Distinguish bipolar and major depressive disorder in adolescents based on multimodal neuroimaging

Liu, Yujun; Chen, Kai; Luo, Yangyang; Wu, Jiqiu; Xiang, Qu; Peng, Li; Zhang, Jian; Zhao, Weiling; Li, Mingliang; Zhou, Xiaobo

Published in:
Digital Health

DOI:
10.1177/20552076221123705

Original Research

DIGITAL HEALTH

Digital Health
Volume 8: 1–12

DOI: 10.1177/20552076221123705
journals.sagepub.com/home/dhj
SAGE

Distinguish bipolar and major depressive disorder in adolescents based on multimodal neuroimaging: Results from the Adolescent Brain Cognitive Development study®

Yujun Liu[1,2], Kai Chen[3], Yangyang Luo[1,2], Jiqiu Wu[1,4], Qu Xiang[1,2], Li Peng[1,2], Jian Zhang[1,2], Weiling Zhao[5], Mingliang Li[1,2] and Xiaobo Zhou[5]

Abstract

Background: Major depressive disorder and bipolar disorder in adolescents are prevalent and are associated with cognitive impairment, executive dysfunction, and increased mortality. Early intervention in the initial stages of major depressive disorder and bipolar disorder can significantly improve personal health.

Methods: We collected 309 samples from the Adolescent Brain Cognitive Development study, including 116 adolescents with bipolar disorder, 64 adolescents with major depressive disorder, and 129 healthy adolescents, and employed a support vector machine to develop classification models for identification. We developed a multimodal model, which combined functional connectivity of resting-state functional magnetic resonance imaging and four anatomical measures of structural magnetic resonance imaging (cortical thickness, area, volume, and sulcal depth). We measured the performances of both multimodal and single modality classifiers.

Results: The multimodal classifiers showed outstanding performance compared with all five single modalities, and they are 100% for major depressive disorder versus healthy controls, 100% for bipolar disorder versus healthy control, 98.5% (95% CI: 95.4–100%) for major depressive disorder versus bipolar disorder, 100% for major depressive disorder versus depressed bipolar disorder and the leave-one-site-out analysis results are 77.4%, 63.3%, 79.4%, and 81.7%, separately.

Conclusions: The study shows that multimodal classifiers show high classification performances. Moreover, cuneus may be a potential biomarker to differentiate major depressive disorder, bipolar disorder, and healthy adolescents. Overall, this study can form multimodal diagnostic prediction workflows for clinically feasible to make more precise diagnose at the early stage and potentially reduce loss of personal pain and public society.

Keywords

Bipolar disorder, major depressive disorder, multimodal, support vector machine, cuneus

Submission date: 11 August 2022; Acceptance date: 16 August 2022

[1]West China Biomedical Big Data Center, West China Hospital, Sichuan University, Chengdu, China
[2]Med-X Center for Informatics, Sichuan University, Chengdu, China
[3]School of Public Health, University of Texas Health Science Center at Houston, Houston, USA
[4]Department of Genetics, University Medical Center Groningen, University of Groningen, Groningen, the Netherlands
[5]Center for Computational Systems Medicine, School of Biomedical Informatics, University of Texas Health Science Center at Houston, Houston, USA

Corresponding authors:
Xiaobo Zhou, University of Texas Health Science Center at Houston, Center for Computational Systems Medicine, School of Biomedical Informatics, 7000 Fannin Street, Houston, Texas 77030, USA.
Email: zhouxb2015@163.com

Yujun Liu, West China Biomedical Big Data Center, West China Hospital/West China School of Medicine, Sichuan University, No. 37 Guoxue Xiang, Chengdu 610041, China.
Email: liuyujun821@gmail.com

Introduction

Psychiatric disorders, such as major depressive disorder (MDD) and bipolar disorder (BD), are prevalent and detrimental in adolescents,[1–3] with a high degree of current and lifetime comorbidities partly due to common etiologies.[4,5] BD is one of the most severe and disabling psychiatric conditions affecting youth, significantly impairing an individual's ability to have relationships with family and friends, function at school, and cope with everyday life.[6] Clinically, the diagnosis of psychiatry disorders mainly relies on the clinician's subjective assessment of the patient's complaints and symptoms. Strikingly, it is estimated that approximately 69% of patients with BD are initially misdiagnosed as MDD,[7] which subsequently leads to inappropriate treatment, exacerbated manic symptoms, worse prognosis, increased health-care costs, and serious adverse events such as increased suicidality.[8,9] This is probably because depressive episodes are the most common mood manifestation, although BD consists of recurring episodes of mania/hypomania and depression.[10,11] Furthermore, childhood- or adolescent-onset are more severe than adult-onset, that is, more prolonged course, worse treatment effect, and more recurrences.[12] Unfortunately, children are often unable to articulate their feelings.[13] Hence, differentiating between BD and MDD in children presents a unique clinical and research challenge. Successfully classification of these two disorders will not only aid diagnostic decisions but also provides insight into their etiology and neuropathological processes.

Recently, researchers have been trying to explore promising neurobiological markers to help identify affective disorders. Accumulating evidence suggests brain magnetic resonance imaging (MRI) presented distinct structural and functional features between healthy individuals and patients with MDD and BD.[14–17] A study on distinguishing bipolar and major depressive disorders by brain structural morphometry demonstrated distinct spatially distributed variations in cortical thickness and surface area in patients with BD and MDD.[18] Compared to controls, a larger surface area in the temporo-parietal regions was observed in BD patients, and thinner cortices in fronto-temporal regions were observed in MDD patients, especially in the medial orbito-frontal area. Another study showed that gray matter volume in the right hippocampus, amygdala, parahippocampal, fusiform gyrus, and insula were reduced in BD patients compared with MDD. The volumes of the aforementioned regions and anterior cingulate cortex were also reduced in BD compared with healthy controls (HC).[19] Current neuroimaging studies are often based on functional connectivity (FC) of functional MRI (fMRI) or anatomical features of structural MRI (sMRI). Brain anatomical features are highly biological heterogeneity, complicating the functional connectivity of patients with psychiatric disorders,[20,21] and two modalities complement each other. The developed machine learning (ML) methods made it possible to translate neuroimaging findings from the research field to clinical practice for diagnosis by providing individual-level classification.[22] Support vector machine (SVM) is useful for high-dimensional and small sample data[23,24] and is one of the most widely used algorithms in diagnostic classification.[25] SVM was used to classify BD and HC with 100% accuracy.[26] In addition, The extensive implementation of ML algorithms on neuroimaging data has gradually identified some markers of psychiatric disorders. A meta-analysis explored the performance of different markers and ML algorithms in classifying several psychiatric disorders and found that the most discriminative features included gray and white matter alterations in the cortico-limbic region, and fMRI activations in this region during emotional tasks.[25] The combination of ML algorithms and neuroimaging data provides objective and reliable analysis, making a growing number of researchers recognize their contribution to psychiatric disorder diagnoses.[27]

Despite these efforts, research on direct comparisons of neuroimaging measures between BD and MDD in youth is sparse and has shown inconclusive and inconsistent findings. Firstly, many neuroimaging studies on MDD and BD patients enrolled adult populations, not youths. It may be related to the difficulty in obtaining neuroimaging data of children. Exploring the changes in brain structures and functions in adolescents with MDD and BD is critical for finding early disease biomarkers and taking timely intervention. Early attention to the abnormality of brain functions and structures in adolescents is of great significance for the study of brain development and affective disorders. Secondly, most previous studies used a single modality in classification tasks.[23,28] Previous reports have shown multimodal models can produce higher classification accuracy, allowing the exploration of the abnormalities of brain structures and functions from multiple perspectives in MDD and BD patients. Thirdly, most of the existing studies are small sample studies (with a sample size of less than 200), which is far from reaching reliable[29] and consistent[30] conclusions. Biomarkers obtained by analyzing the MRI data from only a dozen subjects are also unreliable and usually need to be validated on larger samples. Simultaneously, the sample size will also affect the performance of the model. The small sample size makes the model unable to be fully trained and easily leads to overfitting. A larger sample size will help improve the stability and robustness of the model, thereby enhancing the reliability of the conclusions drawn. Moreover, the images collected by a single site in most studies made the model generalization ability weak and not easy to be extended to new clinical sites.

To address this gap, the study aimed to develop a high-performance differential diagnosis classifier for MDD and BD based on the data from the Adolescent Brain Cognitive Development (ABCD) study. We developed

multimodal models covering five MRI modalities, including functional connectivity of fMRI and cortical thickness, area, volume, and sulcal depth of sMRI. We employed a modality-wise comparison of multimodal and single modality models to obtain the best classifiers. Finally, we discussed a biomarker that can distinguish adolescents with MDD and BD from healthy adolescents based on the best classifiers.

Methods

Dataset and participants

The data used in this study were obtained from the ABCD Data Release 3.0 (https://nda.nih.gov/abcd). The ABCD study is a large longitudinal study of the brain, behavioral development, and child health in the United States, recruiting more than 11,000 adolescent children aged between 9 and 10 years from 21 research sites across the USA.[31] Written and oral informed consents from parents and children have been obtained, respectively.[32] More information is provided at the ABCD website (https://abcdstudy.org) and elsewhere.[33]

In this study, a total of 309 adolescents neuroimaging and demographic data who completed baseline tasks, including 116 BD, 64 MDD, and 129 healthy adolescents, were recruited for this study. The diagnoses of BD and MDD complied with the ABCD youth Diagnostic Interview for Diagnostic and Statistical Manual of Mental Disorders, fifth edition (DSM-5) (abcd_ksad501). According to the ABCD Parent Diagnostic Interview for DSM-5 Background Items (KSADS-5), 10 and 3 adolescents with MDD and 17 and 9 adolescents with BD had a history of hospitalization and medication in this study, respectively. Only participants with current MDD or BD, whose imaging data meet MRI Quality Control criteria (abcd_imgincl01) entered the study. We excluded the subjects in case of the following conditions, serious neurological or psychiatric diagnosis, and any traumatic brain injury (TBI). Cognitive function was assessed by the ABCD Youth NIH TB Summary Scores (abcd_tbss01), which consist of 10 validated and reliable psychometric test scores.

Imaging data acquisition and processing

All imaging was collected using a 3T scanner and the neuroimaging parameters are available: https://abcdstudy.org/images/Protocol_Imaging_Sequences.pdf. The ABCD Data Analysis and Information Center (DAIC) performed a series of standard preprocessing pipelines. More details on rs-fMRI and sMRI preprocessing steps and parameters are described in the Supplemental Method.

Details of ABCD neuroimaging data acquisition and processing methods have been previously published.[33,34] For rs-fMRI data, the time courses of each region of interest (ROI) obtained by preprocessing are sampled onto the FreeSurfers' cortical surface to calculate the average time course of each ROI. We used 13 predefined networks (e.g. default, fronto-parietal, dorsal attention, etc.) within the Gordon parcellation to calculate the Pearson correlation value of each pair of ROIs, which are Fisher transformed to z-statistics and averaged within or between networks to provide summary measures of network correlation strength. For sMRI data, the Desikan atlas was used to segment cortical regions into 68 ROIs for the whole brain after cortical reconstruction,[35] and the morphological values were measured including cortical thickness, area, volume, and sulcal depth. Finally, a total of 363 features were used as input data for the classification task, including 91 functional connectivity of rs-fMRI (13 within-network and 78 between-network) and 68 brain regions in every modality of sMRI.

Feature selection

Feature selection is necessary to avoid overfitting and the influence of redundant noise.[36–38] The recursive feature elimination (RFE) algorithm based on SVM was used to implement this process. RFE is an effective feature selection algorithm that uses the accuracy yielded by SVM to determine which features contribute most to the prediction results.[37,39,40] Training on the original, each feature was assigned a weight coefficient, and those features with the most negligible absolute weight were kicked out of the feature set. After multiple iterations, the features of small weight coefficients were removed until the remaining features reached optimal performance.

The feature selection process was embedded into the main prediction pipeline. The features required to achieve the highest accuracy via 10-fold stratified cross-validation were used as input to the model. This process was achieved by characterizing the learning curve of SVM-RFE, which shows the accuracy corresponding to the number of selected features by the weight vector of 10-fold stratified cross-validation. As a comparison, we measured model prediction performance without RFE.

SVM classification and evaluation

Synthetic Minority Oversampling Technique with Tomek link (SMOTETomek),[41–43] a method that combines over- and under-sampling, was implemented to correct the impacts of imbalanced data. We selected the linear kernel of SVM,[24,44] which could yield the feature importance to the classification. In addition, there is only one parameter C in linear kernel SVM,[43] which reflects the tolerance of the model to the classification error and is used to balance between the correct classification of training samples and the maximization of the margin of the decision boundary. In this study, we set the hyper-parameter C to the default value of 1, as previous reported.[43–45]

Table 1. Demographic and clinical characteristics of participants.

	BD ($n = 116$)	MDD ($n = 64$)	HC ($n = 129$)	*P*-value
Demographic variables				
Age, years	9.94 ± 0.63	9.89 ± 0.67	9.81 ± 0.55	0.28[a]
Female	51 (43.97)	24 (37.50)	67 (51.94)	0.14[b]
Left-handed	10 (8.62)	4 (6.25)	7 (5.43)	0.60[b]
Standard intelligence Score	89.17 ± 10.87	87.16 ± 12.02	91.46 ± 11.00	**<0.05**[a]
Ethnicity				**<0.05**[b]
White	51 (43.97)	24 (37.50)	68 (52.71)	
Black	32 (27.59)	22 (34.38)	19 (14.73)	
Hispanic	20 (17.24)	11 (17.19)	30 (23.26)	
Asian	-	-	3 (2.33)	
Other	13 (11.21)	7 (10.94)	9 (6.98)	
Clinical variables				
Self-report				
Hospitalization history	17 (14.66)	10 (15.63)	-	**<0.05**[b]
Medication history	9 (7.76)	3 (4.69)	-	0.63[b]
BD-I	40 (34.48)	-	-	
manic	22 (18.97)	-	-	
depressed	17 (14.67)	-	-	
hypomanic	1 (0.86)	-	-	
BD-II	22 (18.97)	-	-	
depressed	18 (15.52)	-	-	
hypomanic	4 (3.45)	-	-	
BD-NOS	54 (46.55)	-	-	
MDD	-	-	-	
Present	-	43 (67.19)	-	
partial remission	-	21 (32.81)	-	

Data are expressed as *n* (%) or mean ± standard deviations. Bold represents statistically significant differences.
BD: bipolar disorder; BD-I: bipolar I disorder; BD-II: bipolar II disorder; BD-NOS: unspecified bipolar and related disorder; HC: healthy controls; MDD: major depressive disorder.
[a]*P*-value was obtained from ANOVA analysis.
[b]*P*-value was obtained from χ^2-test

Model performances were measured with standard metrics including accuracy, sensitivity, specificity, AUC (Area Under Curve), F1-score, positive predictive value (PPV), and negative predictive value (NPV), which are measured by 10-fold stratified cross-validation. The images in this study are from multiple sites, thus we perform a leave-one-site-out (LOSO) analysis embedded with 10-fold stratified cross-validation to test the site effects of 21 sites to verify the model's generalization. Specifically, an entire site was excluded from each round and the remaining 20 sites were used as the training set for feature selection. The hold-out site as the test set is used to calculate the performance by the important features selected from the training set to test the generalizability of these features in the new site. Site effect is considered significant if the 95% CI for the performance of any hold-out site does not overlap with other sites.[46,47]

Statistical analysis

A linear model was employed to correct the feature values of fMRI and sMRI. This process was implemented in R statistical software version 3.6.1 (https://cran.r-project.org). All continuous variables were normalized to the standard normal distribution to avoid the influence of non-uniform dimensions. The aforementioned procedures were implemented on Python version 3.8.6 (https://www.python.org). Analysis of variance and χ^2 test were used for continuous and categorical measures, respectively. A two-sided *P*-value less than 0.05 was considered statistically significant.

Results

Demographic and clinical characteristics

Table 1 lists the demographic and clinical characteristics of the three groups. Our data analysis included 64 MDD adolescents, 116 BD adolescents, and 129 healthy controls. Data in the study were baseline data at enrollment. Participants did not differ significantly in age, gender, hand dominance, and medication history, but there were significant differences in standard intelligence scores, ethnicity, and hospitalization history. Given the significant difference in standard intelligence scores, ethnicity, and hospitalization history, we regressed these variables, along with other potentially confounding variables (including gender, age, and medication history) from neuroimaging features by linear models to eliminate confounding effects.[48,49]

Feature importance

The optimal number of input features is 19–64 for five single modality classifiers achieving the highest accuracy, and more features are required, 34–102 features for multimodal classifiers, see Figure 1 and Supplemental Figure 1. The feature selection process significantly improved prediction performance.

The top-ten features of the multimodal SVM classifier are shown in Figure 2, while others are in Supplemental Tables 2 and 3. Cuneus is the only overlapping region among the top-ten features in the multimodal classifier, which was also found when distinguishing depressed BD patients from depressed MDD patients. This indicates that cuneus may be a potential biomarker to differentiate MDD, BD, and healthy adolescents.

The feature selection improved the performances of all modality classifiers, see Figure 3. The accuracies were significantly improved from 51.8–80.0% to 98.5–100% in multimodal classifiers and increased moderately from 45.4–74.6% to 66.5–86.9% in five single modalities, as shown in Supplemental Tables 4 and 5.

Comparing the performance of all classifiers based on feature selection, multimodal classifiers yielded the

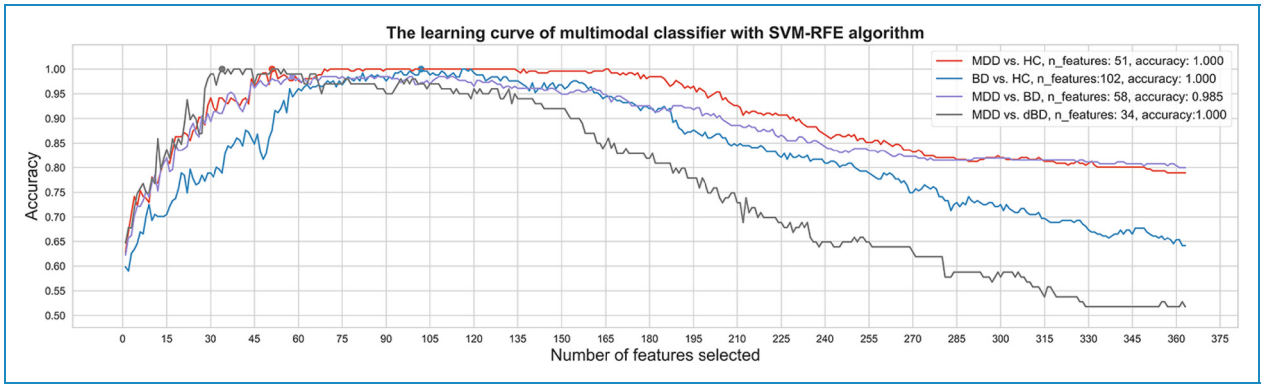

Figure 1. The learning curves of the multimodal classifiers in the 10-fold stratified cross-validation SVM-RFE algorithm. The *x*-axis shows the numbers of selected features, and the *y*-axis shows the corresponding accuracy.
BD: bipolar disorder; dBD: depressed bipolar disorder; MDD: major depressive disorder; HC: healthy controls; SVM-RFE: support vector machine recursive feature elimination.

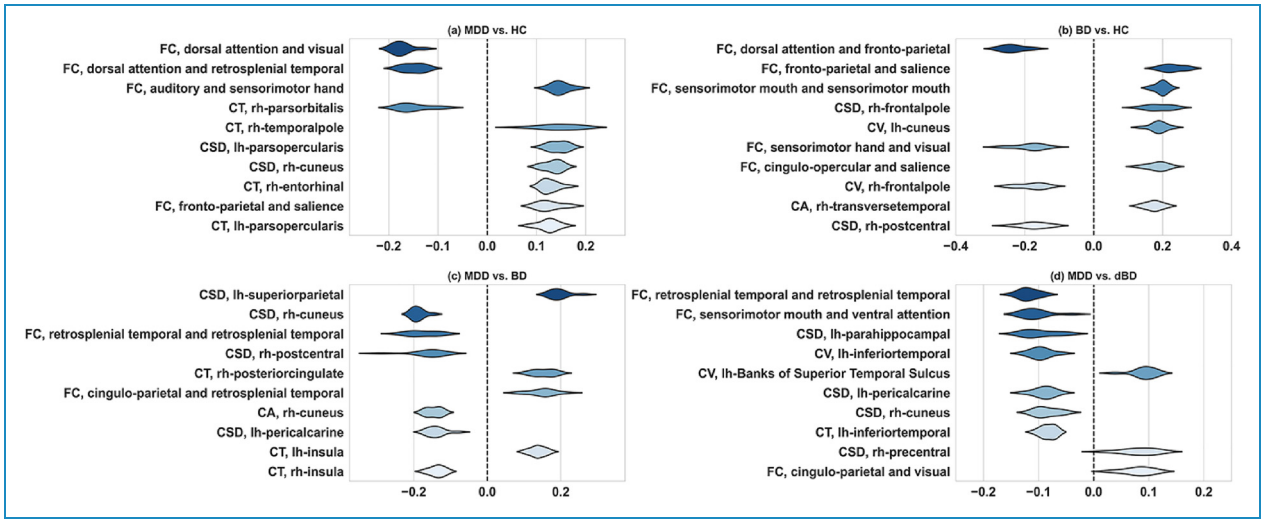

Figure 2. The top-ten features of multimodal classifiers in different classification tasks. The *x*-axis represents the feature's weight coefficient, and the *y*-axis represents the feature name. The average weight coefficient of each feature is determined by 10-fold cross-validation in SVM. Positive and negative coefficients represent positive and negative correlations. The higher the absolute value of the feature coefficient, the more significant impact of the feature on the model classification results.
BD: bipolar disorder; dBD: depressed bipolar disorder; MDD: major depressive disorder; HC: healthy controls; FC: functional connectivity; CV: cortical volume; CT: cortical thickness; CSD: cortical sulcal depth; CA: cortical area; lh: left hemispheric; rh: right hemispheric; SVM: support vector machine.

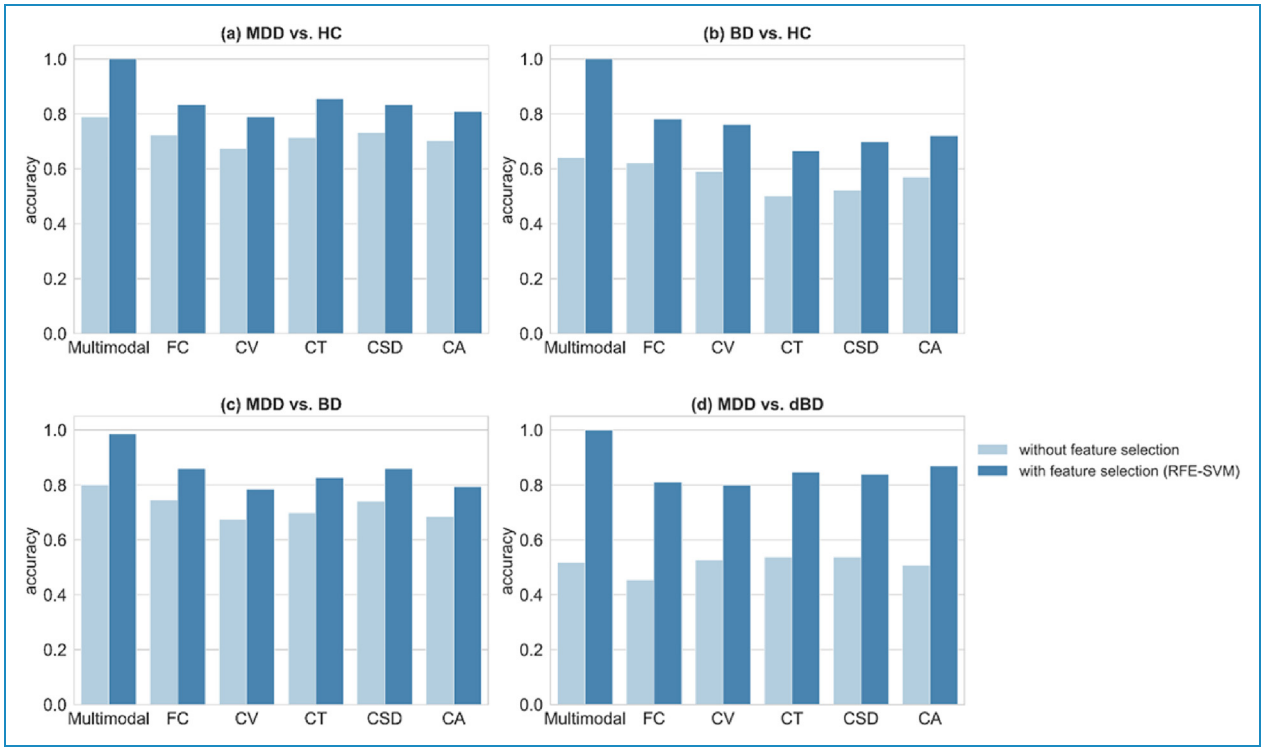

Figure 3. Performance comparison between without and with feature selection.
BD: bipolar disorder; dBD: depressed bipolar disorder; MDD: major depressive disorder; HC: healthy controls.

Table 2. Performance comparison between multimodal and single modality classifiers based on RFE-SVC feature selection.

		Single modality				
	Multimodal	FC/rs-fMRI	CV/sMRI	CT/sMRI	CSD/sMRI	CA/sMRI
MDD versus HC	1.0 (1.0–1.0)	0.833 (0.783–0.885)	0.79 (0.736–0.838)	0.856 (0.808–0.902)	0.832 (0.786–0.879)	0.809 (0.752–0.859)
BD versus HC	1.0 (1.0–1.0)	0.781 (0.742–0.817)	0.761 (0.696–0.825)	0.665 (0.624–0.698)	0.697 (0.661–0.735)	0.721 (0.661–0.78)
MDD versus BD	0.985 (0.954–1.0)	0.859 (0.801–0.918)	0.784 (0.715–0.854)	0.827 (0.769–0.887)	0.859 (0.815–0.901)	0.793 (0.716–0.871)
MDD versus dBD	1.0 (1.0–1.0)	0.81 (0.71–0.9)	0.799 (0.72–0.88)	0.847 (0.783–0.907)	0.839 (0.758–0.92)	0.869 (0.809–0.93)

Data are classification accuracy (95% confidence intervals generated from 1000 bootstrap samples).
rs-fMRI: resting-state functional magnetic resonance imaging; sMRI: structural magnetic resonance imaging; BD: bipolar disorder; dBD: depressed bipolar disorder; MDD: major depressive disorder; HC: healthy controls; FC: functional connectivity; CV: cortical volume; CT: cortical thickness; CSD: cortical sulcal depth; CA: cortical area.

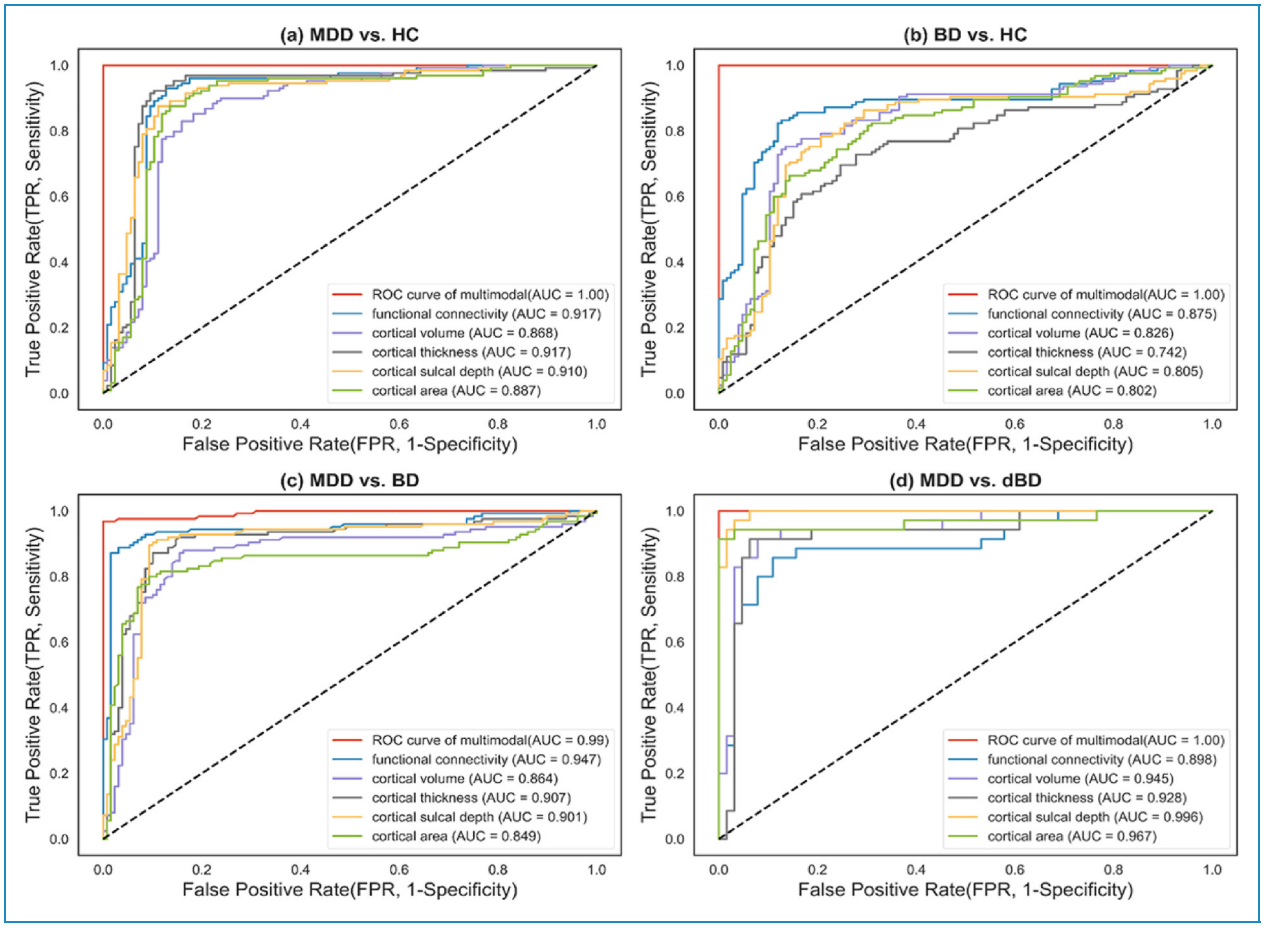

Figure 4. Receiver operating characteristic curves diagram of multimodal and single modality classifiers.
BD: bipolar disorder; dBD: depressed bipolar disorder; MDD: major depressive disorder; HC: healthy controls.

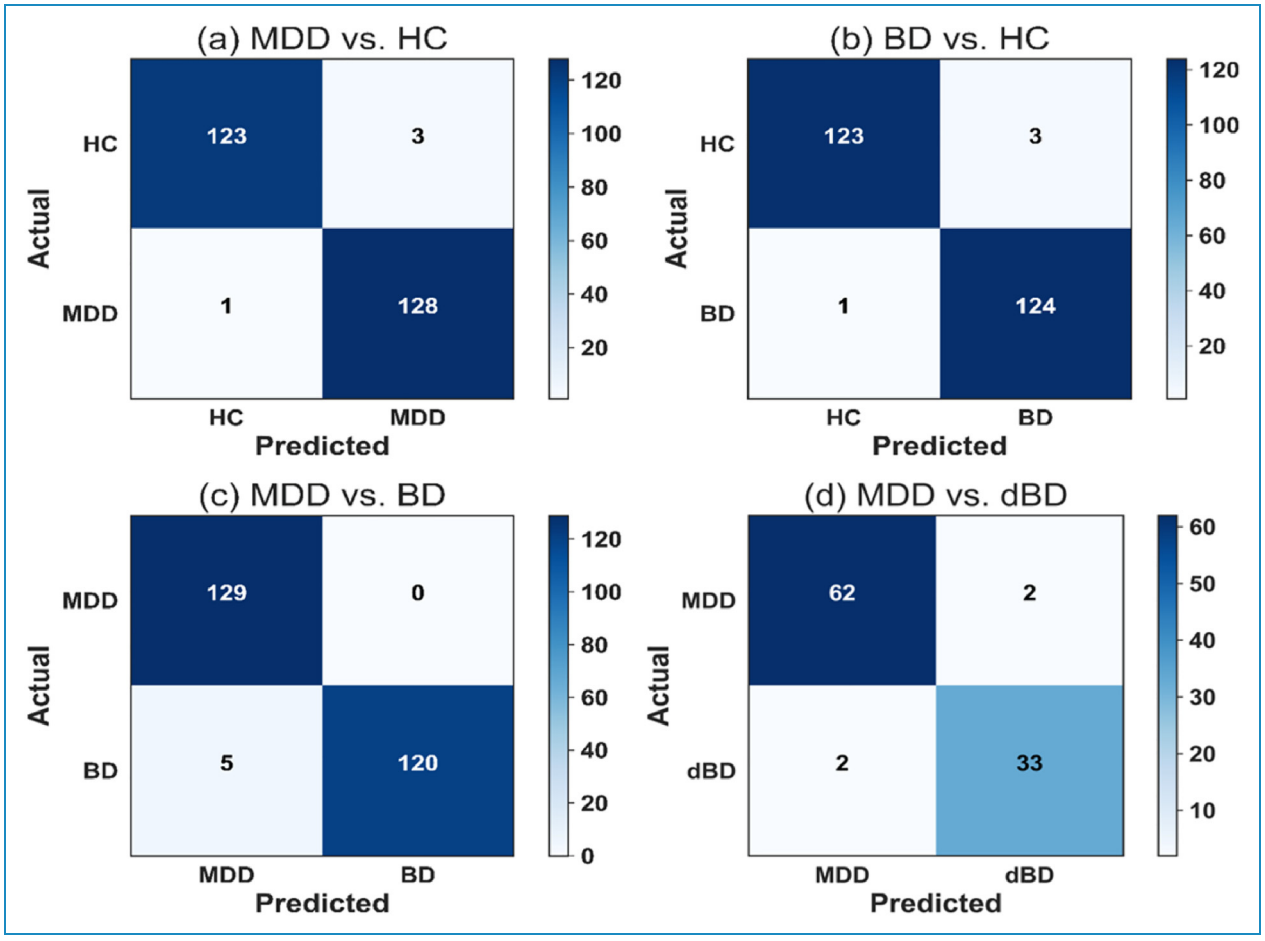

Figure 5. Confusion matrix for multimodal classifiers.
BD: bipolar disorder; dBD: depressed bipolar disorder; MDD: major depressive disorder; HC: healthy controls.

highest scores than others (Table 2). The multimodal performances greatly exceed the chance level of 50%, as well as the clinically relevant threshold of 80%.[44,50,51] The receive operating characteristic (ROC) curves and the area under the curve (AUC) are presented in Figure 4, and the confusion matrix is illustrated in Figure 5.

Sensitivity analyses

Finally, the LOSO analysis results are shown in Supplemental Figure 2. The site effect was not significant at 21 sites.

Discussion

The classification models used in this study to distinguish MDD, BD, and healthy adolescents were developed by 10-fold stratified cross-validation based on RFE-SVM. The accurate classifications of psychiatric disorders in adolescents based on neuroimaging are mainly attributed to the effect of multimodal. We regressed the confounding effects of latent variables using linear models and used the SMOTETomek algorithm to balance the sample size between groups to ensure fair and reliable results. Moreover, we found differences in the predictor required by several classifiers for accurate diagnosis. ML brings neuroimaging analyses to individual subject levels and allows prognostic formulation and treatment prescriptions to be tailored to the individual profile.

Accumulated MRI research studies have gradually improved the performance of machine learning algorithms in adult psychiatric disorder classification. Furthermore, the application of multimodal features has greatly improved model performance because multimodal features can provide multi-view information more than a single modality. A study achieved a high accuracy of 92.1% for distinguishing MDD and BD by using a multimodal method that combines the fractional amplitude of low-frequency fluctuations (fALFFs) of rs-fMRI and the gray matter volume of sMRI.[24] However, this conclusion was based on adult data and has not been validated in children and adolescents. Our multimodal models were constructed based on the adolescent neuroimaging data and obtained superior performance.

The functional and morphological abnormalities of the MDD and BD patients may be reflected in multiple brain regions and morphologies, so the classification performances of five single modality classifiers are slightly inferior. A study used the single modality of sMRI (gray matter volume) to classify BD and MDD with 75% accuracy.[52] Another study used SVM of Elastic Net feature selection to classify MDD and HC on functional connectivity, and obtained 76.1% accuracy.[53] Our study achieved slightly higher performance in the same modality and classification task; that is, 78.4% accuracy uses cortical volume to distinguish MDD and BD, and 83.3% accuracy uses fMRI to distinguish MDD and HC. But it should also be noted that our study focused on populations and brain regions, unlike the above two studies. Grotegerd et al.[45] used fMRI to distinguish between unipolar and bipolar depression (bipolar: 10, unipolar: 10, control: 10) with nearly 90% accuracy. However, the small sample size makes the model unable to be fully trained and easily leads to overfitting, which in turn reduces the stability of the model and the reliability of the conclusion.

In terms of single modality, we did not find any single modality classifier to overperform other classifiers in classifying MDD, BD, and HC. Claude et al.[54] speculated that the classification performances based on fMRI might outshine sMRI. In distinguishing between MDD and BD patients, six studies using sMRI as the input features obtained accuracies ranging from 54.76% to 75.9%,[18,39,52,55–57] compared with 67% to 93.1% in nine fMRI-based studies.[18,45,58–64]

Cuneus may be a potential predictor, which to some certain extent sheds light on the physiological mechanism of MDD and BD. We found that the weight coefficients for cuneus were considerably high for all classifiers. The cuneus is associated with inhibitor control and motor response,[65] which is part of the default mode network (DMN), and its abnormal activation is related to depression. The greater activation of the right cuneus cortex at baseline is associated with a more remarkable improvement in the symptoms of depression and anxiety.[66,67] A global meta-analysis confirmed that the activation of the right cuneus was increased in patients with MDD.[68] Some evidence suggests that individuals with or at risk for depression are associated with structural alterations in the cuneus.[69–71] Previous studies have reported that cuneus gray matter volume is associated with depression[72] and suicidal behavior[73,74] in MDD adults. Kim et al.[75] found that MDD adolescents have a thinner cortical thickness in the cuneus. In addition, two positron emission tomography (PET) studies both show that the metabolic activity of the right cuneus is positively correlated with the severity of apathy,[76–78] which is defined as a negative emotion with decreased feeling, and interest and attention.[77] Recognizing biomarkers can more timely and accurately diagnose MDD and BD, help individualized treatment and treatment response prediction, and provide the possibility for distinguishing subtypes of MDD and BD.[3]

Multimodal features provide rich brain functional and anatomical information, enabling the model to perform well. It is also vital that RFE-SVM select the best predictive features for multiple classification models. In addition, we used LOSO to internally validate the development of several multimodal models. We achieved about 70% accuracy in the LOSO analysis and close to 100% in the aggregate 10-fold cross-validation. In our study, perhaps due to the lack of prior knowledge of the test site,[79] LOSO classification accuracy is somewhat lower than that of all site's aggregate classification. Similarly, an ENIGMA study, which collected MRI data from 13 sites, yielded 58.67% mean accuracy in the LOSO analyses and 65.23% in aggregate 10-fold cross-validation in classifying bipolar and healthy participants.[43] Importantly, our research focuses on adolescents, which is conducive to the early identification of potential high-risk MDD and BD individuals and is essential for early diagnosis and treatment during childhood or adolescence.

This study has limitations and also potential for expansion. Our findings are based primarily on cortical regions, excluding subcortical regions, such as the hippocampus, amygdala, and thalamus, which have been identified as being associated with psychiatric disorders.[80,81] The sample size of this study was modest, and further large-high-quality training datasets from multi-site and external validation are required to improve the generalization of the diagnostic model. Moreover, as a complex disease, multi-omics studies are needed to consider the inclusion of more dimensional predictors such as demographic and clinical variables, genes, blood, and cognition.

Conclusion

In this adolescent study, we combined the functional connectivity of fMRI and four anatomical measurement values of sMRI (cortical thickness, area, volume, and sulcal depth) to build multimodal models. On this basis, we have developed several classification models with high accuracy and good generalization, which can form multimodal diagnostic prediction workflows for clinically feasible and individual-level predictions in MDD and BD adolescents.

Acknowledgements: Data used in the preparation of this article were obtained from the Adolescent Brain Cognitive Development[SM] (ABCD) Study (https://abcdstudy.org), held in the NIMH Data Archive (NDA). This is a multisite, longitudinal study designed to recruit more than 10,000 children aged 9–10 and follow them over 10 years into early adulthood. The ABCD Study® is supported by the National Institutes of Health and additional federal partners under award numbers U01DA041048, U01DA050989, U01DA051016, U01DA041022, U01DA051018, U01DA051037, U01DA050987, U01DA041174, U01DA041106, U01DA041117, U01DA041028, U01DA041134, U01DA050988,

U01DA051039, U01DA041156, U01DA041025, U01DA041120, U01DA051038, U01DA041148, U01DA041093, U01DA041089, U24DA041123, U24DA041147. A full list of supporters is available at https://abcdstudy.org/federal-partners.html. A listing of participating sites and a complete listing of the study investigators can be found at https://abcdstudy.org/consortium_members/. ABCD consortium investigators designed and implemented the study and/or provided data but did not necessarily participate in the analysis or writing of this report. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or ABCD consortium investigators.

Contributorship: YJL and LP researched literature and conceived the study. YJL and KC involved in data analysis. YJL, YYL, JQW, KC, QX, LP, MLL, and JZ involved in clinical investigation and data curation. YJL wrote the first draft of the manuscript. WLZ, YYL, JQW, KC, and JZ involved manuscript proofreading. XBZ supervised overall implementation. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Declaration of conflicting interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethics approval: The study was reviewed and approved by the West China Hospital of Sichuan University Institutional Review Board.

Funding: The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This work was supported by the 1.3.5 project for disciplines of excellence–Clinical Research Incubation Project, West China Hospital, Sichuan University, Center of Excellence-International Collaboration Initiative Grant, West China Hospital, Sichuan University (grant nos. 2019HXFH022, 139170052).

Guarantor: YJL.

ORCID iD: Yujun Liu https://orcid.org/0000-0002-5818-2543

Supplemental material: Supplemental material for this article is available online.

References

1. Cichon L, Janas-Kozik M, Siwiec A, et al. Clinical picture and treatment of bipolar affective disorder in children and adolescents. *Psychiatr Pol* 2020; 54: 35–50.
2. Zhou X, Liu L, Lan X, et al. Polyunsaturated fatty acids metabolism, purine metabolism and inosine as potential independent diagnostic biomarkers for major depressive disorder in children and adolescents. *Mol Psychiatry* 2019; 24: 1478–1488.
3. Lee H-J, Kim SH and Lee MS. *Understanding mood disorders in children*. Singapore: Springer, 2019, pp. 251–261.
4. Rice F, Riglin L, Lomax T, et al. Adolescent and adult differences in major depression symptom profiles. *J Affect Disord* 2019; 243: 175–181.
5. Benazzi F. Bipolar disorder-focus on bipolar II disorder and mixed depression. *The Lancet* 2007; 369: 935–945.
6. Bilska K, Pawlak J, Kapelski P, 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
7. Hirschfeld RM, Calabrese JR, Weissman MM, et al. Screening for bipolar disorder in the community. *J Clin Psychiatry* 2003; 64: 53–59.
8. Kelberman C, Biederman J, Green A, et al. Differentiating bipolar disorder from unipolar depression in youth: a systematic literature review of neuroimaging research studies. *Psychiat Res-Neuroim* 2021; 307: 111201. DOI: ARTN 111201 10.1016/j.pscychresns.2020.111201
9. de Almeida JRC and Phillips ML. Distinguishing between unipolar depression and bipolar depression: current and future clinical and neuroimaging perspectives. *Biol Psychiat* 2013; 73: 111–118.
10. Hirschfeld RM. Differential diagnosis of bipolar disorder and major depressive disorder. *J Affect Disord* 2014; 169: S12–S16.
11. Hashimoto K. Metabolomics of major depressive disorder and bipolar disorder: overview and future perspective. *Adv Clin Chem* 2018; 84: 81–99.
12. Dunn V and Goodyer IM. Longitudinal investigation into childhood- and adolescence-onset depression: psychiatric outcome in early adulthood. *Br J Psychiatry* 2006; 188: 216–222.
13. Lee HJ, Kim SH and Lee MS. Understanding mood disorders in children. *Adv Exp Med Biol* 2019; 1192: 251–261.
14. Bani-Fatemi A, Tasmim S, Graff-Guerrero A, et al. Structural and functional alterations of the suicidal brain: an updated review of neuroimaging studies. *Psychiatry Res Neuroimaging* 2018; 278: 77–91.
15. Bracht T, Jones DK, Muller TJ, et al. Limbic white matter microstructure plasticity reflects recovery from depression. *J Affect Disord* 2015; 170: 143–149.
16. Chang M, Womer FY, Edmiston EK, et al. Neurobiological commonalities and distinctions among three Major psychiatric diagnostic categories: a structural MRI study. *Schizophr Bull* 2018; 44: 65–74.
17. He H, Sui J, Du Y, et al. Co-altered functional networks and brain structure in unmedicated patients with bipolar and major depressive disorders. *Brain Struct Funct* 2017; 222: 4051–4064.
18. Fung G, Deng Y, Zhao Q, et al. Distinguishing bipolar and major depressive disorders by brain structural morphometry: a pilot study. *BMC Psychiatry* 2015; 15: 98.
19. Vai B, Parenti L, Bollettini I, et al. Predicting differential diagnosis between bipolar and unipolar depression with multiple kernel learning on multimodal structural neuroimaging. *Eur Neuropsychopharmacol* 2020; 34: 28–38.
20. Gao S, Calhoun VD and Sui J. Machine learning in major depression: from classification to treatment outcome prediction. *CNS Neurosci Ther* 2018; 24: 1037–1052.
21. He Y, Wang J, Wang L, et al. Uncovering intrinsic modular organization of spontaneous brain activity in humans. *PLoS One* 2009; 4: e5226.

22. Squarcina L, Castellani U, Bellani M, et al. Classification of first-episode psychosis in a large cohort of patients using support vector machine and multiple kernel learning techniques. *Neuroimage* 2017; 145: 238–245.
23. Li H, Cui L, Cao L, et al. Identification of bipolar disorder using a combination of multimodality magnetic resonance imaging and machine learning techniques. *BMC Psychiatry* 2020; 20: 88.
24. Jie NF, Zhu MH, Ma XY, et al. Discriminating bipolar disorder from Major depression based on SVM-FoBa: efficient feature selection with multimodal brain imaging data. *IEEE Trans Auton Ment Dev* 2015; 7: 320–331.
25. Colombo F, Calesella F, Mazza MG, et al. Machine learning approaches for prediction of bipolar disorder based on biological, clinical and neuropsychological markers: a systematic review and meta-analysis. *Neurosci Biobehav Rev* 2022; 135: 104552.
26. Besga A, Termenon M, Grana M, et al. Discovering Alzheimer's disease and bipolar disorder white matter effects building computer aided diagnostic systems on brain diffusion tensor imaging features. *Neurosci Lett* 2012; 520: 71–76.
27. Teixeira AL, Salem H, Frey BN, et al. Update on bipolar disorder biomarker candidates. *Expert Rev Mol Diagn* 2016; 16: 1209–1220.
28. Achalia R, Sinha A, Jacob A, et al. A proof of concept machine learning analysis using multimodal neuroimaging and neurocognitive measures as predictive biomarker in bipolar disorder. *Asian J Psychiatr* 2020; 50: 101984.
29. Delvecchio G, Fossati P, Boyer P, et al. Common and distinct neural correlates of emotional processing in bipolar disorder and Major depressive disorder: a voxel-based meta-analysis of functional magnetic resonance imaging studies. *Eur Neuropsychopharmacol* 2012; 22: 100–113.
30. Han KM, De Berardis D, Fornaro M, et al. Differentiating between bipolar and unipolar depression in functional and structural MRI studies. *Prog Neuropsychopharmacol Biol Psychiatry* 2019; 91: 20–27.
31. Casey BJ, Cannonier T, Conley MI, et al. The adolescent brain cognitive development (ABCD) study: imaging acquisition across 21 sites. *Dev Cogn Neurosci* 2018; 32: 43–54.
32. Auchter AM, Hernandez Mejia M, Heyser CJ, et al. A description of the ABCD organizational structure and communication framework. *Dev Cogn Neurosci* 2018; 32: 8–15.
33. Hagler DJJr., Hatton S, Cornejo MD, et al. Image processing and analysis methods for the adolescent brain cognitive development study. *Neuroimage* 2019; 202: 116091.
34. Thompson WK, Barch DM, Bjork JM, et al. The structure of cognition in 9 and 10 year-old children and associations with problem behaviors: findings from the ABCD study's baseline neurocognitive battery. *Dev Cogn Neurosci* 2019; 36: 100606.
35. Desikan RS, Segonne F, Fischl B, et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest *Neuroimage* 2006; 31: 968–980.
36. Chu S-H, Lenglet C, Schreiner MW, et al. Anatomical biomarkers for adolescent major depressive disorder from diffusion weighted imaging using SVM classifier. 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2018.
37. Sartori JM, Reckziegel R, Passos IC, et al. Volumetric brain magnetic resonance imaging predicts functioning in bipolar disorder: a machine learning approach. *J Psychiatr Res* 2018; 103: 237–243.
38. Salvador R, Radua J, Canales-Rodriguez EJ, et al. Evaluation of machine learning algorithms and structural features for optimal MRI-based diagnostic prediction in psychosis. *PLoS One* 2017; 12: e0175683.
39. Sacchet MD, Livermore EE, Iglesias JE, et al. Subcortical volumes differentiate major depressive disorder, bipolar disorder, and remitted major depressive disorder. *J Psychiatr Res* 2015; 68: 91–98.
40. Saeys Y, Inza I and Larranaga P. A review of feature selection techniques in bioinformatics. *Bioinformatics* 2007; 23: 2507–2517.
41. Chawla NV, Bowyer KW, Hall LO, et al. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res* 2002; 16: 321–357.
42. Lemaitre G, Nogueira F and Aridas CK. Imbalanced-learn: a python toolbox to tackle the curse of imbalanced datasets in machine learning. *J Mach Learn Res* 2017; 18: 1–5.
43. Nunes A, Schnack HG, Ching CRK, et al. Using structural MRI to identify bipolar disorders – 13 site machine learning study in 3020 individuals from the ENIGMA bipolar disorders working group. *Mol Psychiatry* 2020; 25: 2130–2143.
44. Hajek T, Cooke C, Kopecek M, et al. Using structural MRI to identify individuals at genetic risk for bipolar disorders: a 2-cohort, machine learning study. *J Psychiatry Neurosci* 2015; 40: 316–324.
45. Grotegerd D, Suslow T, Bauer J, et al. Discriminating unipolar and bipolar depression by means of fMRI and pattern classification: a pilot study. *Eur Arch Psychiatry Clin Neurosci* 2013; 263: 119–131.
46. Gevers S, Majoie CB, van den Tweel XW, et al. Acquisition time and reproducibility of continuous arterial spin-labeling perfusion imaging at 3T. *AJNR Am J Neuroradiol* 2009; 30: 968–971.
47. Almeida JRC, Greenberg T, Lu H, et al. Test-retest reliability of cerebral blood flow in healthy individuals using arterial spin labeling: findings from the EMBARC study. *Magn Reson Imaging* 2018; 45: 26–33.
48. Squarcina L, Perlini C, Peruzzo D, et al. The use of dynamic susceptibility contrast (DSC) MRI to automatically classify patients with first episode psychosis. *Schizophr Res* 2015; 165: 38–44.
49. Squarcina L, Dagnew TM, Rivolta MW, et al. Automated cortical thickness and skewness feature selection in bipolar disorder using a semi-supervised learning method. *J Affect Disord* 2019; 256: 416–423.
50. Savitz JB, Rauch SL and Drevets WC. Clinical application of brain imaging for the diagnosis of mood disorders: the current state of play. *Mol Psychiatry* 2013; 18: 528–539.
51. Thies B, Truschke E, Morrison-Bogorad M, et al. Consensus report of the working group on: molecular and biochemical markers of Alzheimer's disease. *Neurobiol Aging* 1999; 20: 47.
52. Rubin-Falcone H, Zanderigo F, Thapa-Chhetry B, et al. Pattern recognition of magnetic resonance imaging-based

gray matter volume measurements classifies bipolar disorder and major depressive disorder. *J Affect Disord* 2018; 227: 498–505.
53. Bhaumik R, Jenkins LM, Gowins JR, et al. Multivariate pattern analysis strategies in detection of remitted major depressive disorder using resting state functional connectivity. *Neuroimage Clin* 2017; 16: 390–398.
54. Claude LA, Houenou J, Duchesnay E, et al. Will machine learning applied to neuroimaging in bipolar disorder help the clinician? A critical review and methodological suggestions. *Bipolar Disord* 2020; 22: 334–355.
55. Matsuo K, Harada K, Fujita Y, et al. Distinctive neuroanatomical substrates for depression in bipolar disorder versus Major depressive disorder. *Cereb Cortex* 2019; 29: 202–214.
56. Redlich R, Almeida JJ, Grotegerd D, et al. Brain morphometric biomarkers distinguishing unipolar and bipolar depression. A voxel-based morphometry-pattern classification approach. *JAMA Psychiatry* 2014; 71: 1222–1230.
57. Serpa MH, Ou Y, Schaufelberger MS, et al. Neuroanatomical classification in a population-based sample of psychotic major depression and bipolar I disorder with 1 year of diagnostic stability. *Biomed Res Int* 2014; 2014: 706157.
58. Almeida JR, Mourao-Miranda J, Aizenstein HJ, et al. Pattern recognition analysis of anterior cingulate cortex blood flow to classify depression polarity. *Br J Psychiatry* 2013; 203: 310–311.
59. Li M, Das T, Deng W, et al. Clinical utility of a short resting-state MRI scan in differentiating bipolar from unipolar depression. *Acta Psychiatr Scand* 2017; 136: 288–299.
60. Rive MM, Redlich R, Schmaal L, et al. Distinguishing medication-free subjects with unipolar disorder from subjects with bipolar disorder: state matters. *Bipolar Disord* 2016; 18: 612–623.
61. Gao S, Osuch EA, Wammes M, et al. Discriminating bipolar disorder from major depression based on kernel SVM using functional independent components. IEEE International workshop on machine learning for signal processing. Tokyo, 2017, p. 1–6.
62. Burger C, Redlich R, Grotegerd D, et al. Differential abnormal pattern of anterior cingulate gyrus activation in unipolar and bipolar depression: an fMRI and pattern classification approach. *Neuropsychopharmacology* 2017; 42: 1399–1408.
63. Grotegerd D, Stuhrmann A, Kugel H, et al. Amygdala excitability to subliminally presented emotional faces distinguishes unipolar and bipolar depression: an fMRI and pattern classification study. *Hum Brain Mapp* 2014; 35: 2995–3007.
64. Mourao-Miranda J, Almeida JR, Hassel S, et al. Pattern recognition analyses of brain activation elicited by happy and neutral faces in unipolar and bipolar depression. *Bipolar Disord* 2012; 14: 451–460.
65. Matthews SC, Simmons AN, Arce E, et al. Dissociation of inhibition from error processing using a parametric inhibitory task during functional magnetic resonance imaging. *Neuroreport* 2005; 16: 755–760.
66. Greicius MD, Supekar K, Menon V, et al. Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cereb Cortex* 2009; 19: 72–78.
67. Marwood L, Wise T, Perkins AM, et al. Meta-analyses of the neural mechanisms and predictors of response to psychotherapy in depression and anxiety. *Neurosci Biobehav Rev* 2018; 95: 61–72.
68. Zhang WN, Chang SH, Guo LY, et al. The neural correlates of reward-related processing in major depressive disorder: a meta-analysis of functional magnetic resonance imaging studies. *J Affect Disord* 2013; 151: 531–539.
69. Qiu L, Lui S, Kuang W, et al. Regional increases of cortical thickness in untreated, first-episode major depressive disorder. *Transl Psychiatry* 2014; 4: e378.
70. Yang J, Zhang M, Ahn H, et al. Development and evaluation of a multimodal marker of major depressive disorder. *Hum Brain Mapp* 2018; 39: 4420–4439.
71. Dotson VM, Bogoian HR, Gradone AM, et al. Subthreshold depressive symptoms relate to cuneus structure: thickness asymmetry and sex differences. *J Psychiatr Res* 2021; 145: 144–147.
72. Kroes MC, Rugg MD, Whalley MG, et al. Structural brain abnormalities common to posttraumatic stress disorder and depression. *J Psychiatry Neurosci* 2011; 36: 256–265.
73. Overs BJ, Roberts G, Ridgway K, et al. Effects of polygenic risk for suicide attempt and risky behavior on brain structure in young people with familial risk of bipolar disorder. *Am J Med Genet B Neuropsychiatr Genet* 2021; 186: 485–507.
74. Giakoumatos CI, Tandon N, Shah J, et al. Are structural brain abnormalities associated with suicidal behavior in patients with psychotic disorders? *J Psychiatr Res* 2013; 47: 1389–1395.
75. Kim JH, Suh SI, Lee HJ, et al. Cortical and subcortical gray matter alterations in first-episode drug-naive adolescents with major depressive disorder. *Neuroreport* 2019; 30: 1172–1178.
76. Dujardin K and Defebvre L. Apathy in Parkinson's disease: what are the underlying mechanisms? *Neurology* 2012; 79: 1082–1083.
77. Robert G, Le Jeune F, Lozachmeur C, et al. Apathy in patients with Parkinson disease without dementia or depression: a PET study. *Neurology* 2012; 79: 1155–1160.
78. Voon V, Howell NA and Krack P. Psychiatric considerations in deep brain stimulation for Parkinson's disease. *Brain Stimul* 2013; 116: 147–154.
79. Zhao J, Huang J, Zhi D, et al. Functional network connectivity (FNC)-based generative adversarial network (GAN) and its applications in classification of mental disorders. *J Neurosci Methods* 2020; 341: 108756.
80. Ho NF, Li Hui Chong P, Lee DR, et al. The amygdala in schizophrenia and bipolar disorder: a synthesis of structural MRI, diffusion tensor imaging, and resting-state functional connectivity findings. *Harv Rev Psychiatry* 2019; 27: 150–164.
81. Skatun KC, Kaufmann T, Brandt CL, et al. Thalamo-cortical functional connectivity in schizophrenia and bipolar disorder. *Brain Imaging Behav* 2018; 12: 640–652.