, which can improve the robustness of ROIs detection. In a word, our proposed MSD-MM method yields the best performance on RMSE and CC measures. That demonstrates that simultaneously considering multi-modality imaging data and multi-stage diagnosis status can improve the performances of regression and association analysis between the imaging phenotype and genotype.

4.3. Identification of the related node ROI markers from sMRI data

In addition to improving the measure performances of regression and association analysis, one major purpose of this study is to identify some significant imaging phenotypes, which are highly associated with both risk SNP and multi-stage diagnosis status. In this study, due to the use of multi-modality imaging data, these identified phenotypes offer the possibility of detecting associations between the genotype and the brain structure as well as the function and help researchers further understand the pathogenesis of MDD.

TABLE 4 Top 10 ROIs selected by the node features from sMRI data.

ID	ROI name	Weight
38	Hippocampus.R	6.78
85	Temporal Mid.L	3.73
92	Cerebelum Crus1.R	3.32
87	Temporal Pole Mid.L	3.05
68	Precuneus.R	2.98
10	Frontal Mid Orb.R	2.78
78	Thalamus.R	2.64
49	Occipital Sup.L	2.03
79	Heschl.L	1.90
12	Frontal Inf Oper.R	1.70

For node features from sMRI data, we averaged the weight values by five times repeated five-fold cross validations and selected the top 10 maximum weight ROIs as the significant ROI markers. Table 4 presents the top 10 selected ROIs from sMRI data and the corresponding average weight values. Next, these weight values were mapped onto the human brain, and Figure 2 shows the visualization of the top 10 selected ROIs in all three planes (sagittal, coronal, and axial). In each plane, the colors of the labeled brain regions reflect the average weight values of the corresponding selected ROIs. The remarkable thing is that most of the selected ROIs are consistent with earlier findings, which focus on structural images and have already identified several diagnostic markers of MDD. The literature (26) indicates that patients with MDD exhibit bilateral volume reduction in major hippocampal substructures and identify core hippocampal

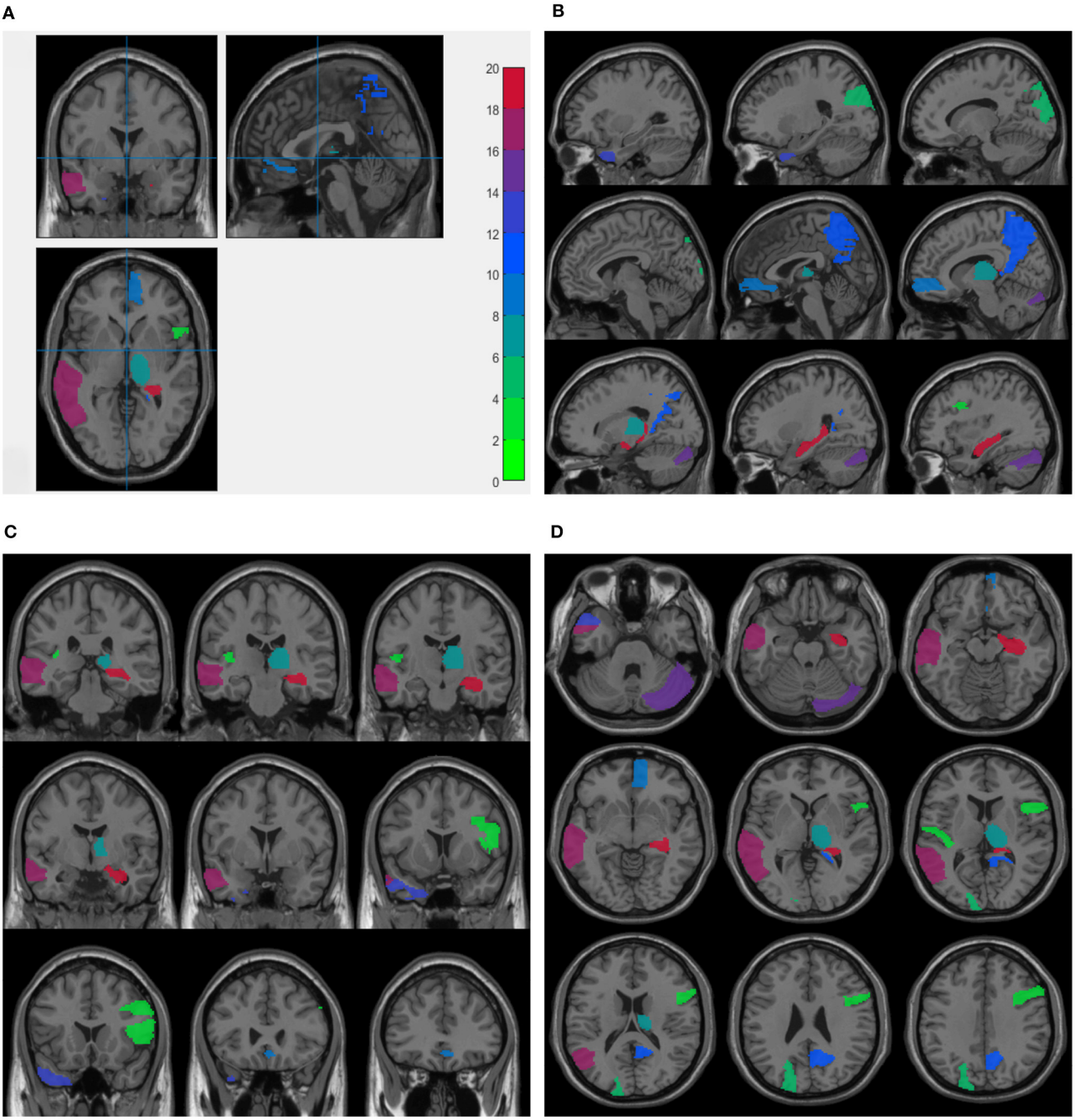


FIGURE 2
Visualization of the top 10 ROIs selected by node features in three planes: Three planes of sMRI scan **(A)**, sagittal plane **(B)**, coronal plane **(C)**, and axial plane **(D)**. The colors represents the average weight values of the corresponding ROI markers. These figures are plotted.

regions in MDD pathology as a potential marker of disease progression in MDD. Moreover, researchers find that reduced hippocampal gray matter volume is a common feature of patients with MDD (27). The bilateral middle frontal gyrus shows that the amplitude of low-frequency fluctuation (ALFF) significantly increases in subjects with subclinical depression (16, 28). Structural abnormalities in the thalamus might be the potential trait marker of MDD at the early stage as in (29, 30). Compared with healthy controls, patients with MDD presents decreased gray matter density in the bilateral temporal pole and right superior temporal gyrus (31).

4.4. Identification of the related edge ROIs marker from rs-fMRI data

The brain system can be simply represented by a brain network model, whose nodes and edges are defined as brain regions and

TABLE 5 Top 10 ROIs selected by the edge features from rs-fMRI data.

ID	ROI name	Weight
85	Temporal Mid.L	8.47
32	Cingulum Ant.R	8.05
89	Temporal Inf.L	6.78
68	Precuneus.R	6.52
54	Occipital Inf.R	5.59
38	Hippocampus.R	2.46
33	Cingulum Mid.L	1.95
35	Cingulum Post.L	1.69
41	Amygdala.L	1.60
30	Insula.R	1.59

connections between brain regions, respectively. In this study, rs-fMRI data are used to construct the brain functional connectivity network. After that, the clustering coefficients are extracted as edge features and are absorbed into each brain region. Thus, the dimension of the obtained edge features is the same as the number of the brain regions, and each dimension corresponds to one brain region. Thus, we can identify the related ROI markers from rs-fMRI data by our proposed method.

For edge features from rs-fMRI data, we also averaged the weight values by five times repeated five-fold cross validations and selected the top 10 maximum weight ROIs as the significant ROI markers, as shown in Table 5. Some existing studies using rs-fMRI data have identified several diagnostic brain region markers for MDD, such as precuneus, cingulum and temporal, which are also consistent with the aforementioned selected relevant ROIs. The study (32) shows that guilt-selective to MDD is associated with functional disconnection of anterior temporal and subgenual cortices. Patients with MDD have an abnormal activity of the precuneus at the resting state in first-episode drug-naive. This indicates that activity within the precuneus may be a potential biomarker for the diagnosis of MDD (33). In literature (34), the increased fractional amplitude of low-frequency fluctuation (fALFF) in the left mid cingulum, right precuneus, and left superior frontal gyrus may serve as a neuroimaging marker for first-episode MDD. The combined use of the increased fALFF in the right precuneus and left superior frontal gyrus obtains the best diagnostic scores. Compared to patients with remitted MDD and to healthy controls, patients with recurrent MDD exhibit decreased fALFF in the right posterior insula and right precuneus and increased fALFF in the left ventral anterior cingulate cortex (35).

Moreover, in order to analyze the functional connectivity of selected brain regions and graphically compare the difference on the functional connectivity network between patients with MDD and HCs, we selected the maximum weight ROI (left temporal) and the minimum weight ROI (right Insula) in Table 4 and then calculated the average edge values of functional connectivity network for MDD and HC group, respectively. Specifically, we first constructed the functional connectivity networks of each subject in MDD and HC group. After that, the average functional connectivity networks were calculated for each group. Finally, we selected seven edges with the highest connection values from all edges of a given brain region (14). Figure 3 graphically presents the top seven average connection value edges on maximum and minimum weight ROIs. As seen in Figure 3, compared to the HC group, the edges of maximum weight ROI in the MDD group have a distinct change, which is that the edge connecting right temporal middle and right frontal superior medial is replaced to the edge connecting right frontal superior medial and left temporal pole superior. However, the edges of the minimum weight ROI in the MDD group are same as those in the HC group. The aforementioned results it demonstrate that the identified significant brain regions (the maximum weight ROI) are highly correlated to the pathogenesis of MDD.

4.5. Identification of the consistent ROIs from multi-modality imaging data

In addition, for identifying some significant ROIs from node and edge features, another advantage of our proposed method is to ensure that these identified brain regions are consistent in both node and edge features. Figure 4 shows all comparisons of weight maps for multi-modality data on 116 ROIs associated with risk SNP TPH1 rs11179027. As seen in Figure 4, SM-based and CM-based methods select a large number of ROIs, while these ROIs are inconsistent across node and edge features. This means that researchers find it hard to use the selected ROIs for further investigation. However, the MSD-MM method is able to identify sparse and consistent ROIs associated with THP1 rs11179027 using multi-modality imaging data. These identified ROIs, such as left temporal, right hippocampus, and right precuneus parts, strongly agree with the existing studies and are highly correlated with MDD (26, 27, 33, 37, 38). To sum up, our proposed method tends to select consistent ROIs associated with risk SNP across multi-modality imaging data, which show a great value to further investigate the mechanism of MDD.

4.6. Identification of new risk SNPs associated with MDD

As is well-known, the pathogenesis of MDD may be caused by a large number of genetic risk SNPs. However, we only explored the relationship between two brain network features and a given MDD-risk SNP, THP1 rs1799913 in the aforementioned experiments and then identified some significant brain ROIs associated with this SNP. Results demonstrate that our proposed method can be used as an effective tool to mine new risk SNPs associated with MDD. In this study, the genotype data contains 5879 SNPs. In this study, we performed the MSD-MM method for each SNP in the whole genotype data. Table 6 shows that four SNPs have similar correlation coefficient with THP1 rs1799913 on node and edge features, respectively. This means that four SNPs may be risk genetic SNP associated with MDD. The literature (39) presents that PIK3R1 rs3730089 is related to schizophrenia and

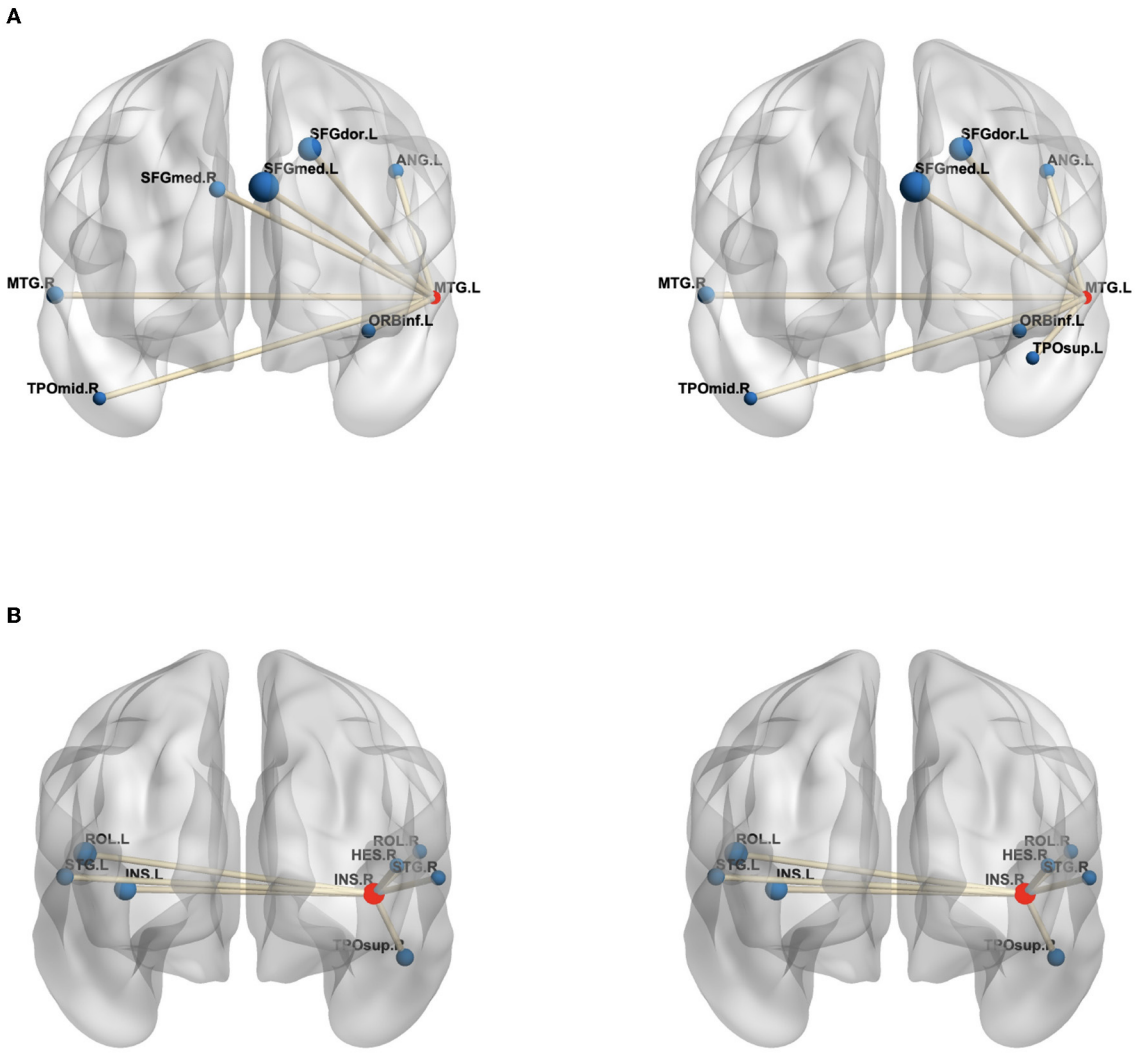


FIGURE 3
The edges of the maximum weight ROI **(A)** and minimum weight ROI **(B)** on the MDD (left) and HC (right) group. The centroid red node represents the selected ROI, and the blue node denotes the corresponding ROI linked by top seven average connection value edges. **(A)** MTG.L, Temporal Mid gyrus.L; MTG.R, Temporal Mid.R; ANG.L, Angular.L; SFGdor.L, Frontal Sup.L; SFGmed.L, Frontal Sup Medial.L; SFGmed.R, Frontal Sup Medial.R; ORBinf.L, Frontal Inf Orb.L; TPOmid.R, Temporal Pole Mid.R; TPOsup.L, Temporal Pole Sup.L; **(B)** INS.R, Insula.R; INS.L, Insula.L; HES.R, Heschl.R; ROL.L, Rolandic Oper.L; ROL.R, Rolandic Oper.R; STG.L, Temporal Sup.L; STG.R, Temporal Sup.R; TPOsup.R, Temporal Pole Sup.R. In the figure, all edges of each brain figure are plotted by BrainNet (36).

bipolar disorder in the Han Chinese population, but patients with schizophrenia, bipolar disorder, and MDD usually show common symptoms, such as anhedonia and amotivation. Recent research studies systematically report that these three mental disorders have a familial clustering character and any two or even three of these disorders could co-exist in some families. In addition, evidence from symptomatology and psychopharmacology also imply that there are intrinsic connections between these three mental disorders (40). Thus, PIK3R1 rs3730089 may be a risk SNP marker in MDD research. Meanwhile, although KDSR rs1138488, LAMA2 rs2229848, and THY1 rs3138094 receive fine correlation coefficients by the MSD-MM method, now there are no medical or biological research studies that directly support three SNPs related to MDD, and we expect that to be verified in future studies. We hope that these new MDD-risk SNPs will be verified in future studies and give more insights to understand the pathogenesis of MDD.

5. Discussion

5.1. Risk MDD SNP vs. non-disease related SNP

In the SNPedia database, TPH1 rs1799913 has been confirmed as a genetic risk variant associated with MDD. We evaluated the performance of regression and association analysis on the risk SNP TPH1 rs1799913 by the MSD-MM method in the aforementioned experiments, as shown in Table 3. To verify that the improvement is brought by only using the risk MDD SNP, we further selected

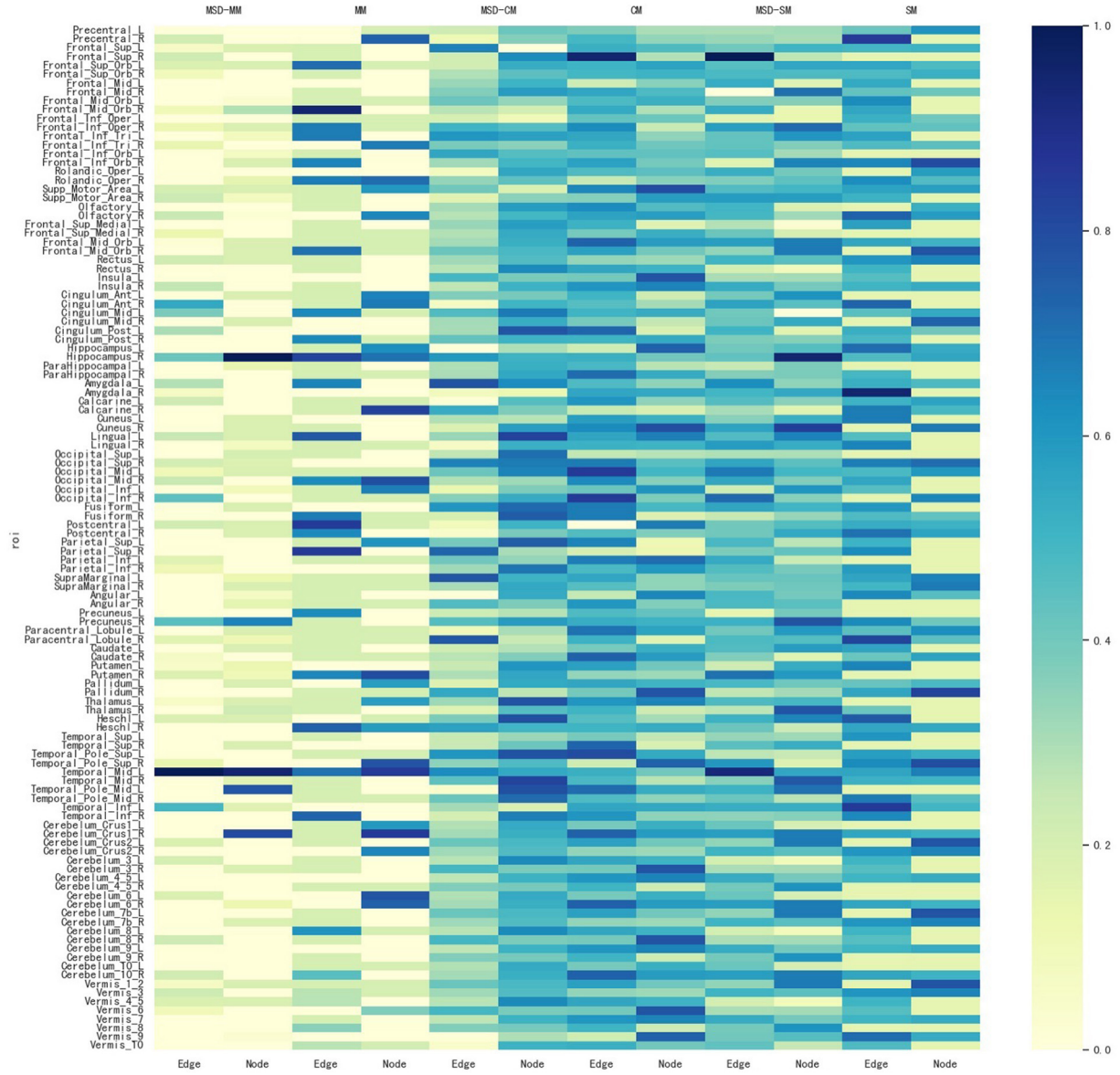


FIGURE 4
Weight maps for the multi-modality imaging data on 116 ROI associated with SNP THP1 rs1799913 respect to different methods.

three non-MDD related SNP for comparison. TPH2 rs11179027 is the nearest SNP to TPH1 rs1799913 (41), APOE rs429358 is a well-known top risk SNP associated with AD (42), and AKT3 rs14403 is a genetic risk SNP for schizophrenia (43). The MSD-MM method evaluated correlation coefficients of three non-MDD related SNPs, and results of all comparisons on the test dataset are shown in Table 7. Compared to the risk SNP THP1 rs1799913 reported in Table 3, all comparisons (including MSD-MM method) receive very low correlation coefficient values. The reason of the poor performance is that the train data containing the non-related MDD SNP lead to an overfitting problem for all comparisons and then lose the power of generalization on the test data. Thus, the contrast experiment shows that the learned consistent multi-modality imaging phenotypes if and only if using the risk MDD SNP can discovery the potential biological pathway from gene to the brain for clinical diagnosis.

TABLE 6 Correlation coefficient of four new SNPs on node and edge features by the MSD-MM method.

SNPs	Edge	Node
rs3730089 (PIK3R1)	0.2033 ± 0.1300	0.2118 ± 0.0355
rs1138488 (KDSR)	0.2032 ± 0.0945	0.2209 ± 0.0245
rs2229848 (LAMA2)	0.2143 ± 0.1003	0.2170 ± 0.1370
rs3138094 (THY1)	0.2248 ± 0.1233	0.2235 ± 0.1281

5.2. The selection of regularization parameters

Our proposed method, MSD-MM, contains two regularization parameters, the sparsity parameter λ_1 and the multi-stage diagnosis information parameter λ_2. The two parameters was used to balance

the relative contributions of three terms in Equation (5). In order to research the effect of two regularization terms on the performance of our proposed method, we set the values of two parameters in the range of $10^{-4}, 3 \times 10^{-4}, 10^{-3}, 3 \times 10^{-3}, ..., 1, 3$, respectively. Figure 5 shows the heat maps of correlation coefficients between parameters λ_1 and λ_2 on the test data. As shown in Figure 5, the MSD-MM method achieved the competitive or better performance than the MM method (reported in Table 3) on all combinations of parameter values, which further indicates the advantages of the multi-stage diagnosis information regularization term. Meanwhile, the areas were bounded by $\lambda_1 < 0.1$ and $0.01 < \lambda_2 < 0.3$ consistently and obviously outperformed than the MM method on node and edge futures. This area was helpful to quickly select the optimal the values of two parameters in future research.

TABLE 7 Correlation coefficients of all comparisons for three non-MDD related SNPs on the test data.

Method		rs11179027	rs429358	rs14403
SM	Node	0.0102 ± 0.0951	−0.0239 ± 0.0236	0.0933 ± 0.0320
SM	Edge	−0.0323 ± 0.1005	0.0482 ± 0.0319	0.0209 ± 0.0334
MSD-SM	Node	0.0290 ± 0.1118	−0.0254 ± 0.0282	0.0089 ± 0.0101
MSD-SM	Edge	−0.0199 ± 0.1366	0.0358 ± 0.0234	0.0357 ± 0.0176
CM	-	0.0156 ± 0.0991	0.0431 ± 0.0341	0.0298 ± 0.0425
MSD-CM	-	0.0095 ± 0.1668	0.0525 ± 0.1934	0.0290 ± 0.2443
MM	Node	0.0346 ± 0.0160	-0.0565 ± 0.0349	0.0233 ± 0.0218
MM	Edge	−0.0211 ± 0.1519	0.0323 ± 0.1843	0.0179 ± 0.0024
MSD-MM	Node	0.0344 ± 0.1274	−0.0887 ± 0.0034	0.0445 ± 0.0367
MSD-MM	Edge	−0.0104 ± 0.1351	0.0522 ± 0.0337	0.0603 ± 0.0285

6. Conclusion

In summary, this study developed a novel imaging genetic association framework to mine the multi-modality phenotype network between a single genetic risk SNP and multi-stage diagnosis status. First, the multimodality phenotype network is constructed by the voxel node features and connectivity edge features from sMRI and rs-fMRI, respectively. After that, an association model incorporated multi-stage diagnosis status is used to fully explore the relationship between the MDD-risk SNP TPH1 rs1799913 and the multi-modality phenotype network. All participants were recruited from two hospitals, and each participant contains sMRI, rs-fMRI, and genotype data. The detailed experimental results show that our proposed method can improve the performance on the metrics of root mean squared error and correlation coefficient compared with other comparisons. Some consistent and stable ROIs biomarkers are identified from voxel node features and connectivity edge features of multi-modality phenotype network. Moreover, an interesting finding is that four new and risk SNPs were discovered highly associated with MDD.

This study is an initial attempt to explore the relationship between a single genetic MDD-risk SNP and multi-modality brain neuroimaging data (sMRI and rs-fMRI). In future, we further investigate the use of other modality brain imaging (e.g., DTI) to directly construct the multi-modality graphical phenotype network

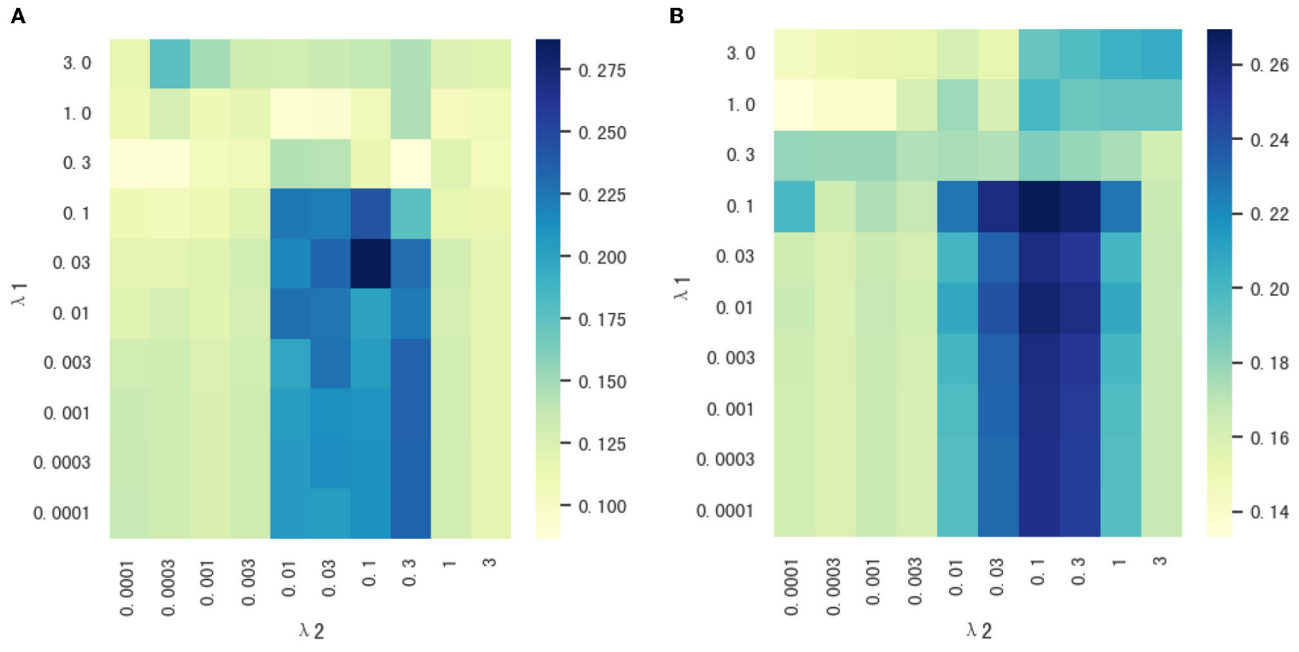


FIGURE 5
(A, B) The heat maps of correlation coefficients between parameters λ_1 and λ_2 on node and edge features.

and mine relationship between multi-modality phenotype network and multi-locus risk SNPs. We hope that more meaningful results are discovered to deeper understand the pathogenesis of MDD and help the diagnosis and treatment of patients with MDD.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: https://doi.org/10.6084/m9.figshare.22015541.v1.

Author contributions

LZ, YY, and DZ conceived and designed this article. LZ and MP proposed the association model, performed the experimental analysis, created figures and tables, and wrote this manuscripts. XL, CX, and ZZ collected the data set and organized it for analyses. MP trained the model and analyzed the experimental results with the help of XH and MW. All authors read, edited, and discussed the article. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the National Natural Science Foundation of China (Nos. 62276088, 62106104, 62136004, 61802193, and 81971277) and also by the Key Research and Development Plan of Jiangsu Province (No. BE2022842).

Conflict of interest

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.

Publisher's note

References

1. Otte C, Gold SM, Penninx BW, Pariante CM, Etkin A, Fava M, et al. Major depressive disorder. *Nat Rev Dis Primers.* (2016) 2:1–20. doi: 10.1038/nrdp.2016.65

2. Bains N, Abdijadid S. Major depressive disorder. In: *StatPearls*. Treasure Island, FL: StatPearls Publishing (2022).

3. Orsolini L, Latini R, Pompili M, Serafini G, Volpe U, Vellante F, et al. Understanding the complex of suicide in depression: from research to clinics. *Psychiatry Investig.* (2020) 17:207. doi: 10.30773/pi.2019.0171

4. Kennedy SH. Core symptoms of major depressive disorder: relevance to diagnosis and treatment. *Dialogues Clin Neurosci.* (2008) 10:271–7. doi: 10.31887/DCNS.2008.10.3/shkennedy

5. Zhang L, Wang M, Liu M, Zhang D. A survey on deep learning for neuroimaging-based brain disorder analysis. *Front Neurosci.* (2020) 14:779. doi: 10.3389/fnins.2020.00779

6. Li W, Xu X, Wang Z, Peng L, Wang P, Gao X. Multiple connection pattern combination from single-mode data for mild cognitive impairment identification. *Front Cell Dev Biol.* (2021) 9:782727. doi: 10.3389/fcell.2021.782727

7. Li WK, Chen YC, Xu XW, Wang X, Gao X. Human-guided functional connectivity network estimation for chronic tinnitus identification: a modularity view. *IEEE J Biomed Health Inform.* (2022) 26:4849–58. doi: 10.1109/JBHI.2022.3190277

8. Yao D, Sui J, Yang E, Yap PT, Shen D, Liu M. Temporal-adaptive graph convolutional network for automated identification of major depressive disorder using resting-state fMRI. In: *International Workshop on Machine Learning in Medical Imaging*. (Lima: Springer) (2020). p. 1–10.

9. Yao D, Yang E, Guan H, Sui J, Zhang Z, Liu M. Tensor-based multi-index representation learning for major depression disorder detection with resting-state fMRI. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Strasbourg: Springer (2021). p. 174–84.

10. Kong Y, Gao S, Yue Y, Hou Z, Shu H, Xie C, et al. Spatio-temporal graph convolutional network for diagnosis and treatment response prediction of major depressive disorder from functional connectivity. *Hum Brain Mapp.* (2021) 42:3922–33. doi: 10.1002/hbm.25529

11. Kong Y, Niu S, Gao H, Yue Y, Shu H, Xie C, et al. Multi-stage graph fusion networks for major depressive disorder diagnosis. *IEEE Trans Affect Comput.* (2022) 13:1917–28. doi: 10.1109/TAFFC.2022.3205652

12. Shen L, Thompson PM. Brain imaging genomics: integrated analysis and machine learning. *Proc IEEE.* (2019) 108:125–62. doi: 10.1109/JPROC.2019.2947272

13. Hao X, Yao X, Yan J, Risacher SL, Saykin AJ, Zhang D, et al. Identifying multimodal intermediate phenotypes between genetic risk factors and disease status in Alzheimer's disease. *Neuroinformatics.* (2016) 14:439–52. doi: 10.1007/s12021-016-9307-8

14. Wang M, Hao X, Huang J, Shao W, Zhang D. Discovering network phenotype between genetic risk factors and disease status via diagnosis-aligned multi-modality regression method in Alzheimer's disease. *Bioinformatics.* (2019) 35:1948–57. doi: 10.1093/bioinformatics/bty911

15. Kennis M, Gerritsen L, van Dalen M, Williams A, Cuijpers P, Bockting C. Prospective biomarkers of major depressive disorder: a systematic review and meta-analysis. *Mol Psychiatry.* (2020) 25:321–38. doi: 10.1038/s41380-019-0585-z

16. Liu X, Hou Z, Yin Y, Xie C, Zhang H, Zhang H, et al. CACNA1C gene rs11832738 polymorphism influences depression severity by modulating spontaneous activity in the right middle frontal gyrus in patients with major depressive disorder. *Front Psychiatry.* (2020) 11:73. doi: 10.3389/fpsyt.2020.00073

17. Tolentino JC, Schmidt SL. DSM-5 criteria and depression severity: implications for clinical practice. *Front Psychiatry.* (2018) 9:450. doi: 10.3389/fpsyt.2018.00450

18. Yan C, Zang Y. DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. *Front Syst Neurosci.* (2010) 4:13. doi: 10.3389/fnsys.2010.00013

19. Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* (2007) 81:559–75. doi: 10.1086/519795

20. Flint J, Kendler KS. The genetics of major depression. *Neuron.* (2014) 81:484–503. doi: 10.1016/j.neuron.2014.01.027

21. Gizatullin R, Zaboli G, Jönsson EG, Åsberg M, Leopardi R. Haplotype analysis reveals tryptophan hydroxylase (TPH) 1 gene variants associated with major depression. *Biol Psychiatry.* (2006) 59:295–300. doi: 10.1016/j.biopsych.2005.07.034

22. Jarien e G, Adomaitien e V, Dunderien e L, Dambrauskien e K, Vilkas V, Rybakova I, et al. The Influence of TPH1 and HTR1A gene polymorphisms for a person's suicide risk. In: *European Journal of Human Genetics: EJHG: Abstracts from the 50th European-Society-of-Human-Genetics (ESHG) Conference: Electronic Posters*. London: Nature Publishing Group (2018).

23. Nielsen DA, Deng H, Patriquin MA, Harding MJ, Oldham J, Salas R, et al. Association of TPH1 and serotonin transporter genotypes with treatment response for suicidal ideation: a preliminary study. *Eur Arch Psychiatry Clin Neurosci.* (2020) 270:633–42. doi: 10.1007/s00406-019-01009-w

24. Liu J, Li M, Pan Y, Lan W, Zheng R, Wu FX, et al. Complex brain network analysis and its applications to brain disorders: a survey. *Complexity.* (2017) 2017:1–27. doi: 10.1155/2017/8362741

25. Chen X, Pan W, Kwok JT, Carbonell JG. Accelerated gradient method for multi-task sparse learning problem. In: *2009 Ninth IEEE International Conference on Data Mining*. (Miami, FL: IEEE) 2009. p. 746–51.

26. Roddy DW, Farrell C, Doolin K, Roman E, Tozzi L, Frodl T, et al. The hippocampus in depression: more than the sum of its parts? Advanced hippocampal substructure segmentation in depression. *Biol Psychiatry*. (2019) 85:487–97. doi: 10.1016/j.biopsych.2018.08.021

27. Brosch K, Stein F, Schmitt S, Pfarr JK, Ringwald KG, Thomas-Odenthal F, et al. Reduced hippocampal gray matter volume is a common feature of patients with major depression, bipolar disorder, and schizophrenia spectrum disorders. *Mol Psychiatry*. (2022) 27:4234–43. doi: 10.1038/s41380-022-01687-4

28. Zhang B, Qi S, Liu S, Liu X, Wei X, Ming D. Altered spontaneous neural activity in the precuneus, middle and superior frontal gyri, and hippocampus in college students with subclinical depression. *BMC Psychiatry*. (2021) 21:280. doi: 10.1186/s12888-021-03292-1

29. Lu Y, Liang H, Han D, Mo Y, Li Z, Cheng Y, et al. The volumetric and shape changes of the putamen and thalamus in first episode, untreated major depressive disorder. *Neuroimage Clin*. (2016) 11:658–66. doi: 10.1016/j.nicl.2016.04.008

30. Hagan CC, Graham JM, Tait R, Widmer B, van Nieuwenhuizen AO, Ooi C, et al. Adolescents with current major depressive disorder show dissimilar patterns of age-related differences in ACC and thalamus. *Neuroimage Clin*. (2015) 7:391–9. doi: 10.1016/j.nicl.2014.12.019

31. Peng J, Liu J, Nie B, Li Y, Shan B, Wang G, et al. Cerebral and cerebellar gray matter reduction in first-episode patients with major depressive disorder: a voxel-based morphometry study. *Eur J Radiol*. (2011) 80:395–9. doi: 10.1016/j.ejrad.2010.04.006

32. Green S, Ralph MAL, Moll J, Deakin JF, Zahn R. Guilt-selective functional disconnection of anterior temporal and subgenual cortices in major depressive disorder. *Arch Gen Psychiatry*. (2012) 69:1014–21. doi: 10.1001/archgenpsychiatry.2012.135

33. Li G, Rossbach K, Zhang A, Liu P, Zhang K. Resting-state functional changes in the precuneus within first-episode drug-naive patients with MDD. *Neuropsychiatr Dis Treat*. (2018) 14:1991. doi: 10.2147/NDT.S168060

34. Gao Y, Wang X, Xiong Z, Ren H, Liu R, Wei Y, et al. Abnormal fractional amplitude of low-frequency fluctuation as a potential imaging biomarker for first-episode major depressive disorder: a resting-state fMRI study and support vector machine analysis. *Front Neurol*. (2021) 12:751400. doi: 10.3389/fneur.2021.751400

35. Liu CH, Ma X, Yuan Z, Song LP, Jing B, Lu HY, et al. Decreased resting-state activity in the precuneus is associated with depressive episodes in recurrent depression. *J Clin Psychiatry*. (2017) 78:22409. doi: 10.4088/JCP.15m10022

36. Xia M, Wang J, He Y. BrainNet Viewer: a network visualization tool for human brain connectomics. *PLoS ONE*. (2013) 8:e68910. doi: 10.1371/journal.pone.0068910

37. Dvorak J, Hilke M, Trettin M, Wenzler S, Hagen M, Ghirmai N, et al. Aberrant brain network topology in fronto-limbic circuitry differentiates euthymic bipolar disorder from recurrent major depressive disorder. *Brain Behav*. (2019) 9:e01257. doi: 10.1002/brb3.1257

38. Wang M, Ju Y, Lu X, Sun J, Dong Q, Liu J, et al. Longitudinal changes of amplitude of low-frequency fluctuations in MDD patients: a 6-month follow-up resting-state functional magnetic resonance imaging study. *J Affect Disord*. (2020) 276:411–7. doi: 10.1016/j.jad.2020.07.067

39. Huang J, Chen Z, Zhu L, Wu X, Guo X, Yang J, et al. Phosphoinositide-3-kinase regulatory subunit 1 gene polymorphisms are associated with schizophrenia and bipolar disorder in the Han Chinese population. *Metab Brain Dis*. (2020) 35:785–92. doi: 10.1007/s11011-020-00552-z

40. Chen X, Long F, Cai B, Chen X, Qin L, Chen G. A novel relationship for schizophrenia, bipolar, and major depressive disorder. Part 8: a hint from chromosome 8 high density association screen. *Mol Neurobiol*. (2017) 54:5868–82. doi: 10.1007/s12035-016-0102-1

41. Abo El Fotoh WMM, Bayomy NR, Kasemy ZA, Barain AM, Shalaby BM, Abd El Naby SA. Genetic variants and haplotypes of tryptophan hydroxylase 2 and reelin genes may be linked with attention deficit hyperactivity disorder in Egyptian children. *ACS Chem Neurosci*. (2020) 11:2094–103. doi: 10.1021/acschemneuro.0c00136

42. Kulminski AM, Shu L, Loika Y, He L, Nazarian A, Arbeev K, et al. Genetic and regulatory architecture of Alzheimer's disease in the APOE region. *Alzheimers Dement*. (2020) 12:e12008. doi: 10.1002/dad2.12008

43. Pergola G, Di Carlo P, Jaffe AE, Papalino M, Chen Q, Hyde TM, et al. Prefrontal coexpression of schizophrenia risk genes is associated with treatment response in patients. *Biol Psychiatry*. (2019) 86:45–55. doi: 10.1016/j.biopsych.2019.03.981

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 23 March 2023
DOI 10.3389/fpsyt.2023.1125339

OPEN ACCESS

EDITED BY
Zhen Zhou,
University of Pennsylvania, United States

REVIEWED BY
Dongren Yao,
Massachusetts Eye and Ear Infirmary and Harvard Medical School, United States
Wenming Zhao,
First Affiliated Hospital of Anhui Medical University, China

*CORRESPONDENCE
Han-Gue Jo
✉ hgjo@kunsan.ac.kr

SPECIALTY SECTION
This article was submitted to Computational Psychiatry, a section of the journal Frontiers in Psychiatry

RECEIVED 16 December 2022
ACCEPTED 02 March 2023
PUBLISHED 23 March 2023

CITATION
Venkatapathy S, Votinov M, Wagels L, Kim S, Lee M, Habel U, Ra I-H and Jo H-G (2023) Ensemble graph neural network model for classification of major depressive disorder using whole-brain functional connectivity. *Front. Psychiatry* 14:1125339. doi: 10.3389/fpsyt.2023.1125339

Ensemble graph neural network model for classification of major depressive disorder using whole-brain functional connectivity

Sujitha Venkatapathy[1], Mikhail Votinov[2,3], Lisa Wagels[2,3], Sangyun Kim[4], Munseob Lee[4], Ute Habel[2,3], In-Ho Ra[1] and Han-Gue Jo[1]*

[1]School of Computer Information and Communication Engineering, Kunsan National University, Gunsan, Republic of Korea, [2]Department of Psychiatry, Psychotherapy and Psychosomatics, Medical Faculty, Uniklinik RWTH Aachen University, Aachen, Germany, [3]Research Center Juelich, Institute of Neuroscience and Medicine: JARA-Institute Brain Structure Function Relationship (INM 10), Juelich, Republic of Korea, [4]AI Convergence Research Section, Electronics and Telecommunications Research Institute, Gwangju, Republic of Korea

Major depressive disorder (MDD) is characterized by impairments in mood and cognitive functioning, and it is a prominent source of global disability and stress. A functional magnetic resonance imaging (fMRI) can aid clinicians in their assessments of individuals for the identification of MDD. Herein, we employ a deep learning approach to the issue of MDD classification. Resting-state fMRI data from 821 individuals with MDD and 765 healthy controls (HCs) is employed for investigation. An ensemble model based on graph neural network (GNN) has been created with the goal of identifying patients with MDD among HCs as well as differentiation between first-episode and recurrent MDDs. The graph convolutional network (GCN), graph attention network (GAT), and GraphSAGE models serve as a base models for the ensemble model that was developed with individual whole-brain functional networks. The ensemble's performance is evaluated using upsampling and downsampling, along with 10-fold cross-validation. The ensemble model achieved an upsampling accuracy of 71.18% and a downsampling accuracy of 70.24% for MDD and HC classification. While comparing first-episode patients with recurrent patients, the upsampling accuracy is 77.78% and the downsampling accuracy is 71.96%. According to the findings of this study, the proposed GNN-based ensemble model achieves a higher level of accuracy and suggests that our model produces can assist healthcare professionals in identifying MDD.

KEYWORDS

major depressive disorder, deep learning, graph neural network, ensemble model, functional connectivity

1. Introduction

Depression is a major source of disability and disease burden worldwide, affecting about 264 million people. Major depressive disorder (MDD) is a serious psychological problem that can cause people to feel sad, lose interest, become listless, and have trouble thinking (1). Individuals who have suffered from MDD typically have difficulty adjusting together with their society. They have a low opinion of themselves, which ultimately leads to a decline in their performance at work. MDD can cause serious emotional problems and suicidal thoughts and behavior if it is not adequately recognized and treated (2). In patients suffering from MDD, abnormalities in large-scale brain connections have been identified more frequently in recent years. Depressed people revealed significantly disrupted connections between the task-related regions of the brain throughout a variety of task-directed functions, such as working memory, executive control, facial emotion perception, and impulse control (3, 4).

In recent years, research on MDD has focused on brain structure and function using morphological or neurobiological features. Functional magnetic resonance imaging (fMRI), magnetoencephalography (MEG), electroencephalography (EEG), and positron emission tomography (PET) are the common physiological methods employed in comparing people with MDD to healthy controls (HCs) (5). Researchers have found that patients with MDD have abnormal communication among the functional brain networks using functional connectivity (FC) of resting-state fMRI (R-fMRI), which detects synchronized and desynchronized spontaneous activity within anatomically diverse networks (6–8). In the present study, whole brain FC is extracted from R-fMRI data in order to determine whether or not the subject is MDD, and for classification between first-episode and recurrent MDD.

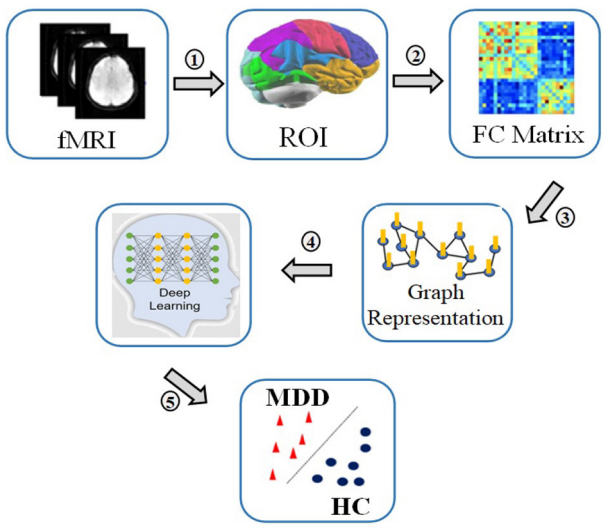


FIGURE 1
The process of MDD and HC classif ication: 1. Recording R-fMRI signal, 2. Extracting time series of regions of interest, 3. Calculating correlation matrices of the time series, 4. Converting correlations in graph representations, 5. Applying deep learning to classify between patients and controls.

Machine learning (ML) methods are of increasing interest for the medical industry at present, and it has emerged as an essential part of the diagnosis and treatment of conditions pertaining to oncology, neurology, and cardiology. The process involved in a typical deep learning pipeline for the identification of MDD can be highlighted as follows: region of interest (ROI) extraction from R-fMRI, functional connectivity matrix generation, graph construction, deep learning model training, and classification (9). Figure 1 depicts the processes required in identifying MDD from HCs.

This study presents a high-performance graph neural network (GNN)-based deep learning method for classifying individuals with MDD using R-fMRI data. In recent years, graph neural network (GNN) has become increasingly popular in graph-based learning. GNN become the optimal deep learning approach for analyzing graph-structured information. GNN algorithm combines node attributes, edge attributes, and graph topology by embedding node characteristics in a neural network and transferring data through the graph's edges. GNN can work well on non-euclidean domains and this is in contrast to traditional convolutional neural networks, which are limited to accepting only euclidean inputs (10). GNN has replaced older ML approaches due to their greater performance in analyzing graph-based information (11, 12).

Studies on MDD have progressed in recent years understanding changes of brain structure and function using morphological or neurobiological features. The studies summarized here used a number of machine learning algorithms, such as support vector machines (SVM), logistic regression, and neural networks, to differentiate between MDD and HCs using fMRI data. Especially, resting-state functional connectivity characteristics of the entire brain were studied in MDD. In order to distinguish individuals with MDD from controls, several studies (13–16) employed SVM-based multivariate pattern analysis (MVPA) techniques, achieving a better classification accuracy. However, there are limitations to this approach that stem from small sample sizes, scanner variability, and the absence of a comprehensive independent data set. By computing the Hurst exponents of resting-state networks, researchers examined their long-term memory for distinguishing depressive patients from HCs (17). Scale-free dynamics of depression-related brain activity were seen as describing the long-term memory of resting-state networks. An SVM-based classifier was used to test the data with a leave-one-out cross-validation (Loocv) method. Others studied the effects of MDD and schizophrenia on whole brain R-fMRI using SVM based MVPA in Yu et al. (18), Lois and Wessa (19), Zhu et al. (20), and Li et al. (21). Again, the dimensionality reduction technique relied on the Loocv strategy due to the small sample size. However, it is essential to evaluate the classification performance of these perspectives using a larger sample of subjects. Others showed that whole-brain R-fMRI connectivity may effectively predict antidepressant medication status in people with serious MDD (22). Medication-naive patients were distinguished from controls by the use of a trained linear SVM classifier based on MVPA technique. A different MVPA strategy based on linear, radial basis function (RBF)-SVM classification with the elastic net feature selection technique could accurately distinguished MDD patients from control subjects (23). The hyper-networks in this study were built using an elastic net and the group lasso technique. Hyper-edge, brain area, and average

metric analyses suggested that the hyper-networks built with elastic net and group lasso differed structurally (24). Further, functional connection density measures derived by R-fMRI have shown to be successful to determine the relationship between the changes in resting-state activities and the responses to electroconvulsive therapy in 23 patients with MDD (25). The neural indices that were discovered as classification criteria were entered into linear SVM based MVPA, which was then used to categorize MDD patients. In another study, the identification of MDD among subjects, was explored *via* different static and dynamic connection metrics retrieved from R-fMRI (26). In this study a feature vector for classification was built by combining features from static and dynamic techniques. To determine the final predictor performance, a Loocv procedure was implemented. Differential sub-graph entropy and dynamic connectivity characteristics have been used in an SVM classifier to distinguish between people with MDD and HCs. A sliding-window approach was implemented to determine functional connectivity in the context of dynamic processes (27). The dynamic functional connectivity matrices were then employed as features in a non-linear SVM model to differentiate MDD patients from controls. Analysis techniques based on the minimum spanning tree need the computation of measurable qualities and the selection of these attributes as features in the classification. Using two feature types to measure two aspects of the network, a multi-kernel SVM classification method (28) allows the use of both brain region features and subgraph features. All studies discussed above made use of an approach that is based on linear SVM, multi-kernel SVM and RBF-SVM to differentiate between MDD patients and HCs.

Some other studies have used different ML methods, such as regression and neural networks, to differentiate between individuals with MDD and HCs. Using partial least squares regression on R-fMRI data, researchers developed a low-dimensional representation that links symptoms to brain activity and predicts clinical measures (29). R-fMRI connectivity in another study was calculated by employing the automated anatomical labeling layout with a partial correlation method (30). The calculations of the metrics and the classification analysis were performed in the frame of neural network. However, the efficacy and sorting of selected features, as well as sample size, kind of classifiers, and distribution of data, all have a role in determining the appropriate amount of features. For example, when developing a classifier for melancholy MDD (31), it has been shown that is it crucial to identify vitally relevant functional connections. It is for such cases recommended to use logistic regression to evaluate the uniformity vs. heterogeneity connectivity hypotheses.

The concept of deep learning has recently received significant attention. Notably, graph-based techniques, such as GNN, have been used to investigate detailed node pair in imaging/nonimaging characteristics among participants, with the goal of identifying significant phenotypes for clinical identification. Successfully applying a whole-brain data-driven approach with R-fMRI, confirmed the use of effective connectivity for MDD detection by calculating its measures *via* a group sparse representation and a structured equation modeling approach (32). Successful integration of effective connectivity and nonimaging phenotypic information allowed the use of spectral graph convolutional networks (GCN) based on a population graph to differentiate drug-naive MDD patients from HCs. Using functional connectivity as a characteristic, Ktena et al. (33) trained a spectral GCN with subjects as nodes. The spectral GCN was used to diagnose the problem by grouping the nodes into their respective categories. Others used a mutual multi-scale triplet GCN (34) for the purpose of analyzing static FC and structural connectivity with the intention of identifying brain disorders. Further a spatio-temporal GCN framework was created to train discriminative features from FC measures for the automated identification and treatment response prediction of MDD (35). The GCN model was developed to every participant's whole-brain functional network in order to differentiate MDD patients from HCs, recognize the most important regions making a contribution to classifying, and investigate the association between structural features of salient regions and clinical features (36).

In this research, our main aim is to employ an ensemble-based GNN framework to perform the primary classification analysis between MDD and HCs as well as subgroup analysis between first-episode and recurrent (REC). Instead of using a single unified GNN model to learn representations for all of the nodes in a large graph, it is better to use ensemble learning methods (37) to improve classification performance. With ensemble learning, many fundamental classifiers are combined to boost the predictive power of the model. Therefore, to enhance the efficiency of scalable GNN, we propose a GNN-based ensemble model that creates customized models.

1.1. Aims

This study provides a methodology for MDD analysis and classification using brain functional networks derived from R-fMRI data. Since the brain is a complex network system, the present study analyses the R-fMRI as the whole brain functional structure rather than individual FCs. In this research, our main aim is to employ an ensemble- based GNN framework to perform the primary classification. Our main scientific contributions are as follows:

- Developing a robust ensemble-based GNN model that takes R-fMRI data for the detection of MDD. A created novel ensemble model is used for identifying individuals with MDD from HCs as well as to perform analysis between two sub groups of MDD patients, namely first episode drug nive (FEDN) and recurrent (REC) MDD patients.
- GCN, GAT, and GraphSAGE models are created as the base line models for the ensemble model, for improving classification accuracy. A developed ensemble model is trained using individual whole-brain functional networks.
- Methods of upsampling and downsampling are employed to achieve balanced sample size. A 10-fold enumeration is used to refine the classification process. Empirical investigations with a large sample size showed that our model is more accurate and beneficial for classification of MDD compared to other models that are currently available.

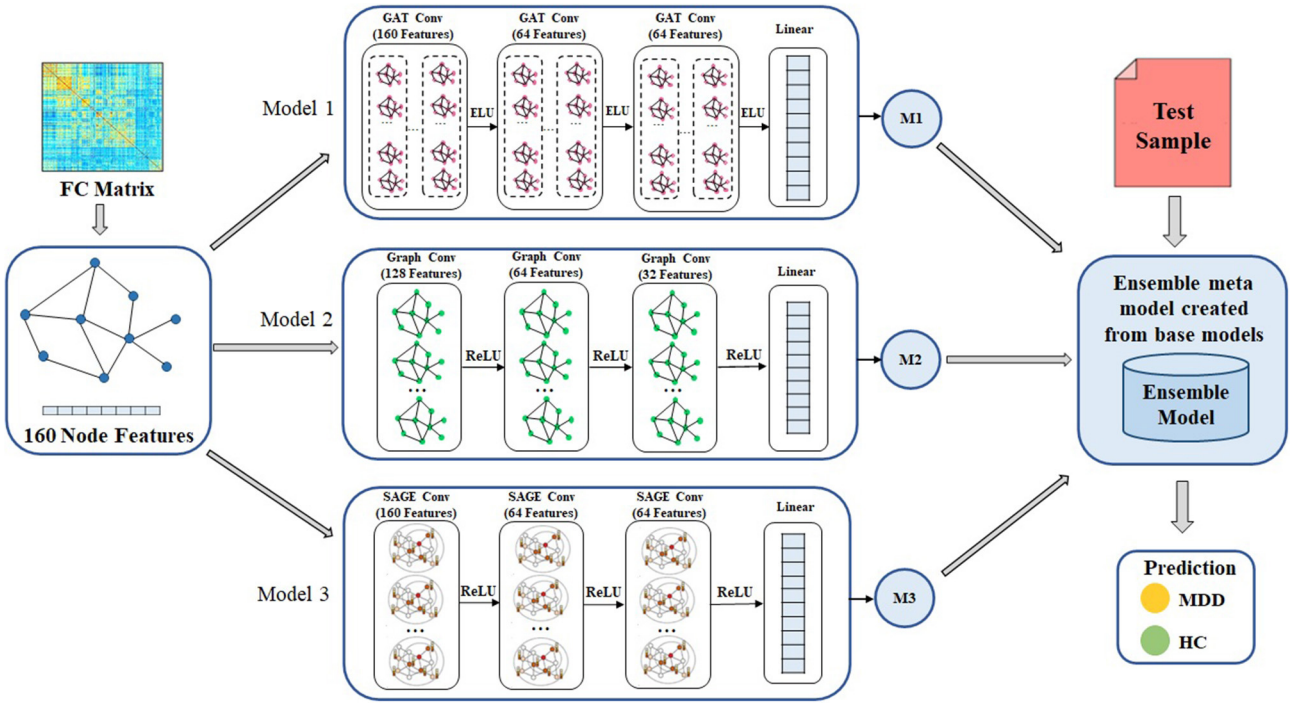


FIGURE 2
Overall framework of the proposed method for MDD classification. Ensemble model is designed to create meta model from the out features of each base models. GAT, graph attention network; GCN, graph convolutional network; SAGE, GraphSAGE technique, ReLU, Rectified Linear Unit; ELU, Exponential Linear Unit, MDD, major depressive disorder; HC, healthy controls.

2. Materials and methods

2.1. Subjects

R-fMRI data from the REST-meta-MDD collaboration (6), which contained 25 datasets totaling 2428 persons (1300 MDD patients and 1128 HCs from 17 hospitals), was used in the present study. There have been 562 patients with MDD who were experiencing their first episode of the disorder, as well as 282 patients with MDD who have been experiencing recurring episodes of the disorder. According to a previous publication on the dataset (6), we used criteria such as missing data, low-quality spatial normalization, insufficient coverage, noticeable head movement, and sites with fewer than 10 subjects to exclude. This was produced in a sample of 821 people with MDD and 765 HCs from 16 different sites. Drug consumption data was submitted by 527 patients; 219 of these individuals are currently using first episode MDD patients without medication treatment was defined as FEDN, and REC is the MDD patients with recurrent episode regardless of medication status. Two research groups (Sites 5 and 13) contributed data on 117 FEDN patients and 72 patients with REC MDD, five research groups (Sites 4, 5, 9, 13, and 16) contributed data on 227 FEDN patients and 388 HCs, and six research groups (Sites 3, 5, 7, 12, 13, and 14) contributed statistics on 189 patients with REC and 423 HCs. The studies involving human participants data were reviewed and approved by the Institutional Review Board of Kunsan National University.

2.2. Preprocessing

Data from R-fMRI and structural MRI were acquired and the DPARSF toolkit (38) was used to perform preprocessing procedure. Slice timing correction, head motion correction, normalization, and the elimination of confounds were the main preprocessing procedures. Dosenbach's atlas was used as a reference point during the process of segmenting the entire brain into 160 distinct regions of interest (ROI) (39). The voxel-level BOLD values were extracted and averaged across all ROIs. The Pearson correlation coefficient of the related time series was used to assess FC between each pair of ROIs. Finally, the correlation estimates were transformed using Fisher's z-transform to generate FC matrix in the range of 160 × 160 for each subject (40).

2.3. Methods

The overall process of the GNN-based ensemble model is shown in Figure 2. The FC matrix the whole brain is initially depicted as a weighted undirected graph $G(N, E)$, where N and E are collections of nodes and edges. Nodes are the 160 brain regions identified by the ROIs, and their characteristics are the matrix representation of the functional connection between them. The link between nodes are represented by a weighted adjacency matrix (A). Each node is linked to its nearest neighbors using a k-nearest neighbors (KNN) technique to establish edges (36). In order

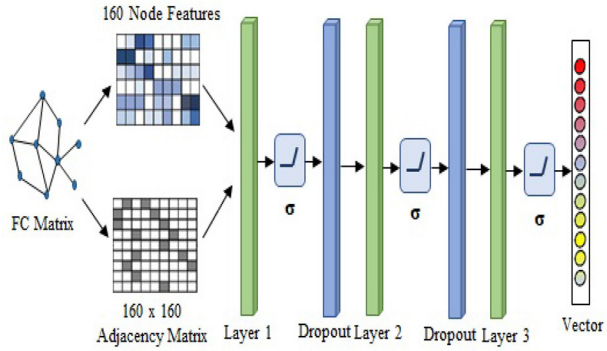


FIGURE 3
Structure of the base model.

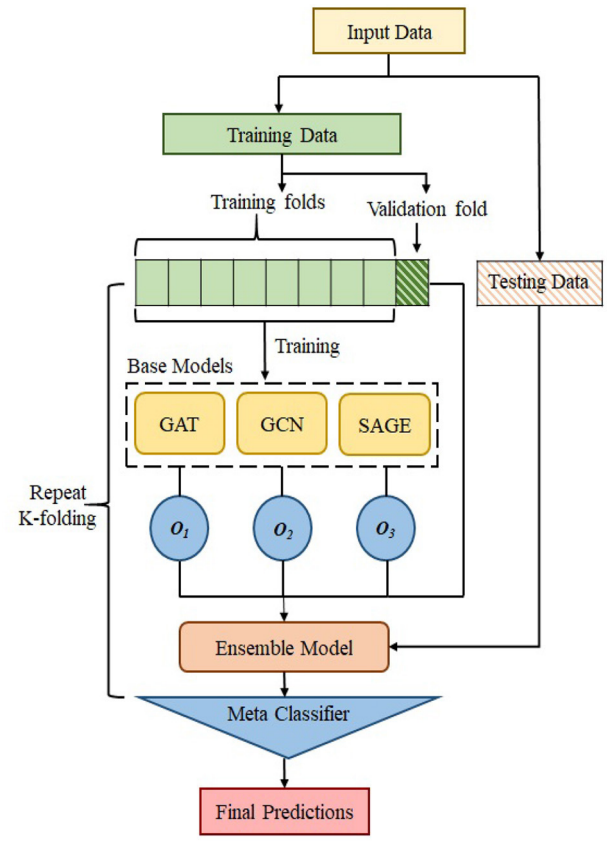


FIGURE 4
Flowchart to show the process of ensemble model.

to construct the GNN-based ensemble model, the GCN, GAT, and SAGE models are used as the base models. The core operation of an ensemble model is to combine out features of base models and apply the softmax activation function to translate final scalers into predicted probabilities of the each class.

2.3.1. Base models

First, the FC matrix was represented as a graph structure, along with an adjacency matrix and node characteristics. The GAT model uses a graph attention layer to learn the node representation, followed by an attention pooling layer and a classification layer to retrieve the node representation and perform the task of learning. We began by stacking three GAT layers with exponential linear unit (eLU) activation functions, then moved on to a global mean pooling layer, then a dropout layer, and finally fed that information into a classification layer. The head size is 8 and the dropout rate is 0.5 in the GAT model. Input layer, graph convolutional hidden layer, fully connected layer, and global average pooling layer are the components that make up GCN model. After each hidden layer, there is a rectified linear unit (ReLU) activation function. The dropout rate is 0.3 in the GCN model. We also use GraphSAGE with three layers ($K = 3$) as a base model. The dimensions of the node embeddings are the same as the size of the hidden units that are used in each GraphSAGE layer, which is 64. The ReLU activation function and a dropout rate of 0.3 are used in each of the GraphSAGE layers. Also, Every model has a weight decay value of 5×10^{-4} and a learning rate of 0.01. Each model in GNN, such as GCN, GAT, and GraphSAGE, is trained with the same set of node features, edge features, weights, and learning rate. Figure 3 illustrates the processes involved in creating a base model.

2.3.2. Ensemble model

We use the ensemble-based GNN model in order to determine the essential features that contribute to the prediction of MDD. The suggested ensemble-GNN procedure is depicted in the flowchart described in Figure 4. We construct a GNN-based ensemble model using GCN, GAT, and GraphSAGE as building blocks. Both the node features and the adjacency matrix are fed into the base models, and the features' identities are then extracted from the respective models. Each model's predicted output features are fed into the ensemble model. Then, for class prediction, we add fully connected layers with a softmax activation function. The cross-entropy loss function is put into effect to this extent. The adam optimizer is used to find the optimal values for each of the model's parameters.

2.4. MDD identification and evaluation

MDD classification employs ensemble-based GNN supervised learning classifiers, and this classification makes it simple to compare the efficacy of various machine learning strategies for processing fMRI data. Both training and testing are required for supervised learning classification. With the help of the samples' class labels, the classifier identifies a decision boundary that divides the input space during the training phase. Once the decision function has been calculated using the training set, it may be applied to unseen testing data to infer the corresponding class label. In order to reduce the overfitting problem and to offer a reliable and generalizable classification performance evaluation, the effectiveness of the classification framework is evaluated using the 10-fold cross validation scheme. In order to balance the sample size, oversampling is accomplished by copying data from minority classes, whereas undersampling is carried out by selecting data from majority classes. The performance of the classification system is measured and analyzed based on its accuracy (ACC), specificity (SPE), sensitivity (SEN), and area under the curve (AUC). The

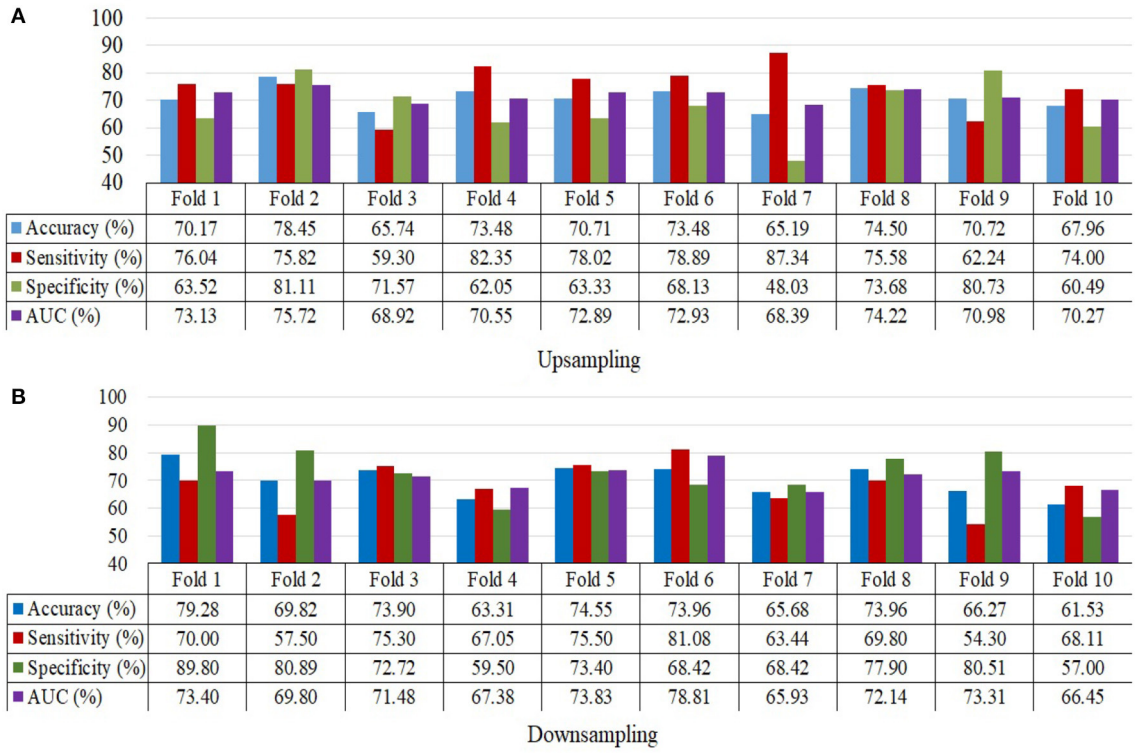


FIGURE 5
10-Fold cross validation for MDD and HC samples depicting accuracy, sensitivity, specificity and the area under the curve. **(A)** Upsampling and **(B)** Downsampling.

diagnostic accuracy of a classifier can be measured with the use of receiver operating characteristic, which is a curve that is generated by graphing the true positive rate against the false positive rate. The process of determining classification ACC, SPE, and SEN is denoted as,

$$ACC = \frac{TP + TN}{TP + FP + FN + TN} \quad (1)$$

$$SPE = \frac{TN}{TN + FP} \quad (2)$$

$$SEN = \frac{TP}{TP + FN} \quad (3)$$

In this case, TP indicates a successful classification of positive samples, TN indicates a successful classification of negative samples, and FP indicates incorrect classification of negative samples as positive, and FN indicates incorrect negative classification.

3. Results

In this section, we validate the efficiency of the suggested MDD identification approach by analyzing the following scenarios: (a) using FC as features; (b) using GCN, GAT, and GraphSAGE as base learners; (c) using an ensemble classification model. We used a 10-fold cross-validation, and in that cross-training, we split the samples of all MDD and HCs into 10 groups. Each time the method is modified, one unit is chosen as the testing dataset for assessing the performance of the model, while the remaining 9 units are used as the training dataset. By stratifying the 10-fold cross-validation, we are able to keep the percentage of samples from each class in every fold equal across the entire sample. In this case, the samples are not balanced, so in order to create samples that are balanced, random upsampling is performed on the majority classes, and random downsampling is performed on the minority classes. For the primary analysis, there were a total of 1,586 participants included in our method (821 patients with MDD and 765 HCs). Based on the clinical data for the patients who are included, 243 were FEDN patients, and 203 are REC patients.

3.1. MDD vs. HC classification

The ensemble model attained an accuracy of 71.8% for upsampling and 70.4% for downsampling when it came to classifying MDD and HC. When using upsampling, the ensemble model achieves an AUC of 76.53%, while using downsampling, it achieves an AUC of 71.27%. Specificity and sensitivity values for upsampling are 74.96 and 68.23%, respectively, whereas the values for downsampling are 67.27 and 72.88%. The findings for upsampling and downsampling based on a 10-fold cross validation for MDD and HC classification are given in Figure 5.

3.2. FEDN vs. HC classification

FEDN could be distinguished from HCs with a classification accuracy of 88.93% for upsampling and 64.17% for downsampling. Classification of FEDN patients with HC achieved an upsampling AUC of 85.75% and a downsampling AUC of 62.25%. Upsampling has a specificity and sensitivity of 89% and 85.79%, whereas downsampling has a specificity and sensitivity of 60.31 and 61.36%. The findings from classifying FEDN and HCs are shown in Table 1.

3.3. REC vs. HC classification

Classification accuracy for upsampling REC with HC is 91.6%, whereas classification accuracy for downsampling REC patients with HC is 68.78%. Using upsampling, we are able to classify REC patients as distinct from HC with an AUC of 88.24%, while using downsampling, we are only able to reach an AUC of 67.11 %. The respective results for specificity and sensitivity while upsampling are 93.15 and 87.20%, while they are 60.96 and 66.92% when downsampling. Table 2 describes the results produced from the classification of REC and HCs.

3.4. FEDN vs. REC classification

The FEDN and the REC are discriminated from each other by the use of a subgroup analysis. There is a 77.78% accuracy rate when upsampling FEDN against REC and a 71.96% rate when downsampling. When we compared the existing approach (36) to the subgroup analysis, we found that the classification performance for the characterization of recurring patients was greater than that of FEDN. Using upsampling, we achieved an AUC of 75.19%, whereas using downsampling, we achieved an AUC of 71.77%. Upsampling yields results of 72.81 and 81.91% for specificity and sensitivity, whereas downsampling yields results of 71.52 and 71.56%. The outcomes of the FEDN and REC classifications are shown in Table 3.

TABLE 1 Ensemble model performance for FEDN vs. HC classification.

Ensemble Model - FEDN vs. HC				
Sampling	ACC	SEN	SPE	AUC
Upsampling	0.8893	0.8900	0.8597	0.8584
Downsampling	0.6417	0.6031	0.6138	0.6225

TABLE 2 Ensemble model performance for REC vs. HC classification.

Ensemble Model - REC vs. HC				
Sampling	ACC	SEN	SPE	AUC
Upsampling	0.9160	0.9315	0.8720	0.8824
Downsampling	0.6878	0.6096	0.6692	0.6711

TABLE 3 Ensemble model performance under FEDN vs. REC classification.

Ensemble Model - FEDN vs. REC				
Sampling	ACC	SEN	SPE	AUC
Upsampling	0.7778	0.7281	0.8191	0.7519
Downsampling	0.7196	0.7150	0.7152	0.7177

4. Discussion

We proposed an ensemble-based GNN method for automatic MDD identification using whole brain functional network features. Using a large open source dataset, the current study employed an ensemble based GNN to classify MDD as well as to classify FEDN with REC, and the resulting upsampling classification performance outperformed typical machine learning approaches by around, 71.18% and 77.78%, respectively. In the analysis process, initially, separate base models are created, and then the classification performance of each model is examined. Models such as GCN, GAT, and GraphSAGE are employed as base line models. The GAT approach is utilized during the training of the model, which resulted in an upsampling accuracy of 66.24% when comparing MDD to HC and 71.67% when comparing FEDN to REC. The GCN technique is also separately applied during the training process of the model, which led to an upsampling accuracy of 64.72% while correlating MDD to HC and 73.58% while comparing FEDN to REC. Also, the GraphSAGE model alone was employed when training the model, which produced in an upsampling accuracy of 64.47 % for MDD among HC classification and 72.78% for FEDN with REC classification. In addition to this, we used an all-individual model to analyze subgroups such as FEDN with HC and REC with HC. The results of the individual base models such as GCN, GAT, and GraphSAGE are listed in Table 4. The GNN-based ensemble model is developed to improve the classification accuracy of primary analysis as well as subgroup analysis in analyzing MDD. Already, the base models are trained independently, and while some findings indicate that GCN produces better results, other findings show that GAT or GraphSAGE produces better outcomes. That indicates that no one model achieves better results for all classes. Because of this, a combined model is developed to produce more accurate results for the whole sample. Figure 6 displays the results of a comparison between the base model and the ensemble model.

In the majority of the earlier investigations, the ML algorithm was employed to differentiate between all MDD and HCs (20, 41–43). To identify brain disorders, researchers have created a number of deep learning techniques, including BrainNetCNN (44) and discriminative/generative long short-term memory (45). However, the sample size that they employed for the investigation was comparatively small. The use of GNN to distinguish between MDD and HCs has been proven effective in a small number of studies. Some previous research (32, 34, 46) has shown that the graph convolution technique can be used to distinguish disorders in patients with HC. This is in contrast to others, who employed MVPA of static or dynamic functional connectivity in the brain network (47, 48), which neglected topological elements

TABLE 4 Base model performance under different group analysis.

Model	Upsampling				Downsampling			
	ACC	SEN	SPE	AUC	ACC	SEN	SPE	AUC
MDD vs. HC								
GCN	0.6472	0.6171	0.6537	0.7122	0.6519	0.5991	0.7041	0.7089
GAT	0.6624	0.6619	0.6667	0.6734	0.6219	0.6176	0.6270	0.6640
SAGE	0.6447	0.5968	0.6715	0.7220	0.6213	0.6713	0.5705	0.6859
FEDN vs. HC								
GCN	0.8593	0.8599	0.8588	0.8543	0.5759	0.7263	0.4188	0.6185
GAT	0.8320	0.8299	0.7929	0.8387	0.5759	0.6446	0.4862	0.6200
SAGE	0.8520	0.8780	0.7986	0.8692	0.5667	0.5304	0.5994	0.6211
REC vs. HC								
GCN	0.8913	0.9397	0.8728	0.9051	0.6363	0.5936	0.6824	0.6625
GAT	0.8527	0.8846	0.7929	0.8534	0.5609	0.5763	0.5828	0.6653
SAGE	0.8746	0.8889	0.8627	0.8946	0.5826	0.5964	0.5705	0.6035
FEDN vs. REC								
GCN	0.7358	0.7069	0.6822	0.7653	0.6651	0.6562	0.6776	0.7101
GAT	0.7167	0.6876	0.7267	0.7113	0.6109	0.6420	0.5945	0.7016
SAGE	0.7278	0.6608	0.8000	0.7539	0.6652	0.7138	0.6166	0.7126

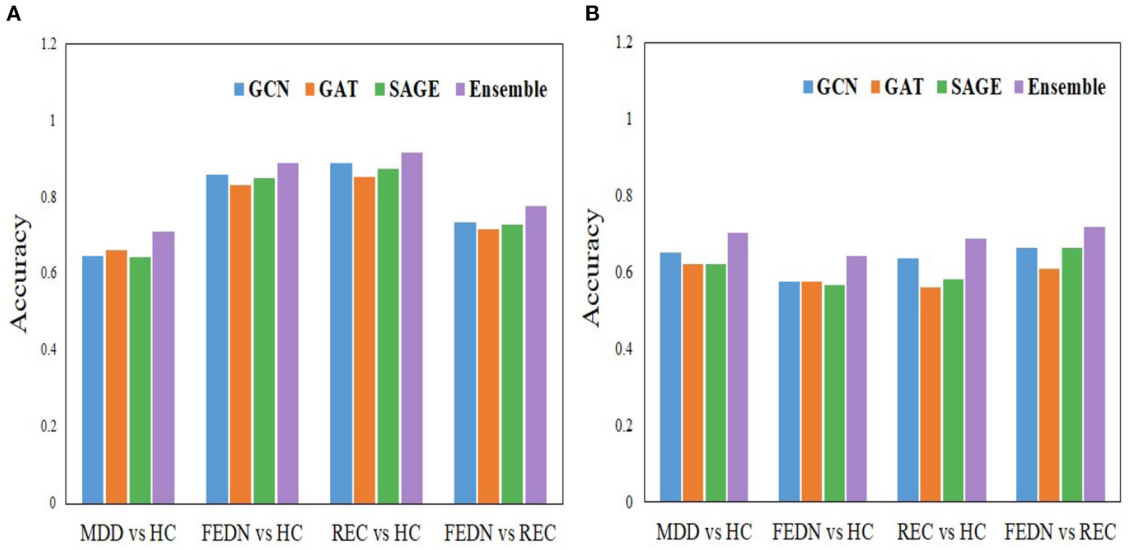


FIGURE 6
Comparison analysis between the individual model and the ensemble model. **(A)** Upsampling and **(B)** Downsampling.

that could provide key clues for diagnosis. The ensemble-GNN algorithm combines the results of numerous GNN classifiers into a single model in order to minimize the impact of overfitting. In addition, by using ensemble-GNN, the deviation caused by a single classifier can be reduced, resulting in improved reliability. When applied to imbalanced datasets with the same number of learning epochs, ensemble-GNN obtains a higher classification accuracy than a single classifier does. We took primary analyses into consideration such as all MDD with HCs as well as subgroup analyses including FEDN among HCs, REC with HCs, and FEDN along with REC. In addition, We also analyzed the model using the Craddock and Automated Anatomical Labeling (AAL) atlases.Our ensemble GNN model achieves an accuracy of 74.75% for the AAL atlas and 73.37% for the Craddock atlas when classifying MDD vs. HC upsampling data. Tables 2, 4 of the Supplementary material contain the primary and all of the sub class analysis results for the AAL and Craddock atlas, respectively.

5. Limitations

Some limitations should be taken into account when evaluating the present findings. We were not able to reach the classification performance in the case of MDD with HC categorization when compared with the previous approach (36). However, our methods achieve higher performance in subgroup classification. In order to solve the problem of imbalanced class representation, we tried both random sampling. Upsampling gives better results, but it can add a redundant samples to the model, which slows down training and vulnerable to overfitting. In our approach, the overfitting problem is reduced by utilizing cross-validation; however, training speed is not taken into account. Also, with respect to the population, most MDD was females which may have a confounding effect on MDD classification. However, it should be noted that the same GNN model structures could also successfully classify between MDD subgroups, FEDN vs. REC, implying that MDD classification is not entirely dependent on the sex effect. Further studies that aim to test this effect requires a much larger sample size for controlling sex or adequate matching.

6. Conclusion

In this study, we effectively created an ensemble model based on GNN for classifying MDD by utilizing R-fMRI data. In particular, we investigated a sub-group analysis between FEDN with REC. The proposed model that employs whole brain functional connectivity classifies MDD patients and healthy individuals with high accuracy. In order to improve the overall performance of the ensemble model, we used three different GNN base models under 10-fold cross-validation. Based on large dataset and a number of different of validation techniques, an ensemble model could classify MDD and HCs with a feasible accuracy of 71.18% for upsampling and 70.14% for downsampling. When compared to earlier approaches, the findings that these methods yield in the subgroup analysis are higher. This method achieves an accuracy of 77.78% for upsampling and 71.96% for downsampling when applied to an analysis of FEDN and REC. The findings of this validation suggest that our model produces a feasible application to assist healthcare professionals in identifying MDD.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the Institutional Review Board of Kunsan National University. Written informed consent from the was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

SV, MV, LW, SK, ML, UH, I-HR, and H-GJ contributed to conception and implementation of the study. SV wrote the first draft. All authors contributed to review, editing, and approved the submitted version.

Funding

This work was supported by Electronics and Telecommunications Research Institute (ETRI) grant funded by the Korean government (23ZK1100, Honam region regional industry-based ICT convergence technology advancement support project), and by the FZJ-NST Bilateral Cooperation Programme funded by the Forschungszentrum Jülich and the National Research Council of Science & Technology (Global-22-001).

Conflict of interest

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyt.2023.1125339/full#supplementary-material

References

1. Yang H, Chen X, Chen ZB, Li L, Li XY, Castellanos FX, et al. Disrupted intrinsic functional brain topology in patients with major depressive disorder. *Mol Psychiatry*. (2021) 26:7363–71. doi: 10.1038/s41380-021-01247-2

2. Rafiei A, Zahedifar R, Sitaula C, Marzbanrad F. Automated detection of major depressive disorder with EEG signals: a time series classification using deep learning. *IEEE Access*. (2022) 10:73804–17. doi: 10.1109/ACCESS.2022.3190502

3. Müller VI, Cieslik EC, Serbanescu I, Laird AR, Fox PT, Eickhoff SB. Altered brain activity in unipolar depression revisited: meta-analyses of neuroimaging studies. *JAMA Psychiatry*. (2017) 74:47–55. doi: 10.1001/jamapsychiatry.2016.2783

4. Cai H, Qu Z, Li Z, Zhang Y, Hu X, Hu B. Feature-level fusion approaches based on multimodal EEG data for depression recognition. *Inf Fusion*. (2020) 59:127–38. doi: 10.1016/j.inffus.2020.01.008

5. Shao X, Sun S, Li J, Kong W, Zhu J, Li X, et al. Analysis of functional brain network in MDD based on improved empirical mode decomposition with resting state EEG data. *IEEE Trans Neural Syst Rehabil Eng*. (2021) 29:1546–56. doi: 10.1109/TNSRE.2021.3092140

6. Yan CG, Chen X, Li L, Castellanos FX, Bai TJ, Bo QJ, et al. Reduced default mode network functional connectivity in patients with recurrent major depressive disorder. *Proc Natl Acad Sci USA*. (2019) 116:9078–83. doi: 10.1073/pnas.1900390116

7. Li G, Liu Y, Zheng Y, Li D, Liang X, Chen Y, et al. Large-scale dynamic causal modeling of major depressive disorder based on resting-state functional magnetic resonance imaging. *Hum Brain Mapp*. (2020) 41:865–81. doi: 10.1002/hbm.24845

8. Kaiser RH, Whitfield-Gabrieli S, Dillon DG, Goer F, Beltzer M, Minkel J, et al. Dynamic resting-state functional connectivity in major depression. *Neuropsychopharmacology*. (2016) 41:1822–30. doi: 10.1038/npp.2015.352

9. Khan DM, Masroor K, Jailani MFM, Yahya N, Yusoff MZ, Khan SM. Development of wavelet coherence EEG as a biomarker for diagnosis of major depressive disorder. *IEEE Sens rs J*. (2022) 22:4315–25. doi: 10.1109/JSEN.2022.3143176

10. Bronstein MM, Bruna J, LeCun Y, Szlam A, Vandergheynst P. Geometric deep learning: going beyond euclidean data. *IEEE Signal Process Mag*. (2017) 34:18–42. doi: 10.1109/MSP.2017.2693418

11. Kim BH, Ye JC. Understanding graph isomorphism network for rs-fMRI functional connectivity analysis. *Front Neurosci*. (2020) 14:630. doi: 10.3389/fnins.2020.00630

12. Li X, Zhou Y, Dvornek N, Zhang M, Gao S, Zhuang J, et al. Braingnn: interpretable brain graph neural network for fmri analysis. *Med Image Anal*. (2021) 74:102233. doi: 10.1016/j.media.2021.102233

13. Bhaumik R, Jenkins LM, Gowins JR, Jacobs RH, Barba A, Bhaumik DK, et al. Multivariate pattern analysis strategies in detection of remitted major depressive disorder using resting state functional connectivity. *Neuroimage Clin*. (2017) 16:390–8. doi: 10.1016/j.nicl.2016.02.018

14. Craddock RC, Holtzheimer III PE, Hu XP, Mayberg HS. Disease state prediction from resting state functional connectivity. *Mag Reson Med*. (2009) 62:1619–28. doi: 10.1002/mrm.22159

15. Zeng LL, Shen H, Liu L, Wang L, Li B, Fang P, et al. Identifying major depression using whole-brain functional connectivity: a multivariate pattern analysis. *Brain*. (2012) 135:1498–507. doi: 10.1093/brain/aws059

16. Guo S, Yu Y, Zhang J, Feng J. A reversal coarse-grained analysis with application to an altered functional circuit in depression. *Brain Behav*. (2013) 3:637–48. doi: 10.1002/brb3.173

17. Wei M, Qin J, Yan R, Li H, Yao Z, Lu Q. Identifying major depressive disorder using Hurst exponent of resting-state brain networks. *Psychiatry Res Neuroimaging*. (2013) 214:306–12. doi: 10.1016/j.pscychresns.2013.09.008

18. Yu Y, Shen H, Zeng LL, Ma Q, Hu D. Convergent and divergent functional connectivity patterns in schizophrenia and depression. *PLoS ONE*. (2013) 8:e68250. doi: 10.1371/journal.pone.0068250

19. Lois G, Wessa M. Differential association of default mode network connectivity and rumination in healthy individuals and remitted MDD patients. *Soc Cogn Affect Neurosci*. (2016) 11:1792–801. doi: 10.1093/scan/nsw085

20. Zhu X, Yuan F, Zhou G, Nie J, Wang D, Hu P, et al. Cross-network interaction for diagnosis of major depressive disorder based on resting state functional connectivity. *Brain Imaging Behav*. (2021) 15:1279–89. doi: 10.1007/s11682-020-00326-2

21. Li J, Chen H, Fan F, Qiu J, Du L, Xiao J, et al. White-matter functional topology: a neuromarker for classification and prediction in unmedicated depression. *Transl Psychiatry*. (2020) 10:1–10. doi: 10.1038/s41398-020-01053-4

22. Qin J, Shen H, Zeng LL, Jiang W, Liu L, Hu D. Predicting clinical responses in major depression using intrinsic functional connectivity. *Neuroreport*. (2015) 26:675–80. doi: 10.1097/WNR.0000000000000407

23. Geng X, Xu J, Liu B, Shi Y. Multivariate classification of major depressive disorder using the effective connectivity and functional connectivity. *Front Neurosci*. (2018) 12:38. doi: 10.3389/fnins.2018.00038

24. Guo H, Li Y, Xu Y, Jin Y, Xiang J, Chen J. Resting-state brain functional hyper-network construction based on elastic net and group lasso methods. *Front Neuroinform*. (2018) 12:25. doi: 10.3389/fninf.2018.00025

25. Wang J, Wei Q, Yuan X, Jiang X, Xu J, Zhou X, et al. Local functional connectivity density is closely associated with the response of electroconvulsive therapy in major depressive disorder. *J Affect Disord*. (2018) 225:658–64. doi: 10.1016/j.jad.2017.09.001

26. Sen B, Cullen KR, Parhi KK. Classification of adolescent major depressive disorder via static and dynamic connectivity. *IEEE J Biomed Health Inform*. (2020) 25:2604–14. doi: 10.1109/JBHI.2020.3043427

27. Yan B, Xu X, Liu M, Zheng K, Liu J, Li J, et al. Quantitative identification of major depression based on resting-state dynamic functional connectivity: a machine learning approach. *Front Neurosci*. (2020) 14:191. doi: 10.3389/fnins.2020.00191

28. Guo H, Yan P, Cheng C, Li Y, Chen J, Xu Y, et al. fMRI classification method with multiple feature fusion based on minimum spanning tree analysis. *Psychiatry Res Neuroimaging*. (2018) 277:14–27. doi: 10.1016/j.pscychresns.2018.05.001

29. Yoshida K, Shimizu Y, Yoshimoto J, Takamura M, Okada G, Okamoto Y, et al. Prediction of clinical depression scores and detection of changes in whole-brain using resting-state functional MRI data with partial least squares regression. *PLoS ONE*. (2017) 12:e0179638. doi: 10.1371/journal.pone.0179638

30. Guo H, Cheng C, Cao X, Xiang J, Chen J, Zhang K. Resting-state functional connectivity abnormalities in first-onset unmedicated depression. *Neural Regenerat Res*. (2014) 9:153. doi: 10.4103/1673-5374.125344

31. Ichikawa N, Lisi G, Yahata N, Okada G, Takamura M, Hashimoto Ri, et al. Primary functional brain connections associated with melancholic major depressive disorder and modulation by antidepressants. *Sci Rep*. (2020) 10:1–12. doi: 10.1038/s41598-020-73436-y

32. Jun E, Na KS, Kang W, Lee J, Suk HI, Ham BJ. Identifying resting-state effective connectivity abnormalities in drug-naïve major depressive disorder diagnosis via graph convolutional networks. *Hum Brain Mapp*. (2020) 41:4997–5014. doi: 10.1002/hbm.25175

33. Ktena SI, Parisot S, Ferrante E, Rajchl M, Lee M, Glocker B, et al. Metric learning with spectral graph convolutions on brain connectivity networks. *Neuroimage*. (2018) 169:431–42. doi: 10.1016/j.neuroimage.2017.12.052

34. Yao D, Sui J, Wang M, Yang E, Jiaerken Y, Luo N, et al. A mutual multi-scale triplet graph convolutional network for classification of brain disorders using functional or structural connectivity. *IEEE Trans Med Imaging*. (2021) 40:1279–89. doi: 10.1109/TMI.2021.3051604

35. Kong Y, Gao S, Yue Y, Hou Z, Shu H, Xie C, et al. Spatio-temporal graph convolutional network for diagnosis and treatment response prediction of major depressive disorder from functional connectivity. *Hum Brain Mapp*. (2021) 42:3922–33. doi: 10.1002/hbm.25529

36. Qin K, Lei D, Pinaya WH, Pan N, Li W, Zhu Z, et al. Using graph convolutional network to characterize individuals with major depressive disorder across multiple imaging sites. *EBioMedicine*. (2022) 78:103977. doi: 10.1016/j.ebiom.2022.103977

37. Shi S, Qiao K, Yang S, Wang L, Chen J, Yan B. Boosting-GNN: boosting algorithm for graph networks on imbalanced node classification. *Front Neurorob*. (2021) 15:154. doi: 10.3389/fnbot.2021.775688

38. Yan C, Zang Y. DPARSF: a MATLAB toolbox for" pipeline" data analysis of resting-state fMRI. *Front Syst Neurosci*. (2010) 4:13. doi: 10.3389/fnsys.2010.00013

39. Dosenbach NU, Nardos B, Cohen AL, Fair DA, Power JD, Church JA, et al. Prediction of individual brain maturity using fMRI. *Science*. (2010) 329:1358–61. doi: 10.1126/science.1194144

40. Achard S, Bullmore E. Efficiency and cost of economical brain functional networks. *PLoS Comput Biol*. (2007) 3:e17. doi: 10.1371/journal.pcbi.0030017

41. Guo M, Wang T, Zhang Z, Chen N, Li Y, Wang Y, et al. Diagnosis of major depressive disorder using whole-brain effective connectivity networks derived from resting-state functional MRI. *J Neural Eng*. (2020) 17:056038. doi: 10.1088/1741-2552/abbc28

42. Nakano T, Takamura M, Ichikawa N, Okada G, Okamoto Y, Yamada M, et al. Enhancing multi-center generalization of machine learning-based depression diagnosis from resting-state fMRI. *Front Psychiatry*. (2020) 11:400. doi: 10.3389/fpsyt.2020.00400

43. Yamashita A, Sakai Y, Yamada T, Yahata N, Kunimatsu A, Okada N, et al. Generalizable brain network markers of major depressive disorder across multiple imaging sites. *PLoS Biol*. (2020) 18:e3000966. doi: 10.1371/journal.pbio.3000966

44. Kawahara J, Brown CJ, Miller SP, Booth BG, Chau V, Grunau RE, et al. BrainNetCNN: convolutional neural networks for brain networks; towards predicting neurodevelopment. *NeuroImage*. (2017) 146:1038–49. doi: 10.1016/j.neuroimage.2016.09.046

45. Dvornek NC, Li X, Zhuang J, Duncan JS. Jointly discriminative and generative recurrent neural networks for learning from fMRI. In: *International Workshop on Machine Learning in Medical Imaging*. Springer (2019). p. 382–90.

46. Parisot S, Ktena SI, Ferrante E, Lee M, Guerrero R, Glocker B, et al. Disease prediction using graph convolutional networks: application to autism spectrum disorder and Alzheimer's disease. *Med Image Anal*. (2018) 48:117–30. doi: 10.1016/j.media.2018.06.001

47. Hultman R, Ulrich K, Sachs BD, Blount C, Carlson DE, Ndubuizu N, et al. Brain-wide electrical spatiotemporal dynamics encode depression vulnerability. *Cell*. (2018) 173:166–80. doi: 10.1016/j.cell.2018.02.012

48. Ramirez-Mahaluf JP, Roxin A, Mayberg HS, Compte A. A computational model of major depression: the role of glutamate dysfunction on cingulo-frontal network dynamics. *Cereb Cortex*. (2017) 27:660–79. doi: 10.1093/cercor/bhv249

frontiers | Frontiers in Psychiatry

TYPE Original Research
PUBLISHED 10 May 2023
DOI 10.3389/fpsyt.2023.1105534

OPEN ACCESS

EDITED BY
Zhi Xu,
Southeast University, China

REVIEWED BY
Ling Sun,
Tianjin Anding Hospital, China
Cao Qingjiu,
Peking University Sixth Hospital, China

*CORRESPONDENCE
Jihua Lu
✉ lujihua@bit.edu.cn
Fan He
✉ hf981207@163.com

RECEIVED 22 November 2022
ACCEPTED 14 April 2023
PUBLISHED 10 May 2023

CITATION
Luo J, Liu M, Feng L, Li Z, Wu Y, Lu J and He F (2023) Multidimensional voiceprint feature assessment system for identifying the depression in children and adolescents: a diagnostic test. *Front. Psychiatry* 14:1105534. doi: 10.3389/fpsyt.2023.1105534

Multidimensional voiceprint feature assessment system for identifying the depression in children and adolescents: a diagnostic test

Jie Luo[1], Mengqi Liu[1], Lihui Feng[2], Zhaojun Li[3], Yuanzhen Wu[1], Jihua Lu[3]* and Fan He[1]*

[1]National Clinical Research Center for Mental Disorders, Beijing Key Laboratory of Mental Disorders, Beijing Anding Hospital, Beijing Institute for Brain Disorders, Capital Medical University, Beijing, China, [2]Beijing Institute of Technology, School of Optics and Photonics, Beijing, China, [3]Beijing Institute of Technology, School of Integrated Circuits and Electronics, Beijing, China

Objective: We designed a diagnostic test to evaluate the effectiveness and accuracy of a multidimensional voiceprint feature diagnostic assessment (MVFDA) system vs. the 24-item Hamilton Rating Scale for Depression (HAMD-24) for adjunctive diagnosis of children and adolescents with major depressive disorder (MDD).

Methods: This study included 55 children aged 6–16 years who were clinically diagnosed with MDD according to the DSM-5 and analyzed by professional physicians, and 55 healthy children (typically developing). Each subject completed a voice recording and was scored on the HAMD-24 scale by a trained rater. We calculated the validity indices, including sensitivity, specificity, Youden's index, likelihood ratio, and other indices including predictive value, diagnostic odds ratio, diagnostic accuracy, and area under the curve (AUC), to assess the effectiveness of the MVFDA system in addition to the HAMD-24.

Results: The sensitivity (92.73 vs. 76.36%) and the specificity (90.91 vs. 85.45%) of the MVFDA system are significantly higher than those of the HAMD-24. The AUC of the MVFDA system is also higher than that of the HAMD-24. There is a statistically significant difference between the groups ($p < 0.05$), and both of them have high diagnostic accuracy. In addition, the diagnostic efficacy of the MVFDA system is higher than that of HAMD-24 in terms of the Youden index, diagnostic accuracy, likelihood ratio, diagnostic odds ratio, and predictive value.

Conclusion: The MVFDA has performed well in clinical diagnostic trials for the identification of MDD in children and adolescents by capturing objective sound features. Compared with the scale assessment method, the MVFDA system could be further promoted in clinical practice due to its advantages of simple operation, objective rating, and high diagnostic efficiency.

KEYWORDS

major depressive disorder, children and adolescents, multidimensional voiceprint feature assessment system, objective diagnosis, diagnostic accuracy

Introduction

Major depressive disorder (MDD) is a type of common mental illness. MDD could become the first-class diseases for the heavy burden worldwide by 2030 (1). However, the prevalence of depression tends to increase among adolescents. In the United States, the lifetime prevalence of MDD among adolescents aged 13 to 18 years is 11.0%, and the 12-month prevalence probability is 7.5% (2). In addition to the high probability of prevalence, the cure rate for adolescent depression is extremely low, and the suicide rate is quite high (3, 4). Moreover, as children and adolescents are in a crucial period of psychophysiological development, their depression has a more serious impact on the performance of both academic and social functions (5, 6). Therefore, early diagnosis and identification of MDD are essential.

There are two major diagnostic systems for depression: the DSM-5 (the Diagnostic and Statistical Manual of Mental Disorders) and the ICD (the International Classification of Diseases) systems. The recognition of depression depends on some basic clinical symptoms. In addition, clinicians will also assist in the diagnosis according to some classical scales, such as the HAMD (7), which can assess the severity of depression. The establishment of the HAMD depression scale includes some basic symptoms of depression, such as decreased interest and fatigue, and some additional symptoms like anxiety/somatization, cognition, and round-the-clock changes (7). However, these scales require rich clinical experiences (8, 9). Even senior physicians are prone to misdiagnosis when patients atypically conceal their illness or symptoms. Moreover, compared with adults, the clinical manifestations of children and adolescents with depressive disorders are more atypical, and the degree of coordination is lower, which makes it more difficult for clinicians, especially primary care medical workers, to identify adolescents with the depressive disorder (10–12). Considering the shortcomings of the scales, research on the objective adjunctive diagnosis of MDD patients has become a hot topic.

Patients with depression display different physiological indicators from healthy individuals, such as altered body posture, facial expressions, and voice. Researchers have extracted some specific features to diagnose depression. Studies show that the voice of patients with depression can change significantly; for instance, voice speed is extremely slow, and pauses are longer, more rigid, and more frequent (13–15). With the development of artificial intelligence (AI), it is easier to design Al-related algorithms based on the voice features of patients with depression that could help identify some classical symptoms of MDD (16). In addition, compared to other traits, sound acquisition is easier and cheaper (15). Therefore, researchers have shown great interest in speech recognition research (SDR) on depression. The current research on depression recognition employing speech includes two main aspects: (1) Analyzing the speech features of patients with depression. (2) Building a speech depression recognition model. Early studies mainly focused on the classical and related features of depression patients. Most studies analyzed the features of their speech in the time domain. Szabadi et al. found that the pause in the voice of depressed patients was prolonged, but the phonation period remained constant (17). Moreover, the length of their voice pause decreased with the improvement of clinical symptoms. Then, Greden and Carroll also confirmed Szabadi's conclusions (18). Hollien showed that patients with depression speak slower than healthy individuals, with a more monotonous intonation (19). However, the features extracted from the time domain are greatly affected by individual differences and cannot fully represent the features of patients with depression.

In addition, researchers have committed to investigating the changes in various acoustic features of speech signals, such as prosody, sound source, composition, and spectrum. The differences between the two types of people in the features of speech signals, such as fundamental frequency (F0), Mel cepstrum coefficient (MFCC), energy, frequency, and formant, have also been studied. Based on these features, a recognition model of depression has been constructed. In recent years, great advancements in SDR have been made possible thanks to the integration of acoustic features. Ooi et al. distinguished depressed adolescents from healthy individuals with a classification accuracy of 73% using four features (progressive, gross, TEO, and spectrum). The recognition effects of multi-channel classification with a weighted decision procedure are superior to all classifications based on a single feature or single channel (20). Cummins et al. combined features of MFCC and formant and used the GMM classifier to obtain an accuracy of 79% (21). Mendirata et al. obtained an accuracy of 80.67% using MFCC features and conducted principal component analysis (PCA), and clustering classification (22). These studies show that the designed recognition system based on the vocal differences of patients with depression has strong feasibility. For the experimental paradigms, the size of the data set and other aspects are difficult to compare horizontally with the results. Therefore, a general and optimal model with higher accuracy is very important. However, the conclusions drawn from studies on the same sound features are not consistent. For example, the experimental results of some studies show that the size of F0 is related to the severity of depression, while other experiments reveal that there is no correlation between F0 and depression (23). Even for the same group of subjects, the relationship between F0 and depression is also affected by the speech content. Therefore, finding a more accurate and scientific diagnostic system is very important. All told, the existing data sets are mostly focused on adults, and the studies are also inconsistent. There are fewer studies on the identification of adolescent speech than there are on adult speech. Due to the physical development of adolescents, the sound of children and adolescents is somewhat different from that of adults in terms of the frequency range, pitch, speech rate, fundamental frequency, resonance peak, and so on. For example, the vocal cords and throat structure of teenagers are not yet fully mature, so their voices tend to have a higher frequency range. In contrast, the voices of adults are more stable, with a relatively smaller frequency range. Teenagers' pitch is usually higher and sometimes less stable, while adults tend to have a more even and stable pitch. In addition, teenagers generally speak at a faster pace than adults, which may be related to their more active bodies and faster thinking ability. Therefore, to obtain a higher identification accuracy, it is essential to specifically recognize adolescent speech.

In response to the above problems, Beijing Anding Hospital and Beijing Institute of Technology have developed a new

multidimensional voice feature diagnostic assessment (MVFDA) system that overcomes the shortcomings of the state-of-the-art algorithms and greatly improves the recognition accuracy based on features extracted from voiceprint. Furthermore, we specifically evaluated the effectiveness of this diagnostic system and compared it with the HAMD clinical scales.

Methods and materials

Study design

Aiming at diagnosing and evaluating the effectiveness of the MVFDA system for MDD and comparing it with the HAMD scale, the design of the trial follows the STARD statement. PASS15 software (NCSS, Kaysville, Utah, USA) was used to calculate samples. Referring to previous related studies, it has been shown that machine learning models have promising results with AUC values generally greater than 0.9. Under the conditions of $\alpha = 0.05$ (unilateral), $\beta = 0.1$, and a 1:1 ratio between groups, 55 subjects were enrolled in each group of the MDD and healthy control (HC) groups. Each subject was evaluated and compared using the MVFDA system and HAMD scale.

Subjects

Subjects in the MDD group were inpatients at the Beijing Anding Hospital. The inclusion criteria were as follows: (a) meet the DSM-5 diagnostic criteria for MDD; (b) be between the ages of 6 and 16, with no gender restriction; (c) be able to cooperate to complete the study; and (e) sign the informed consent. Exclusion criteria were as follows: (a) severe physical illness, such as pharyngeal edema, pharyngeal foreign bodies, hoarseness; (b) comorbid other psychiatric disorders, such as bipolar disorder and schizophrenia, and developmental disorders (e.g., autism spectrum disorder, intellectual impairment); and (c) other conditions deemed inappropriate for inclusion in the group by the investigators. Healthy controls were recruited from the community. Typically developing children and adolescents aged 10 to 18 years with no other conditions, and who were able to cooperate to complete all the requirements were recruited from the community and schools. Ultimately, we successfully recruited 110 subjects according to the study plan, with no subjects dropping out midway through. The project was conducted in accordance with the ethical standards of the Declaration of Helsinki and its subsequent amendments and was approved by the Ethics Committee of the Beijing Institute of Technology (Ethical number: BIT-EC-H-2022120). All subjects and their family members signed an informed consent form prior to the trial. All subjects were able to comply with the MVFDA system and HAMD assessment requirements during the study, and the data collected were valid and reliable.

Gold standard

MDD is diagnosed according to the recommended guidelines and is based on the patient's medical history, clinical symptoms, disease course, and relevant examinations. In this study, two senior experts conducted detailed clinical interviews with each subject, obtained their medical history, performed a psychiatric evaluation, and combined these to form a diagnostic opinion according to the DSM-5 diagnostic criteria. Finally, the unanimous opinion of the two experts served as the gold standard for the complete diagnosis.

MVFDA system

This algorithm achieves the recognition of depression based on voice data. Specifically, as shown in Table 1. First, multi-dimensional features are extracted from voice data, including energy-related, spectral, voice-related, and statistical information. Low-level descriptors (LLDs) of voiceprints are manually designed and generally calculated from a frame of voice. Various statistical functions are then calculated based on the LLD. Then, a more comprehensive feature set is constructed, and features with significant differences between classes are retained through feature screening. The next step is to extract features from the transform domain to optimize their differences. Finally, the integrated classification method is used to achieve high classification accuracy.

HAMD scale

The most widely used scale for assessing depression was created in 1960. There are three versions of this scale: the 17-item, 21-item, and 24-item. In this study, we used the 24-item HAMD. The HAMD scale is administered and independently scored by two trained raters. Most of the items are rated on a 5-point scale (0 to 4), and a few items are rated on a 3-point scale (0 to 2). Items 8, 9, and 11 of the scale are rated based on observation of the patient; the remaining items are rated based on the patient's own verbal narrative; item 1 requires a combination of the two. In addition, for items 7 and 22, information has to be collected from the patient's family or ward staff, while item 16 is based on weight records and can also be rated on the basis of the patient's complaints and information provided by their family or ward staff. According to Davis JM's delineation score of 24 items total, major depressive disorder is possible with a score >20. The scale is widely used in clinical practice and has high reliability and validity (24–27). All interviewers had passed the HAMD-24 consistency training. Additionally, the group's intraclass correlation coefficient (ICC) was higher than 0.8.

Processing

We used the recording function of the MacBook Air 2020 to gather the voice data of all subjects. The data were collected in a quiet room at Beijing Anding Hospital to evaluate the MVFDA system. First, the entire evaluation process was briefly introduced and described. Then, the investigator read part of the story or some words to demonstrate. The recording was initiated once the participant comprehended the lists that needed to be read. Children were required to read one story and six groups of words in

TABLE 1 Description of the features applied in the MVFDA system.

Types of features		Names of features	Meaning
LLD	Energy-related LLD	Sum of the auditory spectrum	The auditory spectrum includes information from the time and frequency domains of a sound signal
		RMS energy	Root mean square value of all samples in a frame
		Zero crossing rate	Number of times the signal crosses the horizontal axis
		Sum of RASTA-style filtered auditory spectrum	Auditory spectrum after RASTA filtering
	Spectral LLD	Mel Frequency Cepstrum Coefficient (MFCC) Energy, variance, and kurtosis of the spectral	Coefficients of Mel frequency cepstrum Energy and variance of spectral, and kurtosis of each spectral line
	Voice-related LLD	Fundamental frequency (F0) Jitter and shimmer Harmonics-to-Noise ratio (HNR)	The frequency of the fundamental tone in a polyphony Basic frequency and amplitude change of acoustic wave between adjacent periods Ratio of harmonic to noise components
Statistical features		Mean, Maximum, Minimum Variance Linear regression slope	Average, Maximum, Minimum, and Variance of samples Slope of linear regression line

succession. For the purpose of feature extraction and classification, speech differences between depressed adolescents and healthy ones were applied. A 10-min break was taken after completing the MVFDA system assessment, and then two professional raters assessed them using the HAMD scale.

Statistical analysis

Data were analyzed using SPSS25.0 software (IBM, Armonk, NY, USA). Age differences between the two groups were analyzed by *t*-test, and gender differences were analyzed by chi-squared. Descriptive statistics were conducted to analyze the medication use of patients in the MDD group. The consistency of the results between the two HAMD scale raters (ML and JLuo) was determined by kappa analysis. A series of indicators, including sensitivity, specificity, the Youden index, diagnostic odds ratios, likelihood ratios, predictive values, and diagnostic accuracy, were calculated to evaluate the diagnostic validity of MVFDA compared with HAMD. Drawing the ROC curve, calculating the area under the curve (AUC), and applying Z-tests should all be performed to test for differences in the AUC, as it is generally believed that the closer the AUC is to 1, the more reliable the test method is.

Results

Subject characteristics

A total of 55 subjects, 20 boys and 35 girls, with a mean age of 14.40 ± 1.72 years, were included in the MDD group of this trial. A total of 55 subjects, 22 boys and 33 girls, with a mean age of 14.83 ± 1.50 years, were included in the HC group. There was no statistical difference in age ($p = 0.16$) or gender ($p = 0.43$) between the two groups. The medication use of the patients in the MDD group is shown in Table 2.

Diagnostic efficacy

The kappa analysis result was 0.964, indicating that the HAMD scale findings were reliable and had a good agreement.

TABLE 2 Medication use by patients in the MDD group.

Medication	MDD ($n = 55$)
Antidepressants, *n* (%)	55 (100.00%)
SSRI *(sertraline, escitalopram oxalate, fluvoxamine)*	45 (81.8%)
SNRI *(duloxetine)*	2 (3.6%)
NDRI *(bupropion)*	5 (9.1%)
Others *(vortioxetine, agomelatine)*	3 (5.5%)
Antipsychotics, *n* (%)	32 (58.2%)
Aripiprazole	11 (20.0%)
Lurasidone	7 (12.7%)
Quetiapine	8 (14.5%)
Others *(olanzapine, paliperidone, peropirox)*	6 (10.9%)
Mood stabilizers, *n* (%)	18 (32.7%)
Lithium	7 (12.7%)
Sodium valproate	6 (10.9%)
Lamotrigine	5 (9.1%)

SSRI, selective serotonin reuptake inhibitor; SNRI, serotonin-norepinephrine reuptake inhibitor; NDRI, norepinephrine-dopamine reuptake inhibitor.

TABLE 3 MVFDA system and HAMD results for all subjects.

	MVFDA system		HAMD	
	Positive (*n*)	Negative (*n*)	Positive (*n*)	Negative (*n*)
MDD group ($n = 55$)	51	4	43	12
HC group ($n = 55$)	5	50	8	47
Total ($n = 110$)	56	54	51	59

As shown in Table 3, in the MDD group, the MVFDA system's correct/incorrect diagnosis was 51/4, and the HAMD scale's correct/incorrect diagnosis was 43/12; in the control group, the MVFDA system's correct/incorrect diagnosis was 50/5, and the HAMD scale's correct/incorrect diagnosis was 47/8. The evaluation indices are shown in Table 4. The sensitivity (92.73 vs. 76.36%,

10.3389/fpsyt.2023.1105534

TABLE 4 Calculation results of the diagnostic efficacy of MVFDA and HAMD.

	Sensitivity (%)		Specificity (%)		Youden index	PLR		NLR		Diagnostic odds ratio	PPV (%)		NPV (%)		Diagnostic accuracy (%)
	Point value	95%CI	Point value	95%CI		Point value	95%CI	Point value	95%CI		Point value	95%CI	Point value	95%CI	
MVFDA	92.73	81.57–97.64	90.91	79.30–96.60	0.84	10.2	4.40–23.60	0.08	0.03–0.20	127.5	91.07	79.63–96.67	92.59	81.26–97.60	91.82
HAMD	76.36	70.70–91.80	85.45	77.07–95.49	0.62	5.25	3.57–16.46	0.28	0.10–0.34	18.98	80.77	75.97–95.22	81.03	72.07–92.23	80.91

NLR, negative likelihood ratio; NPV, negative predictive value; PLR, positive likelihood ratio; PPV, positive predictive value.

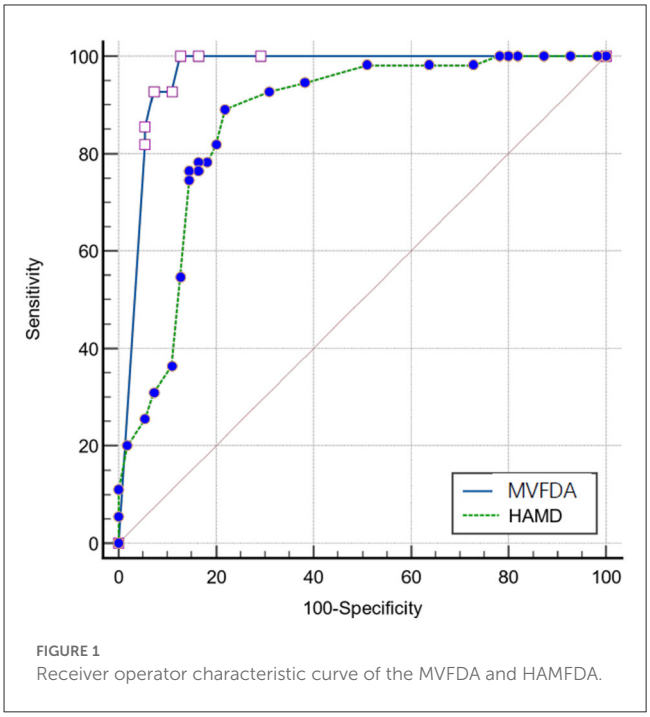


FIGURE 1
Receiver operator characteristic curve of the MVFDA and HAMFDA.

$P = 0.04$) and the specificity (90.91 vs. 85.45%, $P > 0.05$) of MVFDA system are significantly and slightly higher than those of the HAMD scale. However, the difference is not statistically significant. The Youden index, likelihood ratios, predictive value, and diagnostic accuracy of the MVFDA system are also higher than the corresponding items of the HAMD, as shown in Table 3. In addition, the ROC curves of MVFDA and HAMD are shown in Figure 1. The AUC of the MVFDA (0.962) was greater than that of the HAMD (0.962), and the difference between them was statistically significant ($p = 0.012$), indicating that the diagnostic efficacy of the MVFDA system is significantly higher than that of the HAMD.

Discussion

A novel voiceprint retrieval algorithm is proposed for the diagnosis of depression in children and adolescents. We found that the developed MVFDA system in this paper has exceptional diagnostic utility. Compared to the voiceprint system, the HAMD scale has lower sensitivity, likelihood ratios, predictive accuracy, and other indicators. This result is due to the limitations of the HAMD scale in its clinical application to children and adolescents with depressive disorders. The HAMD has the following drawbacks: (1) The evaluation takes a long time. The HAMD evaluation process takes at least 15–20min, it can be challenging for children and adolescents with depression to maintain steady attention during the interviews, which affects the quality of the evaluation (28). (2) The Hamilton Depression Scale raters are highly professional, necessitating professional, consistent training to ensure the quality of evaluation (29, 30). Compared with the typical depressive symptoms of adults, the clinical features of children and adolescents with depression are relatively complex.

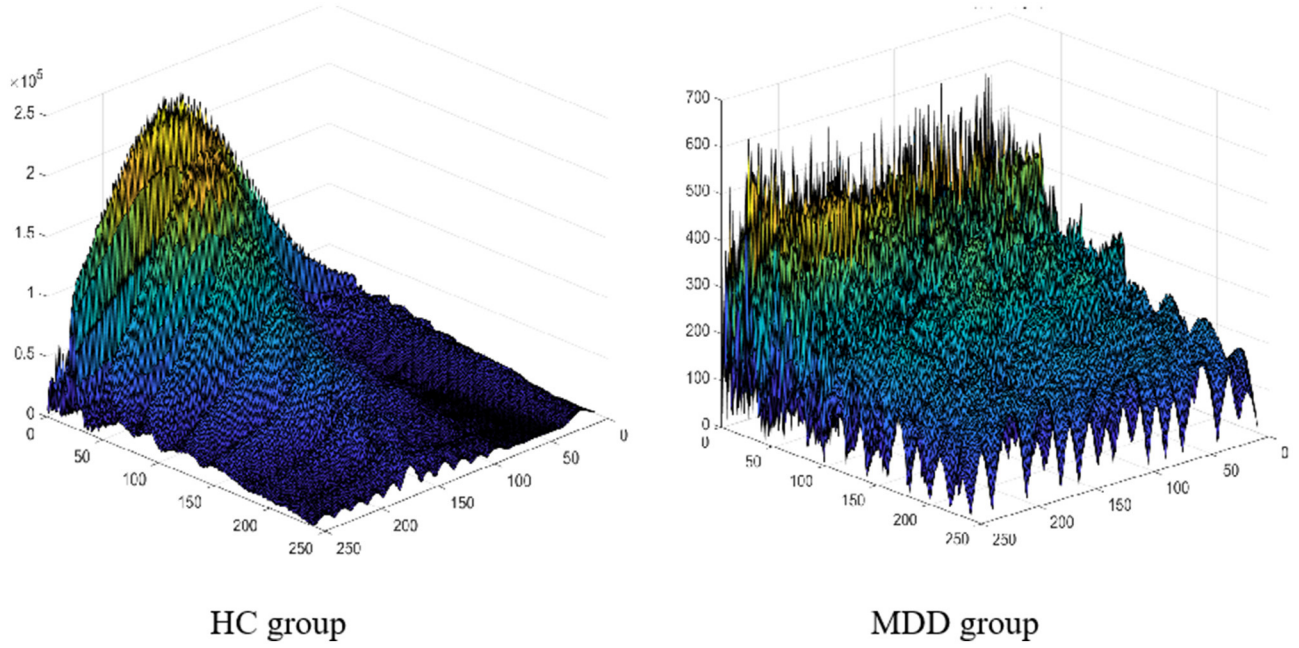


FIGURE 2
The feature visualization results of the HC and MDD groups.

For example, adolescents with depressive disorders may be more likely to have physical complaints, psychomotor agitation, anxiety, and other manifestations, and they might have trouble responding to the item "depressed" (31). For inexperienced evaluators, it can often be difficult to accurately assess the patient's disease status, which leads to misestimation. (3) The level of the patient's cooperation must be high. As an interview measure, the HAMD scale requires effective communication between the evaluator and patient to gather medical history and assess clinical symptoms. Children and adolescents often struggle to describe their own disease condition accurately (10). Communication skills are required to get the patient to cooperate with the scale's assessment. If the patient is unwilling to cooperate with the assessment due to unfamiliarity or other reasons, the assessment of HAMD will not proceed smoothly.

The inadequacies of the HAMD may also be present in other scales; however, the ability to recognize depressive disorders through speech effectively overcomes these problems. First, the assessment phase of the MVFDA system is short, lasting less than 5 min, which means children "and adolescents" attention remains relatively steady. Second, the operation is simple. The auxiliary staff only needs to guide the subjects correctly to record their voices without any formal training. Third, without complex cooperation in question and answer with raters, the patients only complete the test by reading aloud the required paradigm. The data collected by the test is directly extracted from the patients. Analyzing the characteristics of the disease-related voice ensures the objectivity of the evaluation and prevents the subjectivity of the rater's evaluation.

In addition, the research findings that we developed based on the proposed algorithm could have a high degree of diagnostic accuracy for children and adolescents. In previous studies, the voiceprint system was mostly built based on data from adults with depression (32). However, due to the differences between adults, children, and adolescents, the voiceprints used in previous research were not compatible with our collected data. Furthermore, the number of voiceprint features used in state-of-the-art studies is small. Also, few attributes could deliver good results for specific groups, or even specific environments of audio acquisition, who bears poor robustness (33). In previous studies, a single voice was used as a sample, so multiple voices of the same person were used for both training and testing. Because the association between the voices of the same person is ignored, the accuracy of the system recognition is relatively high; however, in the actual application scenario, the accuracy is drastically decreased (34). To solve the aforementioned issues, we created an effective data set specifically for children and adolescents, based on resources acquired from the hospital. Additionally, in terms of feature extraction, LLD and its statistical features related to energy, spectrum, and rhythm were used to obtain a more comprehensive feature set, breaking the defect of a single feature promoted by existing methods, which leads to the difficulty in extracting common features of patients with depression. In addition to the time and frequency domains, our algorithm is extended to the wavelet domain to maximize the difference between the two groups of people, as shown by the feature visualizations in Figure 2. To achieve effective feature screening, the proposed algorithm applies KS test combined with maximum information coefficient (MIC), where KS test is used to screen features with large differences between classes, and MIC is used to remove features with high correlation. Finally, based on sufficient samples, and taking into account the

correlation among the voices of the same person, the proposed algorithm treats a person as a sample when training the model, and the voice of the same person can only be applied for training or testing to achieve higher accuracy, which is more consistent with the actual application scenarios. Improvements in ensuring the effective identification of children and adolescents with depressive disorder.

The MVFDA system is the first generation; hence, it has some shortcomings. First of all, this study used a single device to capture speech, which may not reflect the impact of different devices on recording quality. Second, the dimensionality of the features extracted by the algorithm is not sufficiently streamlined. In the future, another dataset might be established where speech is collected using multiple recording devices and in different environments to investigate the impact of recording devices and environments on the MVFDA. Besides, further consideration is given to screening and reducing redundant features in addition to ensuring high classification accuracy by maintaining the features with the greatest differences between classes.

This test verified the scientific validity and reliability of the algorithm, and the results showed that it is very suitable for clinical diagnosis and application promotion. The detection system can reduce the need for a depression examination facility. Under the premise of solving privacy issues and other issues, the complete voiceprint evaluation system could be used by families, schools, and primary medical units. Voices are collected through a simple paradigm. Data are collected and transferred to the cloud devices to obtain a report, which is then sent to the doctors for guidance. Its convenience and efficiency are used for large-scale early screening of MDD to reduce the medical burden and greatly improve the diagnostic efficiency of clinicians.

Conclusions

The MVFDA system performed well in clinical diagnostic studies for identifying MDD in children and adolescents by capturing objective voiceprint features. The algorithm's scientific and dependable characteristics were verified. Simulations revealed that the sensitivity, the Youden Index, likelihood ratios, predictive value, and diagnostic accuracy of the MVFDA system were higher than those of the HAMD. The specificity of the MVFDA system was also slightly lower than that of the HAMD scale. Furthermore, considering the ROC and the AUC of the MVFDA system, the diagnostic efficacy of MVFDA is significantly higher than that of the HAMD. Compared with the scale assessment, the MVFDA system deserves to be further promoted in clinical practice for its advantages in terms of simple operation, objective evaluation, and high diagnostic efficiency.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Beijing Institute of Technology (Ethical Number: BIT-EC-H-2022120). Written informed consent to participate in this study was provided by the participants' legal guardian/next of kin.

Author contributions

JLuo, ML, LF, ZL, YW, JLu, and FH are involved in the collection of the data. JLuo performed the data analysis and wrote the final manuscript. JLu made substantial revisions to the first draft of the manuscript. ML, LF, and ZL were involved throughout the data collection process. JLu was primarily involved in the design of the MVFDA system. FH made practical contributions to the conduct and development of the trial. All authors contributed to the article and approved the submitted version.

Funding

This study was funded by the Beijing Hospitals Authority Clinical Medicine Development of Special Funding Support, code: ZYLX202128, by the Natural Science Foundation of China (Grant No. 61902379), by the China Academic Degrees and Graduate Education Association (Grant No. 2020MSA178), and by the NSFC with (Grant No. 61675025 of 2020).

Acknowledgments

The authors would like to thank all participants for their support.

Conflict of interest

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.

Publisher's note

References

1. The global burden of disease: 2004 update. Geneva: World Health Organization [EB/OL].

2. Avenevoli S, Swendsen J, He J-P, Burstein M, Merikangas KR. Major depression in the national comorbidity survey-adolescent supplement: prevalence, correlates, and treatment. *J Am Acad Child Adolesc Psychiatry.* (2015) 54:37–44.e2. doi: 10.1016/j.jaac.2014.10.010

3. Dwyer JB, Stringaris A, Brent DA, Bloch MH. Annual research review: defining and treating pediatric treatment-resistant depression. *J Child Psychol Psychiatry.* (2020) 61:312–32. doi: 10.1111/jcpp.13202

4. Asarnow JR, Baraff LJ, Berk M, Grob C, Devich-Navarro M, Suddath R, et al. Pediatric emergency department suicidal patients: two-site evaluation of suicide ideators, single attempters, and repeat attempters. *J Am Acad Child Adolesc Psychiatry.* (2008) 47:958–66. doi: 10.1097/CHI.0b013e3181799ee8

5. Weissman MM, Wolk S, Goldstein RB, Moreau D, Adams P, Greenwald S, et al. Depressed adolescents grown up. *JAMA.* (1999) 281:1707–13. doi: 10.1001/jama.281.18.1707

6. Fergusson DM, Woodward LJ. Mental health, educational, and social role outcomes of adolescents with depression. *Arch Gen Psychiatry.* (2002) 59:225–31. doi: 10.1001/archpsyc.59.3.225

7. Hamilton M. *The Hamilton Rating Scale for Depression*. In Sartorius N, Ban TA, editor. Assessment of Depression, Berlin, Heidelberg: Springer Berlin Heidelberg. 1986:143-152. doi: 10.1007/978-3-642-70486-4_14

8. Faries D, Herrera J, Rayamajhi J, DeBrota D, Demitrack M, Potter WZ. The responsiveness of the Hamilton depression rating scale. *J Psychiatr Res.* (2000) 34:3–10. doi: 10.1016/S0022-3956(99)00037-0

9. Bagby RM, Ryder AG, Schuller DR, Marshall MB. The Hamilton depression rating scale: has the gold standard become a lead weight? *Am J Psychiatry.* (2004) 161:2163–77. doi: 10.1176/appi.ajp.161.12.2163

10. Rice F, Riglin L, Lomax T, Souter E, Potter R, Smith DJ, et al. Adolescent and adult differences in major depression symptom profiles. *J Affect Disord.* (2019) 243:175–81. doi: 10.1016/j.jad.2018.09.015

11. Miller L, Campo JV. Depression in adolescents. *N Engl J Med.* (2021) 385:445–9. doi: 10.1056/NEJMra2033475

12. Malhi GS, Mann JJ. Depression. *Lancet.* (2018) 392:2299–312. doi: 10.1016/S0140-6736(18)31948-2

13. Pampouchidou A, Simantiraki O, Vazakopoulou C-M, Chatzaki C, Pediaditis M, Maridaki A, et al. Facial geometry and speech analysis for depression detection. *Annu Int Conf IEEE Eng Med Biol Soc.* (2017) 2017:1433–6. doi: 10.1109/EMBC.2017.8037103

14. Mundt JC, Vogel AP, Feltner DE, Lenderking WR. Vocal acoustic biomarkers of depression severity and treatment response. *Biol Psychiatry.* (2012) 72:580–7. doi: 10.1016/j.biopsych.2012.03.015

15. Wu P, Wang R, Lin H, Zhang F, Tu J, Sun M. Automatic depression recognition by intelligent speech signal processing: a systematic survey. *Caai Transactions on Intelligence Technology.* (2022). doi: 10.1049/cit2.12113

16. Graham S, Depp C, Lee EE, Nebeker C, Tu X, Kim H-C, et al. Artificial intelligence for mental health and mental illnesses: an overview. *Curr Psychiatry Rep.* (2019) 21:116. doi: 10.1007/s11920-019-1094-0

17. Szabadi E, Bradshaw CM, Besson JaO. Speech in depressive states. *BJPsych.* (1977) 131:109–10. doi: 10.1192/S0007125000013507

18. Greden JF, Carroll BJ. Decrease in speech pause times with treatment of endogenous depression. *Biol Psychiatry.* (1980) 15:575–87.

19. Hollien H. Vocal indicators of psychological stress. *Ann NY Acad Sci.* (1980) 347:47–72. doi: 10.1111/j.1749-6632.1980.tb21255.x

20. Ooi KEB, Lech M, Allen NB. Multichannel weighted speech classification system for prediction of major depression in adolescents. *IEEE Transactions on Biomedical Engineering.* (2013) 60:497–506. doi: 10.1109/TBME.2012.2228646

21. Cummins N, Epps J, Breakspear M, Goecke R. An investigation of depressed speech detection: features and normalization[C]. *12th Annual Conference of the International-Speech-Communication-Association.* (2011) 2011:2997–3000. doi: 10.21437/Interspeech.2011-750

22. Mendiratta A, Scibelli F, Esposito AM, Capuano V, Likforman-Sulem L, Maldonato MN, et al. Automatic Detection of Depressive States from Speech. In: *Multidisciplinary Approaches to Neural Computing*. Cham: Springer International Publishing (2018). p. 301–14. doi: 10.1007/978-3-319-56904-8_29

23. Hussenbocus AY, Lech M, Allen NB. Statistical differences in speech acoustics of major depressed and non-depressed adolescents. In: *9th International Conference on Signal Processing and Communication Systems (ICSPCS).* (2015). doi: 10.1109/ICSPCS.2015.7391781

24. Aben I, Verhey F, Lousberg R, Lodder J, Honig A. Validity of the beck depression inventory, hospital anxiety and depression scale, SCL-90, and Hamilton depression rating scale as screening instruments for depression in stroke patients. *Psychosomatics.* (2002) 43:386–93. doi: 10.1176/appi.psy.43.5.386

25. Pan S, Liu Z-W, Shi S, Ma X, Song W-Q, Guan G-C, et al. Hamilton rating scale for depression-24 (HAM-D(24)) as a novel predictor for diabetic microvascular complications in type 2 diabetes mellitus patients. *Psychiatry Res.* (2017) 258:177–83. doi: 10.1016/j.psychres.2017.07.050

26. Zhu G, Yin Y, Xiao C-L, Mao R-J, Shi B-H, Jie Y, et al. Serum DHEAS levels are associated with the development of depression. *Psychiatry Res.* (2015) 229:447–53. doi: 10.1016/j.psychres.2015.05.093

27. Sun XY, Li YX, Yu CQ, Li LM. Reliability and validity of depression scales of Chinese version: a systematic review. *Zhonghua Liu Xing Bing Xue Za Zhi.* (2017) 38:110–16. doi: 10.3760/cma.j.issn.0254-6450.2017.01.021

28. Müller MJ, Dragicevic A. Standardized rater training for the Hamilton Depression Rating Scale (HAMD-17) in psychiatric novices. *J Affect Disord.* (2003) 77:65–9. doi: 10.1016/S0165-0327(02)00097-6

29. Cicchetti DV, Prusoff BA. Reliability of depression and associated clinical symptoms. *Arch Gen Psychiatry.* (1983) 40:987–90. doi: 10.1001/archpsyc.1983.01790080069009

30. Demitrack MA, Faries D, Herrera JM, DeBrota D, Potter WZ. The problem of measurement error in multisite clinical trials. *Psychopharmacol Bull.* (1998) 34:19–24.

31. Loades ME, Clair MCS, Orchard F, Goodyer I, Reynolds S, I CONSORTIUM. Depression symptom clusters in adolescents: a latent class analysis in a clinical sample. *Psychother Res.* (2022) 32:860–73. doi: 10.1080/10503307.2022.2030498

32. Yalamanchili B, Kota NS, Abbaraju MS, Nadella VSS, Alluri SV. Real-time acoustic based depression detection using machine learning techniques[C]// (2020). *International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE).* (2020). doi: 10.1109/ic-ETITE47903.2020.394

33. Low L-SA, Maddage NC, Lech M, Sheeber LB, Allen NB. Detection of clinical depression in adolescents' speech during family interactions. *IEEE Trans Biomed Eng.* (2010) 58:574–86. doi: 10.1109/TBME.2010.2091640

34. Zhao Y, Liang Z, Du J, Zhang L, Liu C, Zhao L. Multi-Head attention-based long short-term memory for depression detection from speech. *Front Neurorobot.* (2021) 2021:111. doi: 10.3389/fnbot.2021.684037

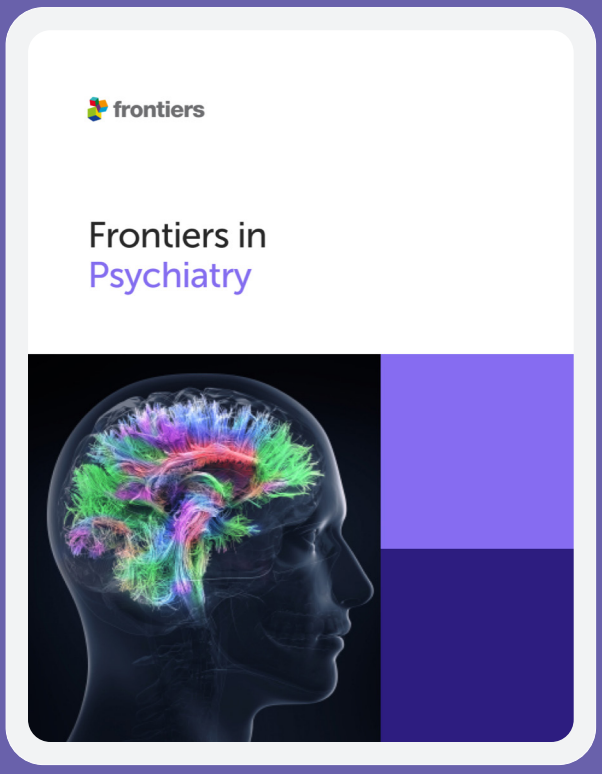
Frontiers in Psychiatry

Explores and communicates innovation in the field of psychiatry to improve patient outcomes

The third most-cited journal in its field, using translational approaches to improve therapeutic options for mental illness, communicate progress to clinicians and researchers, and consequently to improve patient treatment outcomes.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact