

ARTICLE



Predicting treatment outcomes in major depressive disorder using brain magnetic resonance imaging: a meta-analysis

Fenghua Long^{1,2,4}, Yufei Chen^{1,2,4}, Qian Zhang^{1,2,4}, Qian Li^{1,2}, Yaxuan Wang^{1,2}, Yitian Wang^{1,2}, Haoran Li^{1,2}, Youjin Zhao^{1,2}, Robert K. McNamara³, Melissa P. DelBello³, John A. Sweeney^{1,3}, Qiyong Gong^{1,2} and Fei Li^{1,2}✉

© The Author(s), under exclusive licence to Springer Nature Limited 2024

Recent studies have provided promising evidence that neuroimaging data can predict treatment outcomes for patients with major depressive disorder (MDD). As most of these studies had small sample sizes, a meta-analysis is warranted to identify the most robust findings and imaging modalities, and to compare predictive outcomes obtained in magnetic resonance imaging (MRI) and studies using clinical and demographic features. We conducted a literature search from database inception to July 22, 2023, to identify studies using pretreatment clinical or brain MRI features to predict treatment outcomes in patients with MDD. Two meta-analyses were conducted on clinical and MRI studies, respectively. The meta-regression was employed to explore the effects of covariates and compare the predictive performance between clinical and MRI groups, as well as across MRI modalities and intervention subgroups. Meta-analysis of 13 clinical studies yielded an area under the curve (AUC) of 0.73, while in 44 MRI studies, the AUC was 0.89. MRI studies showed a higher sensitivity than clinical studies (0.78 vs. 0.62, $Z = 3.42$, $P = 0.001$). In MRI studies, resting-state functional MRI (rsfMRI) exhibited a higher specificity than task-based fMRI (tbfMRI) (0.79 vs. 0.69, $Z = -2.86$, $P = 0.004$). No significant differences in predictive performance were found between structural and functional MRI, nor between different interventions. Of note, predictive MRI features for treatment outcomes in studies using antidepressants were predominantly located in the limbic and default mode networks, while studies of electroconvulsive therapy (ECT) were restricted mainly to the limbic network. Our findings suggest a promise for pretreatment brain MRI features to predict MDD treatment outcomes, outperforming clinical features. While tasks in tbfMRI studies differed, those studies overall had less predictive utility than rsfMRI data. Overlapping but distinct network-level measures predicted antidepressants and ECT outcomes. Future studies are needed to predict outcomes using multiple MRI features, and to clarify whether imaging features predict outcomes generally or differ depending on treatments.

Molecular Psychiatry; <https://doi.org/10.1038/s41380-024-02710-6>

INTRODUCTION

Major depressive disorder (MDD) is a highly prevalent and disabling condition [1, 2]. Established treatments for MDD include psychological interventions such as cognitive behavioral therapy (CBT) and interpersonal psychotherapy, antidepressant medications, and somatic non-pharmacological treatments including electroconvulsive therapy (ECT), repetitive transcranial magnetic stimulation (rTMS), and direct current stimulation. However, treatment outcomes for MDD patients are highly variable and have been shown to be influenced by variables including age [3], sex [4], disease duration, and symptom severity [5]. It is difficult to predict using clinical and demographic features, and approximately 30–50% of patients with MDD do not respond to first-line medication or psychotherapy [6]. Therefore, treatment selection often begins with a “trial and error” approach, with weeks or months long trials until an effective and well-tolerated treatment is found. Consequently, several studies have investigated the potential of pretreatment features to guide personalized medicine approaches that can speed optimal treatment selection and positive clinical outcomes.

Multiple features have been evaluated for predicting treatment outcomes including clinical [7] and neuroimaging features [8]. A previous study demonstrated that utilizing pretreatment clinical features, such as the scores of apparent sadness, reported sadness, and inability to feel in the Montgomery-Åsberg Depression Rating Scale (MADRS), can successfully predict the ECT treatment outcomes [9]. The pretreatment scores of the Beck Depression Inventory (BDI), neuroticism, extraversion, depression, anxiety, and stress can also predict treatment outcomes of rTMS [10]. Clinical features, including baseline symptom severity, suicidality, and appetite changes, and demographic features such as age, sex, and ethnicity were also significant predictors for predicting antidepressant treatment outcomes [11]. However, the predictive accuracy using pretreatment clinical features varied from 44.3% to 94.3% (Table S6), and clinical heterogeneities between studies are important obstacles in the generalizability of diagnostic models.

Neuroimaging using magnetic resonance imaging (MRI) employs non-invasive techniques to evaluate brain anatomy and function whose predictive utility can be optimized using machine

¹Department of Radiology and Huaxi MR Research Center (HMRC), Functional and Molecular Imaging Key Laboratory of Sichuan Province, West China Hospital, Sichuan University, Chengdu 610041 Sichuan Province, PR China. ²Research Unit of Psychoradiology, Chinese Academy of Medical Sciences, Chengdu 610041 Sichuan, PR China. ³Department of Psychiatry and Behavioral Neuroscience, University of Cincinnati, Cincinnati, OH 45219, USA. ⁴These authors contributed equally: Fenghua Long, Yufei Chen, Qian Zhang. ✉email: charlie_lee@qq.com

Received: 18 January 2024 Revised: 18 August 2024 Accepted: 19 August 2024

Published online: 26 August 2024

learning [12–15]. Numerous studies have predicted treatment outcomes of antidepressants, ECT, and other treatments using pretreatment brain structural and functional MRI features. For example, resting state functional connectivity (rsFC) between dorsolateral prefrontal cortex (dlPFC) and visual regions evaluated with resting state functional MRI (rsfMRI) has shown a predictive utility [16], as well as activation of rostral anterior cingulate cortex (ACC) in task-based fMRI (tbfMRI) [17]. Structural MRI (sMRI) studies have shown a predictive utility for gray matter volume (GMV) of hippocampal subfields [18] and cortical thickness (CTH) of supplementary motor area [19]. These features have shown varying levels of sensitivity (ranging from 0.74 to 0.84) and specificity (ranging from 0.67 to 0.97) in predicting treatment outcomes. The differences in sensitivity and specificity among studies may arise from variations in interventions, sample cohort features, small study samples, varying MRI modalities, different tasks in tbfMRI, acquisition parameters, and analysis methods. Given these factors and the urgent need for predictive features for MDD treatment, a meta-analysis of the literature is needed to advance progress in this field.

Prior works have assessed utilizing neuroimaging features for treatment outcome prediction. A meta-analysis explored prediction based on electroencephalogram (EEG) and MRI, achieving an area under the curve (AUC) of 0.85 [8]. This meta-analysis didn't specify neuroimaging techniques, limiting insights into the specific role of MRI features for MDD treatment outcomes. Moreover, a meta-analysis of brain MRI features used for outcome prediction in MDD reported an AUC of 0.84 [20]. However, the expanding literature, along with limitations and unresolved research questions of prior work, emphasizes the need for further investigation. For example, this meta-analysis omitted separate subgroup meta-analyses for fMRI and sMRI studies [20], preventing a detailed understanding of the performance of different MRI modalities. Moreover, for functional MRI studies, the potential differential predictive utility of rsfMRI and tbfMRI has not been systematically examined. Furthermore, no meta-analyses have compared the predictive potential of MRI and clinical features. Therefore, an updated meta-analysis and systematic review are warranted to comprehensively understand the performance of brain features in predicting treatment outcomes using more available published studies.

The primary objective of the present meta-analysis was to evaluate the overall performance of clinical and brain MRI features for predicting treatment outcomes for MDD. A secondary objective was to explore the utility of different MRI modalities for predicting treatment outcomes and determine variations in predictive performance for different interventions. Our primary hypotheses were that: (a) MRI data would have greater performance for predicting treatment outcomes than clinical features in patients with MDD, and (b) the predictive performance in MRI studies would differ across imaging modalities and interventions.

MATERIALS AND METHODS

Search strategy and selection criteria

Our study followed the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guideline [21] (Table S1) and was registered on PROSPERO (CRD42022376797). Three authors (FHL, YFC, and QZ) conducted a literature search in PubMed, Embase, Web of Science, and Science Direct databases from inception to July 22, 2023.

Inclusion criteria were used to select studies that: (a) included participants meeting diagnostic and statistical manual of mental disorders (DSM) or international classification of diseases (ICD) criteria for MDD, (b) evaluated prediction of treatment outcome using pretreatment clinical (including severity ratings, duration, and demographics) or brain MRI data, and (c) provided a specific evidence-based clinical intervention such as antidepressants, ECT, or CBT. We excluded: (a) theoretical papers, case reports, reviews, and meta-analyses, (b) animal studies, (c) studies of

samples with mean age younger than 18 or older than 65 years, and (d) confusion matrix data for a study sample, which depicts the differences between model predictions and actual outcomes using true/false positives/negatives, was not available even after contacting the authors. Details of the search strategy and selection criteria (Figure S1) and quality assessment and data extraction (Tables S2–S5) are provided in the Supplementary Methods.

Meta-analysis

Data analysis was conducted in R (version 4.2.1) using the *mada* [22] and *glmnet* [23] packages. First, we calculated I^2 as a measure of study heterogeneity and classified it as low, moderate, and high for I^2 values of < 50%, 50–75%, and > 75% [24]. If the data were low in heterogeneity, a fixed effects model would be employed to estimate the results, i.e., here as the logarithm of the diagnostic odds ratio [log(DOR)]. Otherwise, a random effects model was employed. Log(DOR) greater than zero indicates that the predictive model can discriminate between responders/remitters and non-responders/non-remitters, and higher values indicate better predictive performance [25]. Second, bivariate analyses for sensitivity and specificity were implemented using the approach of Reitsma [26]. Because the pooled studies use different diagnostic thresholds (such as varying treatment response definitions), this could lead to heterogeneity in the estimates of sensitivity and specificity. To examine the potential impact of variable outcome thresholding approaches, we computed a summary receiver operating characteristic (SROC) curve that displays predictive utility across the range of potential thresholds. When the correlation coefficient (r) between sensitivity and false positive rate exceeds 0.6, it signifies a considerable threshold effect [27]. For the pooled meta-analysis and all subgroup meta-analyses in the present study, no considerable threshold effect was observed (all $r < 0.6$; r range, $-0.08 \sim 0.52$). Of note, AUC of an SROC curve ranging from 0.75 to 0.92 is generally considered to indicate good predictive performance [28]. When the lower bound of the 95% confidence interval [CI] for AUC exceeds 0.5 (chance performance), it signifies that the current prediction is significantly better than chance [29].

We conducted separate pooled meta-analyses for prediction models based on clinical features, encompassing demographics and severity ratings (Hamilton Depression Rating Scale [HDRS] scores, Hamilton Anxiety Rating Scale [HARS] scores, illness duration, sleep disruption score, etc.), and models based on brain MRI features. Subsequently, MRI studies were categorized into subgroups based on modalities (rsfMRI, tbfMRI, sMRI, and diffusion tensor imaging [DTI]) and interventions (antidepressant, ECT, rTMS, and CBT). In studies utilizing multiple MRI modalities, those exclusively relying on features from a single modality to predict treatment outcomes were categorized within the corresponding modality subgroup. One study that combined features from multiple modalities to collectively predict treatment outcomes was classified under the multiple MRI modalities group [30]. The tbfMRI studies were categorized into two subsets based on the emotional and cognitive tasks employed. We subdivided the antidepressant studies into those using selective serotonin reuptake inhibitors (SSRIs) and those using other medications. Separate meta-analyses were performed in subgroups that included five or more original studies ($n \geq 5$) to provide a reasonable statistical power, noting the preliminary nature of such analyses [31]. If a study reports on both response and remission, or multiple studies use the same patient data, we calculate a weighted average of their 2×2 tables, with weights determined by sample size [20]. We used a systematic review to summarize the current state of the field with regard to less studied imaging modalities and intervention approaches.

Next, we assessed publication bias by Deeks' funnel plot asymmetry test, considering it present if there was a nonzero slope coefficient ($P < 0.05$). The diagnostic original studies usually have an imbalance in negative samples (i.e., patients with poor treatment outcomes) over positive ones. Therefore, this imbalance tends to inflate the calculated DOR above 1, potentially affecting the standard errors and increasing the likelihood of a false positive result if it is calculated by Begg's and Egger's methods. The Deeks' test conducted a linear regression of log(DOR) and precision (the reciprocal of the square root of the effective sample size) to investigate the relationship between effect size and effective sample size, which further evaluated the impact of publication bias. The effective sample size is more appropriate than the sample size to summarize the precision due to the unequal numbers of responder/remitter and non-responder/non-remitter in each study [32].

Then, bivariate meta-regression was used firstly to explore the effects of covariables (sample size, age, sex ratio, illness duration, HDRS score, and

publication year) on sensitivity and specificity of each meta-analysis [33]. Baseline BDI and MADRS scores were transformed to approximate HDRS categories using published transformation rules [34, 35]. The bivariate meta-regression was also used to test for the differential effects between clinical and MRI studies, as well as among pairs of modality and intervention subgroups within MRI studies, respectively. The likelihood ratio test was conducted to determine the statistical significance of the differences in estimated variances of logit sensitivity and logit specificity before and after adding covariates (i.e., imaging modalities or interventions) in bivariate meta-regression [36]. For example, to assess the influence of interventions on predictive performance, the antidepressant subgroup was taken as a reference group, and an ECT subgroup was added as a covariate to analyze the statistical significance of changes in the estimated variance. Therefore, we conducted bivariate meta-regression on the following studies, including: (a) clinical and MRI studies, (b) three different MRI modalities (rsfMRI, tbfMRI, and sMRI) to evaluate potential imaging modality variations and provide additional information for outcome prediction, (c) tbfMRI studies employing emotional and cognitive tasks to understand the impact of tasks on prediction, (d) antidepressant and ECT subgroups to assess their distinct predictive performance, and (e) specific SSRIs and ECT to contrast predictive features performance. For multiple comparisons in the analyses among three different MRI modalities, we used the Bonferroni correction, and results were considered significant if the P value was less than $0.05/3 = 0.017$.

Finally, the elastic net algorithm was employed to construct a multivariate regression model for predicting log(DOR) in each study, investigating the impact of methodological and clinical variables on the prediction. Methodological variables mainly encompassed data types (e.g., rsfMRI data), validation status, magnetic field strength, and predicted methodologies (e.g., support vector machine [SVM]). Clinical variables included sample size, age, intervention type, etc. To address missing values in variables, we utilized the k -nearest neighbors algorithm for imputation. We utilized nested cross-validation (CV) with 10-fold for the inner CV and leave-one-out for the outer CV, aiming to select alpha and lambda values that minimized the root mean squared error. Correlation analysis was conducted between the predicted log(DOR) and true log(DOR) to assess the reliability of the predictive model.

RESULTS

Characteristics of included studies

We included 13 studies that used clinical features to predict treatment outcomes, covering 4301 patients (mean age, 45.1 years; male/female, 1753/2548); and 44 MRI studies recruited 2623 patients (mean age, 38.2 years; male/female, 1109/1514) (Table 1 and S6). Within MRI studies, 19 rsfMRI, 13 tbfMRI, and ten sMRI studies were included in modality subgroups (Table S7); 27 MRI studies utilized antidepressants and nine utilized ECT which were included in intervention subgroup analyses (Table S8). Detailed characteristics of included studies are provided in the Supplementary Results.

Pooled meta-analysis

Due to the high heterogeneity observed with the fixed effects model, a random effects model was employed for the present study. The overall log(DOR) of clinical studies for treatment outcome prediction was 1.62 (95% CI 1.16–2.09; Fig. 1). The AUC of SROC curve was 0.73 (95% CI 0.67–0.81; Fig. 2), sensitivity was 0.62 (95% CI 0.48–0.74), and specificity was 0.76 (95% CI 0.64–0.85). No covariates were identified to impact the sensitivity and specificity ($P > 0.05$). There was a low heterogeneity observed among studies ($I^2 = 42.4\%$). Deeks' funnel plot asymmetry test did not reveal significant publication bias in the included studies ($\beta = 0.008$, $P = 0.51$; Fig. S2). No significant correlation was observed between the predicted log(DOR) and true log(DOR) in clinical studies ($r = 0.12$, $P = 0.71$; Fig. S3).

The pooled meta-analysis of all included MRI studies revealed an overall log(DOR) of 2.53 (95% CI 2.22–2.84; Fig. 1). The AUC of the SROC curve was 0.89 (95% CI 0.87–0.91; Fig. 2), indicating performance better than chance. Sensitivity was 0.78 (95% CI 0.75–0.81), and specificity was 0.75 (95% CI 0.71–0.79). No

covariates had a significant impact on overall sensitivity and specificity ($P > 0.05$). There was no evidence of heterogeneity observed among studies (Fig. 1). Deeks' funnel plot asymmetry test did not demonstrate significant publication bias in the included studies ($\beta = 0.001$, $P = 0.93$; Fig. S2). In the meta-regression comparing clinical and MRI studies, we identified significant differences in predicting treatment outcomes ($\text{Chi}^2 = 6.53$, $P = 0.03$), with the MRI studies exhibiting higher sensitivity ($Z = 3.42$, $P = 0.001$).

For the pooled MRI studies, we employed the elastic net algorithm with an average of $\alpha = 0.5$ and $\lambda = 0.21$ across all CV-folds. The predicted log(DOR) showed a significant correlation with the true log(DOR) ($r = 0.39$, $P = 0.02$). Six variables were identified based on the absolute value of their coefficients. Specifically, "data: tbfMRI", "scanner: 1.5 T", and "sample size" were linked to lower log(DOR), while "data: rsfMRI", "method: ROC curve analysis", and "illness duration" demonstrated associations with higher log(DOR) (Fig. 3).

Modality subgroup outcomes in MRI studies

Meta-analysis. The rsfMRI subgroup consisted of a total of 1130 patients (mean age, 40.7 years; male/female, 440/690). The outcome prediction model had a log(DOR) of 2.74 (95% CI 2.39–3.08), sensitivity of 0.80 (95% CI 0.75–0.84), specificity of 0.79 (95% CI 0.75–0.82), and an AUC of 0.90 (95% CI 0.87–0.93).

For tbfMRI studies, which included 891 participants (mean age, 34.2 years; male/female, 394/497), the log(DOR) was 2.00 (95% CI 1.28 to 2.72), sensitivity was 0.74 (95% CI 0.67–0.81), specificity was 0.69 (95% CI 0.56 to 0.79), and AUC was 0.85 (95% CI 0.78–0.92). Regarding different tasks employed in tbfMRI, the log(DOR) of emotional task subgroup was 1.78 (95% CI 0.91 to 2.65), sensitivity, specificity, and AUC were 0.77 (95% CI 0.68–0.84), 0.63 (95% CI 0.48–0.75), and 0.84 (95% CI 0.73–0.99). The cognitive task subgroup has a log(DOR) of 2.35 (95% CI 1.69–3.02), sensitivity of 0.77 (95% CI 0.69–0.84), specificity of 0.73 (95% CI 0.63–0.81), and AUC of 0.88 (95% CI 0.80–0.97).

The sMRI studies included 347 patients (mean age, 39.5 years; male/female, 146/201). The log(DOR) was 2.63 (95% CI 1.96–3.30), and sensitivity, specificity, and AUC were 0.79 (95% CI 0.71 to 0.86), 0.73 (95% CI 0.63–0.81), and 0.91 (95% CI 0.86–0.96), respectively (Fig. S4a and Table S9).

The heterogeneity test showed low heterogeneity in rsfMRI subgroup ($I^2 = 1.57\%$) and no evidence of heterogeneity was observed in the tbfMRI and sMRI subgroups (Fig. S6). No publication bias was found in any of these three subgroup analyses (Fig. S5a). Using meta-regression analysis, we found a significant difference in the sensitivity and specificity of outcomes predicted by the rsfMRI and tbfMRI subgroups ($\text{Chi}^2 = 8.70$, uncorrected $P = 0.013$ survived with the Bonferroni correction). Specifically, while the sensitivity was similar ($Z = -1.13$, $P = 0.26$), rsfMRI showed higher specificity than tbfMRI ($Z = -2.86$, $P = 0.004$). There was no significant difference in sensitivity and specificity in prediction of treatment outcome between sMRI and rsfMRI subgroups ($\text{Chi}^2 = 1.00$, $P = 0.61$), between sMRI and tbfMRI subgroups ($\text{Chi}^2 = 1.70$, $P = 0.43$), or between emotional and cognitive tbfMRI ($\text{Chi}^2 = 1.61$, $P = 0.45$). Furthermore, we found that HDRS score was a significant covariate influencing the prediction in emotional task subgroup ($\text{Chi}^2 = 6.34$, $P = 0.04$), with a negative impact on its specificity ($Z = -2.88$, $P = 0.004$).

Predictive brain features in models including all studies using each type of MRI protocol. Analysis of features selected for outcome predictions indicated that predictive brain regions were predominantly located within the limbic and default mode networks (DMN) for both rsfMRI and tbfMRI studies. The rsfMRI features included rsFC between ACC and middle frontal gyrus [37], amygdala [38], and dlPFC [39], as well as between medial PFC and posterior cingulate cortex (PCC) [40]. Predictive features for

Table 1. Characteristics of all included studies in the present meta-analysis.

Study and year	N	Mean age (SD)	M/F	Symptom severity	Features	Intervention	Outcome	Analysis method
Clinical studies (<i>n</i> = 13)								
Y.L. Zhou et al. [7]	187	33.6 (12.1)	92/95	HDRS-17, 22.9	clinical, demographic	antidepressant	response, remission	logistic regression
A. Kautzky et al. [11]	518 ^a	51.8 (14.3)	205/347	MADRS, 35.1	clinical, demographic	antidepressant	response	random forest
A. Kautzky et al. [88]		49.6 (14.3)	151/329	HDRS, -	clinical, demographic	antidepressant	remission	random forest
S. Wagner et al. [54]	617	40.9 (11.6)	264/353	HDRS-17, -	clinical, demographic	antidepressant	remission	logistic regression
R. Perlis et al. [89]	461	40.0 (13.4)	188/273	QIDS-SR, 14.9	clinical, demographic	antidepressant	remission	logistic regression
M. Riedel et al. [5]	1014	45.1 (11.9)	412/602	HDRS-21, 24.8	clinical, demographic	antidepressant	response	logistic regression
Liang S. et al. [90]	237	47.8 (16.6)	88/149	HDRS-17, 26.4	clinical, demographic	ECT	remission	logistic regression
L. Carstens et al. [9]	96	52.6 (14.8)	38/58	MADRS, 30.2	clinical, demographic	ECT	response	logistic regression
M. Rezaei et al. [91]	116	40.5 (11.6)	40/76	BDI-II, 34.0	clinical	tDCS	response	logistic regression
N. Krepel et al. [10]	196 ^a	43.6 (-)	97/99	BDI-II, 31.2	clinical	rTMS	response	linear discriminant analysis
L. Donse et al. [92]		43.2 (12.9)	98/98	BDI-II, 31.3	clinical, demographic	rTMS	response, remission	ROC curve analysis
E.L. Brakemeier et al. [83]	75 ^a	49.1 (14.3)	36/43	HDRS, 21.0	clinical	rTMS	response	logistic regression
E.L. Brakemeier et al. [82]		49.5 (14.2)	44/26	HDRS, 26.1	clinical	rTMS	response	logistic regression
MRI studies (<i>n</i> = 44)								
Xue S.W. et al. [93]	36	28.0 (10.1)	10/26	HDRS-24, 27.8	rsfMRI	antidepressant	response	ROC curve analysis
T. Nakamura et al. [38]	15	45.9 (12.5)	6/9	HDRS-17, 27.6	rsfMRI	antidepressant	response	ROC curve analysis
Xiao H.Q. et al. [63]	36	35.1 (10.5)	17/19	HDRS-17, 22.1	rsfMRI	antidepressant	response	ROC curve analysis
Zhang F. et al. [50]	40 ^a	36.5 (12.2)	15/24	MADRS, 31.5	rsfMRI	antidepressant	remission	ROC curve analysis
Zhang F. et al. [51]		36.2 (11.4)	16/24	MADRS, 32.1	rsfMRI	antidepressant	response	ROC curve analysis
Ye Y.X. et al. [62]	66	27.7 (9.3)	25/41	HDRS-17, 23.1	rsfMRI	antidepressant	remission	ROC curve analysis
Wu H. et al. [94]	67	27.0 (7.2)	20/47	HDRS-17, 21.4	rsfMRI	antidepressant	remission	SVM
Pei C. et al. [95]	98	34.1 (10.1)	44/54	HDRS-6, 24.7	rsfMRI	antidepressant	response	SVM
Tian S. et al. [96]	106	32.7 (8.8)	55/51	HDRS-17, 24.0	rsfMRI	antidepressant	response	SVM
Hou Z.H. et al. [65]	81	48.0 (15.2)	22/59	HDRS-24, 29.4	rsfMRI	antidepressant	response	logistic regression analysis
Zhu J.J. et al., [97]	47	46.4 (13.5)	20/27	HDRS-17, 30.3	rsfMRI	antidepressant	response	logistic regression analysis
H.J. Hopman et al. [37]	61	42.5 (11.5)	14/47	MADRS, 29.5	rsfMRI	rTMS	response	SVM
R.F.H. Cash et al. [98]	33	43.0 (12.0)	14/19	MADRS, 33.9	rsfMRI	rTMS	response	SVM
A.T. Drysdale. [87]	124	40.6 (-)	53/71	HDRS-17, 20.4	rsfMRI	rTMS	response	SVM
Ge R.Y. et al. [99]	50	43.7 (12.1)	21/29	HDRS-17, 21.9	rsfMRI	rTMS	response, remission	ROC curve analysis

Table 1. continued

Study and year	N	Mean age (SD)	M/F	Symptom severity	Features	Intervention	Outcome	Analysis method
S.N. Gosnell et al. [30]	175	31.1 (12.1)	89/86	PHQ-9, -	rsfMRI, DTI	unspecified	response	logistic regression analysis
Sun H.L. et al. [100]	122	56.3 (15.9)	43/79	HDRS-17, 25.6	rsfMRI	ECT	response, remission	logistic regression analysis
M. Moreno-Ortega et al. [16]	18	52.0 (12.0)	7/11	HDRS-24, 26.5	rsfMRI	ECT	remission	logistic regression analysis
J.A. van Waarde et al. [39]	45	56.6 (14.1)	17/28	MADRS, 36.5	rsfMRI, sMRI	ECT	remission	SVM
A.M. Leaver et al. [86]	41 ^a	41.7 (6.8)	21/25	HDRS-17, 24.0; MADRS, 37.5; QIDS-SR, 19.7	rsfMRI, sMRI, asfMRI	ECT	response	SVM
B.S.C. Wade et al. [80]	18/16	41.7 (13.5)	18/16	HDRS-17, 24.3; QIDS-SR, 20.6; MADRS, 38.4	sMRI	ECT	response	SVM
B.S.C. Wade et al. [49]	89	41.0 (14.0)	21/21	HDRS-17, -	sMRI	ECT	remission	random forest
P.M. Fisher. et al. [81]	27/62	27.0 (8.3)	27/62	HDRS-17, 22.7	tbfMRI	antidepressant	response, remission	random forest
B.M. Meyer et al. [101]	22	31.5 (7.7)	11/11	MADRS, 27.0	tbfMRI	antidepressant	remission	logistic regression analysis
B.R. Godlewska et al. [17]	32	28.4 (6.1)	14/18	HDRS, 23.3; BDI-I, 32.0	tbfMRI	antidepressant	response	logistic regression analysis
N.A. Crane et al. [44]	29	34.5 (11.2)	12/17	HDRS, 19.5	tbfMRI	antidepressant	remission	logistic regression analysis, random forest
A.F. Marquand et al. [45]	20	43.7 (8.6)	6/14	HDRS-17, 21.2	tbfMRI	antidepressant	response	SVM
ISPOT-D	145 ^a							
T.A. Braund et al. [41]	106/120	33.4 (11.7)	106/120	HDRS-17, 21.7; QIDS-SR16, 14.2	tbfMRI	antidepressant	response	SVM
M.S. Korgaonkar et al. [42]	80/83	34.1 (11.4)	80/83	HDRS-17, 23.6; HDRS-6, 12.0	tbfMRI	antidepressant	remission	logistic regression analysis
A.N. Goldstein-Piekarski et al. [40]	36/39	31.7 (11.4)	36/39	HDRS-17, 21.1; QIDS-SR16, 13.9	tbfMRI	antidepressant	remission	logistic regression analysis
A.N. Goldstein-Piekarski et al. [102]	35/35	32.3 (12.4)	35/35	HDRS-17, 21.1; QIDS-SR16, 13.8	tbfMRI	antidepressant	remission	logistic regression analysis
L.M. Williams et al. [103]	40/40	32.8 (12.4)	40/40	HDRS-17, 21.1; QIDS-SR16, 13.8	tbfMRI	antidepressant	response	linear discriminant analysis
M.S. Korgaonkar et al. [104]	40/40	33.8 (13.1)	40/40	HDRS-17, 21.0	DTI	antidepressant	remission	logistic regression analysis
F. Queiraza et al. [105]	37	39.2 (12.7)	19/18	BDI-II, 27.8	tbfMRI	CBT	response	SVM, logistic regression analysis
S.G. Costafreda et al. [46]	16	40.0 (9.4)	3/13	HDRS-17, 20.6	tbfMRI	CBT	remission	SVM
G.J. Siegle et al. [43]	32	36.1 (11.1)	5/27	BDI-I, 30.7; HDRS, 20.6	tbfMRI	CBT	response, remission	random forest
Wu P.Y. et al. [19]	63	33.1 (9.9)	25/38	HDRS-17, 22.0	sMRI	antidepressant	response	ROC curve analysis
Hu X.X. et al. [18]	38	36.2 (12.1)	22/16	HDRS-17, 22.9	sMRI	antidepressant	response	ROC curve analysis
Gong Q.Y. et al. [106]	46	39.8 (12.7)	24/22	HDRS-17, 23.9	sMRI	antidepressant	response	SVM
I. Nouredinov et al. [107]	18 ^a	44.2 (8.2)	4/14	HDRS-17, 21.1	sMRI	antidepressant	remission	SVM
S.G. Costafreda et al. [48]	41	44.2 (8.2)	4/14	HDRS-17, 21.1	sMRI	antidepressant	remission	SVM
Xu J.P. et al. [108]	41	37.7 (10.8)	9/32	HDRS-17, 22.9	sMRI	ECT	remission	SVM

Table 1. continued

Study and year	N	Mean age (SD)	M/F	Symptom severity	Features	Intervention	Outcome	Analysis method
Cao B. et al. [47]	24	31.3 (10.8)	10/14	HDRS, 31.3	sMRI	ECT	remission	SVM
R. Redlich et al. [52]	23	45.7 (9.8)	9/14	HDRS-17, 26.0; BDI, 30.9	sMRI	ECT	response	SVM

as/rs/tbfMRI Arterial spin labeling/resting-state/task-based functional magnetic resonance imaging, BDI Beck depression inventory, CBT Cognitive behavioral therapy, DTI Diffusion tensor imaging, ECT Electroconvulsive therapy, HDRS Hamilton depression rating scale, ISPT-D International study to predict optimized treatment in depression, MADRS Montgomery-Asberg depression rating scale, M/F Male/Female, N Patient number, PHQ Patient health questionnaire, QIDS-SR Quick inventory of depressive symptomatology-self-report, ROC Receiver operating characteristic, rTMS Repetitive transcranial magnetic stimulation, SD Standard deviation, sMRI Structural magnetic resonance imaging, SVM Support vector machine, tDCS Transcranial direct current stimulation.

^aRepresents a weighted average number of study samples utilizing the same patient dataset; -represents that the relevant information or data is not reported.

tbfMRI included task-based FC between limbic and somatomotor networks [41], as well as within DMN [42]. Activation of ACC [43, 44] and precuneus [45, 46] in tbfMRI studies also contributed to prediction. The sMRI predictive features for all treatments predominantly included brain regions within limbic network not the DMN, including GMV of hippocampus [18, 47], GM density of ACC [48], and CTh of hippocampus [49] (Table S10, Fig. 4).

Intervention subgroup outcomes in MRI studies

Meta-analysis. Studies including 1700 patients using antidepressants (mean age, 35.5 years; male/female, 736/964) showed a log(DOR) of 2.48 (95% CI 2.05–2.91), sensitivity of 0.78 (95% CI 0.74–0.82), specificity of 0.74 (95% CI 0.68–0.80), and an AUC of 0.89 (95% CI 0.86–0.92). The log(DOR), sensitivity, specificity, and AUC for studies with patients only administered SSRIs were 2.68 (95% CI 1.81–3.56), 0.79 (95% CI 0.72–0.84), 0.75 (95% CI 0.61 to 0.85), and 0.91 (95% CI 0.86–0.96), respectively. ECT studies included 395 participants (mean age, 44.9 years; male/female, 155/240). The log(DOR) for ECT studies was 2.56 (95% CI 1.90–3.22), sensitivity and specificity were 0.83 (95% CI 0.69 to 0.91) and 0.74 (95% CI 0.65 to 0.82), and AUC was 0.89 (95% CI 0.80 to 1.00; Fig. S4b and Table S9).

No publication bias or evidence of heterogeneity was found in any intervention subgroup (Figs. S5b and S7). Meta-regression showed no significant differences between antidepressant (SSRI and other antidepressant studies combined) and ECT subgroups (combined across imaging modalities) in sensitivity and specificity ($\text{Chi}^2 = 0.98$, $P = 0.61$), as well as in SSRIs and ECT ($\text{Chi}^2 = 0.10$, $P = 0.95$). We observed that sample size significantly affected the predictive efficacy of ECT treatment outcomes ($\text{Chi}^2 = 7.98$, $P = 0.02$), negatively influencing its sensitivity ($Z = -3.50$, $P < 0.001$).

Predictive brain features. Through a systematic review, we found that features for antidepressants, including SSRIs examined separately, were distributed in the limbic network and DMN. Predictive features included rsFC between hippocampus and angular gyrus [50], and between ACC and supplementary motor area [51]. The task-based FC between DMN and somatomotor networks [42], and activation of medial PFC [40], were also significant predictors. In terms of ECT, features related to treatment outcome were mainly found in the limbic network, including rsFC between ACC and dlPFC [39], GMV of subgenual ACC [52], as well as CTh of hippocampus [49] (Table S11, Fig. 4).

DISCUSSION

Findings of the present meta-analysis highlight the potential of utilizing pretreatment brain MRI data to predict treatment outcomes for MDD patients, outperforming clinical features. Pretreatment alterations in functional and structural brain features may explain in part the wide heterogeneity of clinical response to antidepressant therapies. In imaging modality subgroups, rsfMRI outperformed tbfMRI in specificity, more accurately identifying true negatives (i.e., non-responders and non-remitters) among patients. Outcome prediction of sMRI features did not differ significantly from either of the two fMRI modalities. No significant differences were found among the different intervention subgroups in accuracy of outcome prediction. Although outcome prediction features mainly involved DMN and limbic networks, predictive neuroimaging features differed somewhat among modality and intervention subgroups.

Overall prediction performance

Our meta-analysis utilized brain MRI data to predict treatment outcomes and revealed a superior predictive performance compared to that reported in a meta-analysis of EEG data, with an AUC of 0.89 versus 0.76 [53]. Our overall predictive performance is similar to the findings reported by Lee et al. for

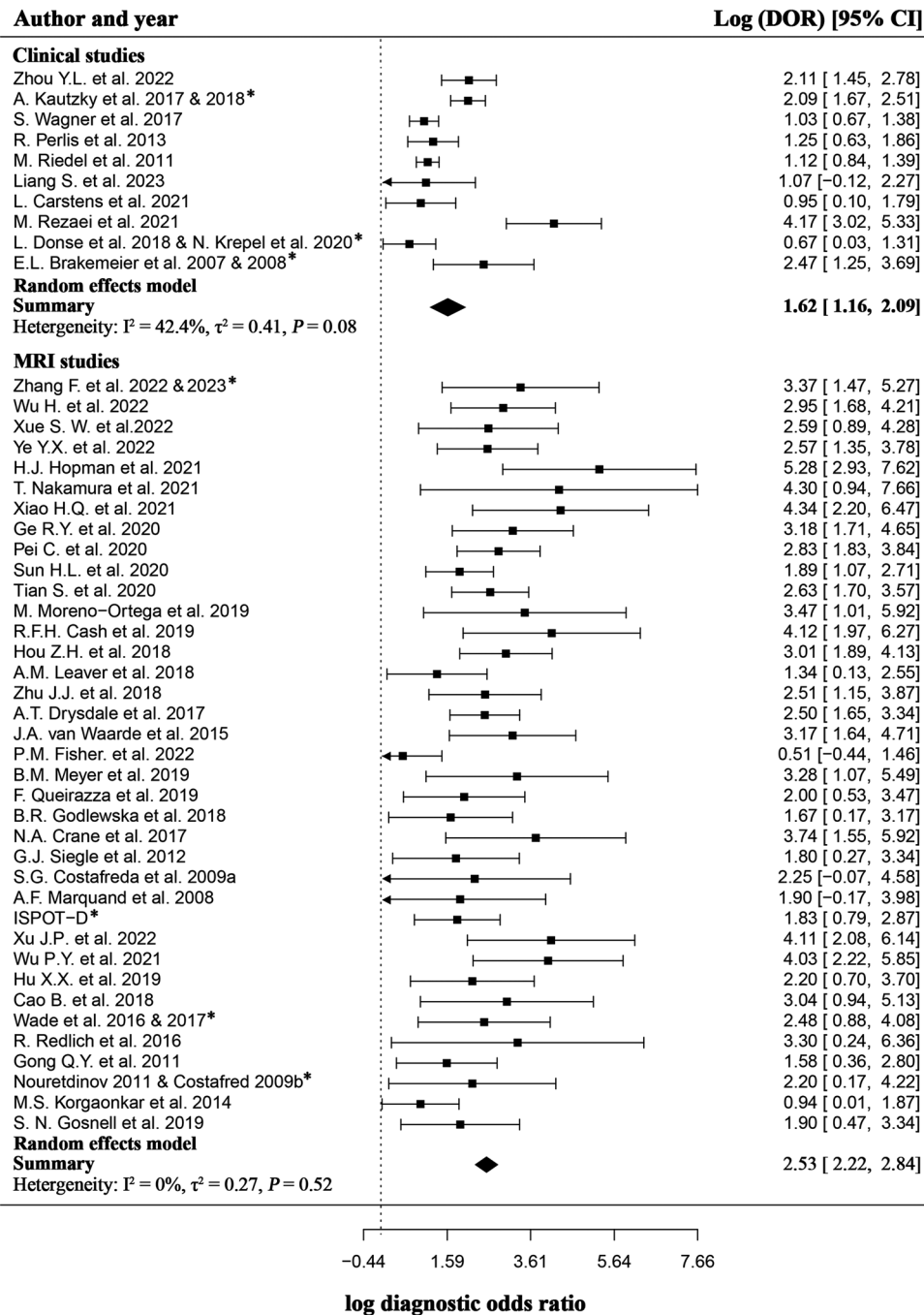


Fig. 1 Overall random effects model forest plot of the logarithm of diagnostic odds ratios in clinical and MRI studies. CI confidence interval; log(DOR), the logarithm of diagnostic odds ratios. Notes: * represents data after weighted averaging of studies that used repeated samples. ISPOD-D included six studies utilizing the international study to predict optimized treatment in depression data.

EEG and MRI-based predictive markers (AUC: 0.89 vs. 0.85) [8]. Furthermore, we observed comparable overall performance to another meta-analysis (AUC: 0.89 vs. 0.84) [20], further emphasizing the robust potential of MRI data in predicting treatment outcomes in MDD patients. In addition, with the leverage of more recent studies to increase samples, we were able to compare different imaging modalities and treatment types. In terms of comparisons of MRI modalities, we observed rsfMRI outperformed tbfMRI in predicting treatment outcomes. We additionally utilized a multivariate regression model to identify predictive factors associated with predictive accuracy. The results further support our findings, indicating an association between rsfMRI utility and

higher log(DOR), while tbfMRI demonstrates a correlation with lower log(DOR).

Pretreatment clinical features are readily accessible features for patients, but the performance of models based on them in predicting MDD treatment outcomes is limited. This may be attributed to the diversity of clinical features introducing heterogeneity. Although all clinical studies included scores of symptom severity scales, the feature sets for prediction were not uniform (i.e., not all studies including comorbidities or illness duration). A prior study revealed that combining pretreatment clinical features with early response factors (i.e., severity ratings at two weeks after medication administration) significantly improves

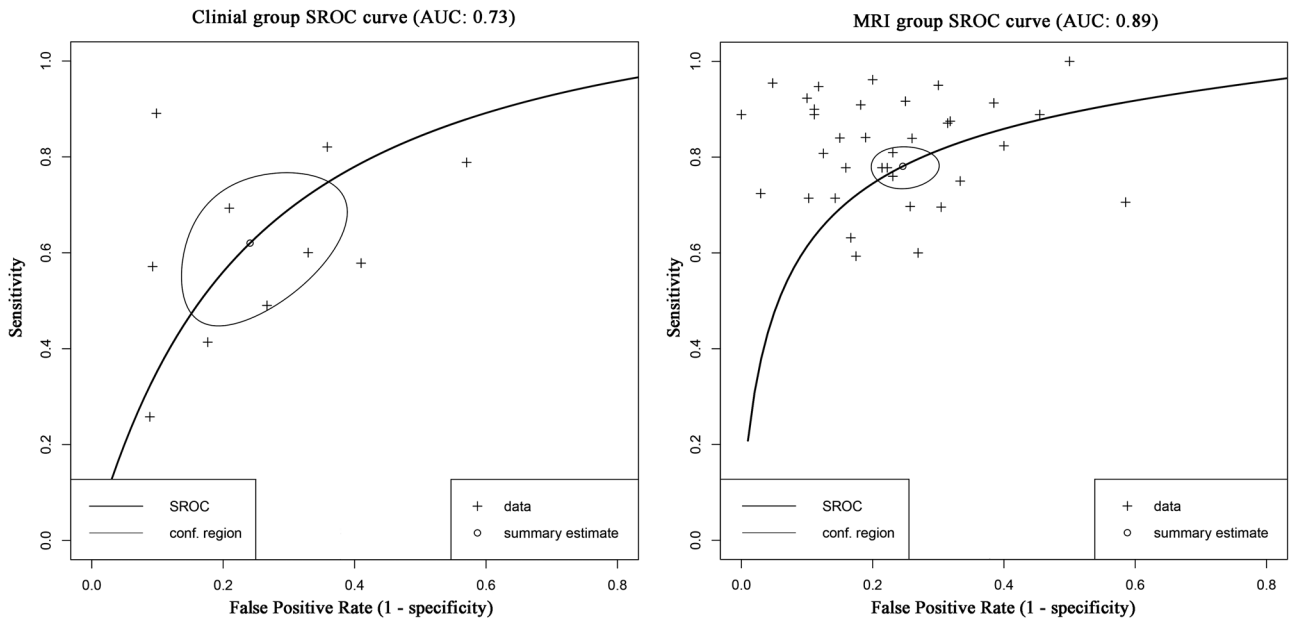


Fig. 2 Summary receiver operator characteristic (SROC) curve within clinical and MRI studies. AUC Area under the curve, CI Confidence interval; Conf. region, the region of confidence interval, MRI Magnetic resonance imaging.

treatment outcome prediction, raising specificity from 30% to 90% [54]. Furthermore, integrating clinical features with patient genetics, metabolomics, and other features is poised to further enhance prediction accuracy [55, 56]. Collectively, our study highlights the superior suitability of pretreatment brain MRI relative to clinical features for predicting treatment outcomes. It is worth noting however that it was not possible to directly compare clinical and MRI features for prediction in the same sample, and warrants additional further validation in future studies.

As key brain networks that have been implicated in MDD, the limbic network and DMN play crucial roles in various cognitive processes, reward regulation, and emotion homeostasis [57, 58]. Previous studies have suggested that MDD patients exhibit decreased activation in the limbic network, which has been hypothesized to be related to abnormal reward-related behaviors and reduced hedonic tone [59]. Furthermore, reduced rsFC within DMN can disrupt the ability to disengage from internal emotions and cognitive processes, impairing their ability to focus on external tasks and experiences [57]. Our meta-analysis demonstrated the predictive power of alterations in these networks for treatment outcomes, showing good predictive performance with AUC ranging from 0.84 to 0.91 that varied within that range depending on the MRI modality and intervention subgroups. Previous studies have reported that treatment-resistant patients with MDD demonstrated decreased rsFC and regional homogeneity within DMN [42, 60–62], and reduced rsFC between DMN and other brain regions [51, 63], as compared to successfully treated individuals. In parallel, increased fractional amplitude of low-frequency fluctuations and rsFC within the limbic network [38, 64], along with elevated rsFC between limbic network and other brain regions [65], were also found to be greater in treatment-resistant patients. These findings are reflected in aggregate in our meta-analysis findings.

Some factors influenced the prediction of treatment outcomes using brain MRI data. Previous studies demonstrated that longer MDD duration is associated with reduced GMV in hippocampus and ACC [66], along with weakened FC between ACC and DMN [67]. This is consistent with crucial regions in DMN and the limbic network that we systematically identified for prediction, and MRI features yield a higher predictive accuracy for patients with a longer illness duration. Moreover, the negative correlation

between sample size and predictive accuracy may be influenced by the clinical and neurophysiological heterogeneity in MDD patients [68]. Regarding methodology, the 1.5 T MRI data has a low signal-to-noise ratio and spatial resolution [69], which might contribute to the observed low predictive accuracy in our results. Lastly, prediction using feature-based ROC curve analysis, which typically selects MRI features with optimal performance, exhibited high predictive accuracy. However, as the performance of this method does not stem from its learning capabilities, our findings do not imply its superiority over other machine learning methods.

Varied MRI modalities

The superior predictive utility of rsfMRI relative to tbfMRI is based on the studies that have been done previously. While methods across studies can vary in terms of how measurements are performed, for tbfMRI there is an additional major consideration of what task is being performed during scans. While many studies using a range of tasks were examined, the potential utility of different tasks is unknown. Other methodological features for all fMRI include factors such as the focus on regional activity vs. FC between pairs of regions vs. whole brain connectome analysis using graph theory approaches [70]. Sufficient evidence does not yet exist to clarify which of these approaches has the best utility considered individually or in combination. Further, a previous study suggested that increased severity of MDD correlates with reduced hippocampus and amygdala activation during emotional processing [71]. Patients with severe symptoms and varied treatment outcomes may exhibit comparable lower baseline activation in these regions, which was consistent with our findings and might be one possible explanation for the negative impact of heightened symptom severity on specificity by using emotional tbfMRI for treatment outcome prediction.

There are similar methodological issues for anatomic imaging, as cortical thickness, surface area, gyral features, and volume measurements can all provide somewhat independent information. In our meta-analysis, we found that sMRI predictors of treatment outcome were more commonly observed within the limbic network rather than DMN, though from a broader perspective, we did not observe significant differences in predictive performance between sMRI and the two fMRI modalities. The differences in findings across MRI modalities are

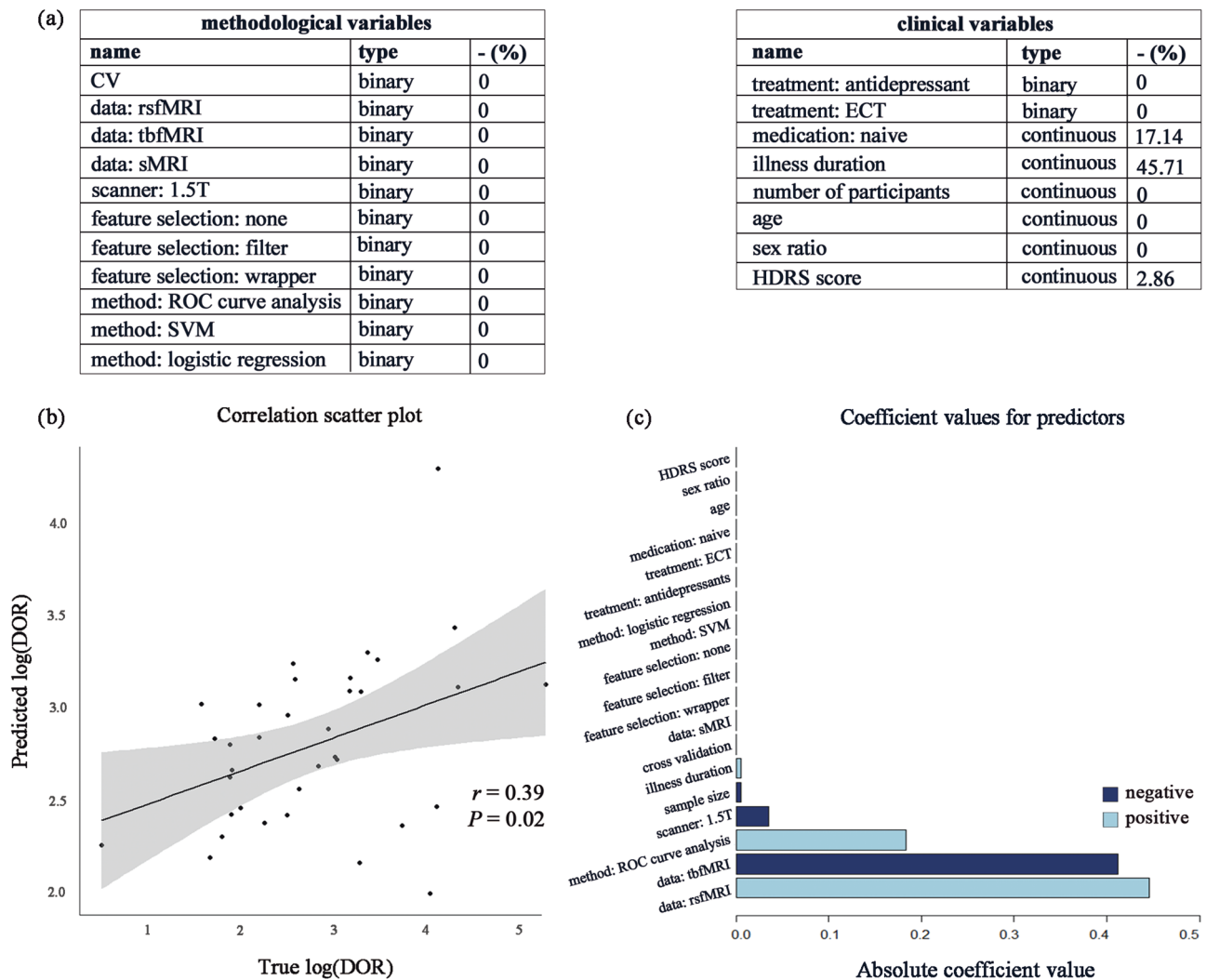


Fig. 3 Results of multiple regression model in MRI group by the elastic net algorithm predicting log(DOR) of individual studies. **a** Nineteen methodological and clinical variables were included in elastic net algorithm. **b** The predicted log(DOR) was significantly correlated with true log(DOR) ($r = 0.39$, $P = 0.02$). **c** Six variables with non-zero coefficients were important predictors for log(DOR) prediction, ranked by their absolute value of coefficient values from the lowest to the highest. CV Cross validation, ECT Electroconvulsive therapy, HDRS Hamilton depression rating scale, log(DOR), the logarithm of diagnostic odds ratios; MRI Magnetic resonance imaging, ROC receiver operating characteristic, rs/tbfMRI resting-state/task-based functional magnetic resonance imaging; sMRI, structural magnetic resonance imaging; SVM support vector machine. Note: -, represents that the relevant information or data is not reported.

of mechanistic interest for MDD, as local anatomic alteration can induce functional alterations elsewhere in regions with which they are connected. They also are of interest in supporting the need for more multimodal imaging studies [72–76] that use diverse MRI data in an integrated way to perhaps improve prediction as well as understanding of treatment response in MDD.

Different interventions

Antidepressant medication is commonly used to treat depression. They aid in restoring the functionality of brain networks by modulating interaction of neurotransmitter systems [77]. ECT is a stimulation therapy that produces therapeutic effects by directly stimulating neural activity. Animal models have shown that repeated electrical stimulation induces neurogenesis, synaptogenesis, and synaptic plasticity in the brain [78]. In the intervention subgroup meta-regression, despite the different mechanisms of antidepressants and ECT, similar sensitivity, specificity, and AUC were observed in treatment outcome prediction using overlapping brain features. Our study demonstrated that the primary distribution of brain-predicting outcomes of antidepressant

therapy was observed in the limbic network and DMN, while those for ECT were more specific to the limbic network. Although FC of the PCC and precuneus contributed to outcomes prediction for both antidepressants and ECT, the structural characteristics and functional activation within DMN were exclusive to predicting antidepressant outcomes. Interpretation of this similarity is also complicated by the fact that many receiving ECT are also administered antidepressants before and during ECT treatment.

While ECT is widely recognized for its high efficacy in treating therapy-resistant patients, it is not a first line treatment due to costs, adverse effects on memory, and stigma [79]. In this context, identifying brain features that predict a poor response to antidepressants and a more positive outcome to ECT would be advantageous clinically by suggesting a need for ECT rather than a trial with a different antidepressant medication for treatment nonresponsive patients.

Limitations

Certain limitations should be noted in interpreting the present findings. First, confusion matrices essential for our primary analysis

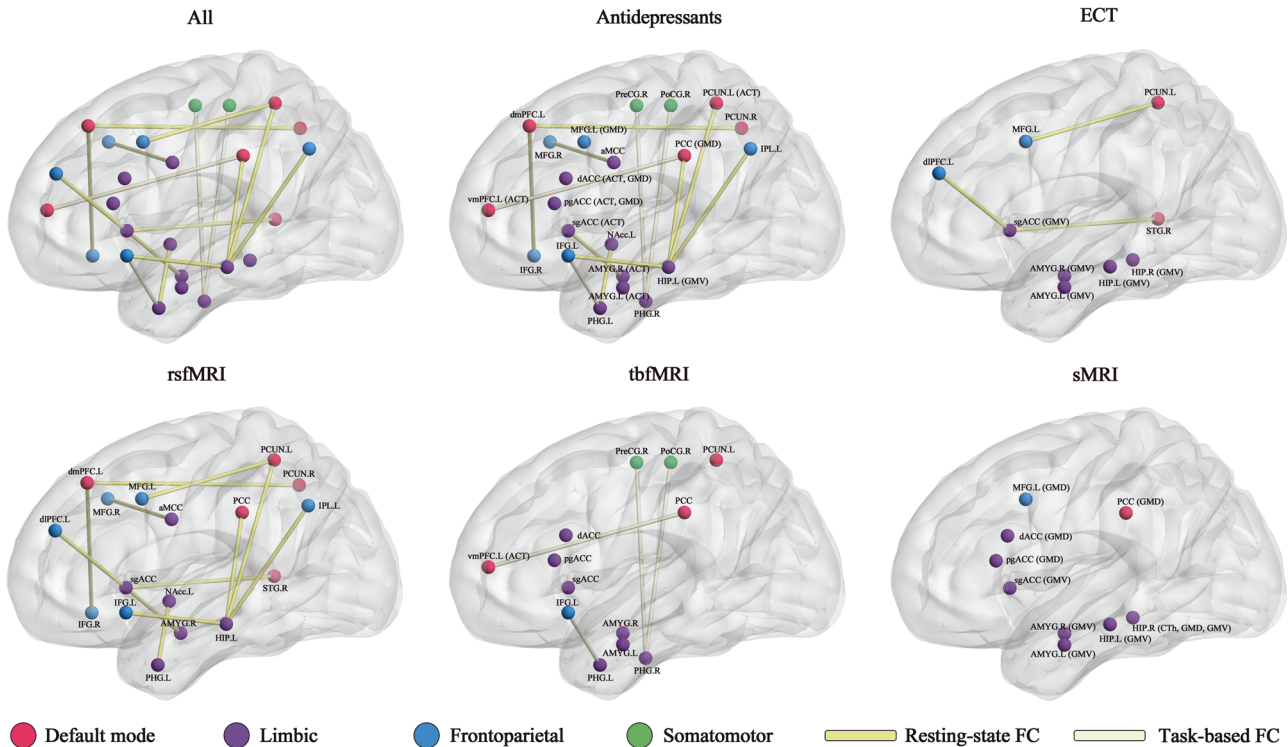


Fig. 4 Summary schematic representation of brain MRI features predicting treatment outcomes of major depressive disorder. Illustrated are all predictive brain MRI features and features for different MRI modalities and interventions. Node color illustrates the brain networks the nodes belong to. Yellow/white edges represent resting-state/task-based FC. ACT activation, aMCC anterior midcingulate cortex, AMYG amygdala, CTh cortex thickness, d/pgACC dorsal/pregenual/subgenual anterior cingulate cortex, dl/dm/vmPFC dorsolateral/dorsomedial/ventromedial prefrontal cortex, ECT electroconvulsive therapy, FC functional connectivity, GMD/V gray matter density/volume, HIP hippocampus, I/MFG inferior/middle frontal gyrus, IPL inferior parietal lobule, L left, NAcc nucleus accumbens, PCC posterior cingulate cortex, PCUN precuneus, PHG parahippocampal gyrus, Po/PreCG post/precentral gyrus, R right, rs/tbfMRI resting-state/task-based functional magnetic resonance imaging, STG Superior temporal gyrus, sMRI structural magnetic resonance imaging.

were unavailable in some included studies even after contacting authors. We used scatter plots of all participants and high-resolution ROC to estimate them in order to include more studies in our analysis [44, 50, 80, 81]. Second, two of thirteen clinical studies included patients with bipolar depression (with a proportion of no more than 15%) [82, 83]. This overlap in illness risk may affect the predictive performance for pure MDD patients and also the comparisons of clinical and MRI studies. Further, many clinical studies (more than 20 studies published to date based on our rough search, such as these studies [84, 85]) used early response (i.e., clinical ratings after two or four weeks of treatment) rather than pretreatment baseline clinical ratings as input features to predict treatment response in MDD. To be consistent with the inclusion and exclusion criteria of MRI studies and to compare the clinical predictive results to MRI results that were predicted by using pretreatment brain MRI features, we only found and included 13 clinical studies taking pretreatment clinical ratings as predictors. Thus, due to the variations in study numbers between clinical and MRI studies in the present study, more future studies using pretreatment clinical features are needed to validate the present results. Third, similarly, the variations in patient numbers between fMRI and sMRI subgroups, as well as antidepressants and ECT subgroups, limit the ability to draw conclusive differences in predictive performance among treatment and imaging modality subgroups. Additionally, given the influence of sample size on ECT subgroup sensitivity, caution is warranted in interpreting our results. Fourth, subgroup analyses of rTMS, CBT, and DTI were not performed, due to the insufficient number of included studies ($n \leq 4$) to analyze and draw a robust SROC. Fifth, although included studies have defined treatment response and remission based on previous research or

professional consensus, there are still variations in definitions due to the use of different rating scales that impact the ability to model binarized outcome prediction [86]. Sixth, we did not incorporate validation as inclusion criteria to ensure a larger sample size, allowing for a comprehensive and robust analysis with enhanced statistical power and reliability. Most included studies utilized separate datasets for training, testing, and validating predictive models. Only two studies employed independent external validation datasets [43, 87]. While we were able to include more studies in our more up-to-date review of prior work, limitations in the available literature may lead to some degree of overfitting that would increase estimated model utility. Finally, the current meta-analysis combined studies employing varying modalities and interventions, and validation was not universal across all included studies. Consequently, the applicability of the current model to clinical practice remains limited.

CONCLUSION

The present findings revealed that pretreatment brain MRI features outperformed clinical characteristics in predicting short-term treatment outcomes in patients with MDD. The observed variations between rsfMRI and tbfMRI are also noteworthy. We found that rsfMRI biomarkers have higher accuracy in predicting non-responders/non-remitters than tbfMRI. These findings may be helpful in the early identification of patients who may not benefit from treatment, potentially aiding clinicians in considering alternative treatment options. Additional research is required to validate and expand upon these findings, particularly in exploring the predictive capabilities of specific MRI modalities and specific interventions.

DATA AVAILABILITY

The data that support the findings of this study, along with the code utilized in the Methods section, are available from the corresponding authors upon reasonable request.

REFERENCES

- World Health Organization. Depression fact sheet. World Health Organization. Published December 2019. Accessed November 20, 2022. <https://www.who.int/mediacentre/factsheets/fs369/en/>.
- Chen X, Lu B, Li HX, Li XY, Wang YW, Castellanos FX, et al. The DIRECT consortium and the REST-meta-MDD project: towards neuroimaging biomarkers of major depressive disorder. *Psychoradiology*. 2022;2:32–42.
- Jain FA, Connolly CG, Reus VI, Meyerhoff DJ, Yang TT, Mellon SH, et al. Cortisol, moderated by age, is associated with antidepressant treatment outcome and memory improvement in Major Depressive Disorder: A retrospective analysis. *Psychoneuroendocrinology*. 2019;109:104386.
- Sramek JJ, Murphy MF, Cutler NR. Sex differences in the psychopharmacological treatment of depression. *Dialogues Clin Neurosci*. 2016;18:447–57.
- Riedel M, Möller HJ, Obermeier M, Adli M, Bauer M, Kronmüller K, et al. Clinical predictors of response and remission in inpatients with depressive syndromes. *J Affect Disord*. 2011;133:137–49.
- Davis AK, Barrett FS, May DG, Cosimano MP, Sepeda ND, Johnson MW, et al. Effects of Psilocybin-Assisted Therapy on Major Depressive Disorder: A Randomized Clinical Trial. *JAMA Psychiatry*. 2021;78:481–89.
- Zhou Y, Zhang Z, Wang C, Lan X, Li W, Zhang M, et al. Predictors of 4-week antidepressant outcome in patients with first-episode major depressive disorder: An ROC curve analysis. *J Affect Disord*. 2022;304:59–65.
- Lee Y, Ragugett RM, Mansur RB, Boutilier JJ, Rosenblat JD, Trevizol A, et al. Applications of machine learning algorithms to predict therapeutic outcomes in depression: A meta-analysis and systematic review. *J Affect Disord*. 2018;241:519–32.
- Carstens L, Hartling C, Stippl A, Domke A-K, Herrera-Mendez AL, Aust S, et al. A symptom-based approach in predicting ECT outcome in depressed patients employing MADRS single items. *Eur Arc Psychiatry Clin Neurosci*. 2021;271:1275–84.
- Krepel N, Rush AJ, Iseger TA, Sack AT, Arns M. Can psychological features predict antidepressant response to rTMS? A Discovery-Replication approach. *Psychol Med*. 2020;50:264–72.
- Kautzky A, Dold M, Bartova L, Spies M, Vanicek T, Souery D, et al. Refining Prediction in Treatment-Resistant Depression: Results of Machine Learning Analyses in the TRD III Sample. *J Clin Psychiatry*. 2018;79:16m11385.
- Li F, Sun H, Biswal BB, Sweeney JA, Gong Q. Artificial intelligence applications in psychoradiology. *Psychoradiology*. 2021;1:94–107.
- Luo L, You W, DelBello MP, Gong Q, Li F. Recent advances in psychoradiology. *Phys Med Biol*. 2022;67:23TR01.
- You W, Luo L, Yao L, Zhao Y, Li Q, Wang Y, et al. Impaired dynamic functional brain properties and their relationship to symptoms in never treated first-episode patients with schizophrenia. *Schizophrenia (Heidelb)*. 2022;8:90.
- Luo L, Li Q, Wang Y, He N, Wang Y, You W, et al. Shared and Disorder-Specific Alterations of Brain Temporal Dynamics in Obsessive-Compulsive Disorder and Schizophrenia. *Schizophr Bull*. 2023;49:1387–98.
- Moreno-Ortega M, Prudic J, Rowny S, Patel GH, Kangarlu A, Lee S, et al. Resting state functional connectivity predictors of treatment response to electroconvulsive therapy in depression. *Sci Rep*. 2019;9:5071.
- Godlewska BR, Browning M, Norbury R, Igomenuou A, Cowen PJ, Harmer CJ. Predicting Treatment Response in Depression: The Role of Anterior Cingulate Cortex. *Int J Neuropsychopharmacol*. 2018;21:988–96.
- Hu X, Zhang L, Hu X, Lu L, Tang S, Li H, et al. Abnormal Hippocampal Subfields May Be Potential Predictors of Worse Early Response to Antidepressant Treatment in Drug-Naïve Patients With Major Depressive Disorder. *J Magn Reson Imaging*. 2019;49:1760–68.
- Wu P, Zhang A, Sun N, Lei L, Liu P, Wang Y, et al. Cortical Thickness Predicts Response Following 2 Weeks of SSRI Regimen in First-Episode, Drug-Naïve Major Depressive Disorder: An MRI Study. *Front Psychiatry*. 2021;12:751756.
- Cohen SE, Zantvoord JB, Wezenberg BN, Bockting CLH, van Wingen GA. Magnetic resonance imaging for individual prediction of treatment response in major depressive disorder: a systematic review and meta-analysis. *Transl Psychiatry*. 2021;11:168.
- Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Ann Intern Med*. 2009;151:W65–94.
- Doebler P. mada: Meta-Analysis of Diagnostic Accuracy. R package version 0.5.11. <https://r-forge.r-project.org/projects/mada/>. 2022.
- Friedman J, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw*. 2010;33:1–22.
- Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *Bmj*. 2003;327:557–60.
- Glas AS, Lijmer JG, Prins MH, Bossuyt PM. The diagnostic odds ratio: a single indicator of test performance. *J Clin Epidemiol*. 2003;56:1129–35.
- Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol*. 2005;58:982–90.
- Devillé WL, Buntinx F, Bouter LM, Montori VM, de Vet HC, van der Windt DA, et al. Conducting systematic reviews of diagnostic studies: didactic guidelines. *BMC Med Res Methodol*. 2002;2:9.
- Jones CM, Athanasiou T. Summary receiver operating characteristic curve analysis techniques in the evaluation of diagnostic tests. *Ann Thorac Surg*. 2005;79:16–20.
- Gowin JL, Manza P, Ramchandani VA, Volkow ND. Neuropsychosocial markers of binge drinking in young adults. *Mol Psychiatry*. 2021;26:4931–43.
- Gosnell SN, Curtis KN, Velasquez K, Fowler JC, Madan A, Goodman W, et al. Habenular connectivity may predict treatment response in depressed psychiatric inpatients. *J Affect Disord*. 2019;242:211–19.
- Jackson D, Turner R. Power analysis for random-effects meta-analysis. *Res Synth Methods*. 2017;8:290–302.
- Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *J Clin Epidemiol*. 2005;58:882–93.
- Macaskill PTY, Deeks JJ, Gatsonis C. Chapter 9: Understanding meta-analysis. Draft version (4 October 2022) for inclusion in: Deeks JJ, Bossuyt PM, Leeflang MM, Takwoingi Y, editor(s). *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 2*. London: Cochrane.
- Leucht S, Fennema H, Engel RR, Kaspers-Janssen M, Szegedi A. Translating the HAM-D into the MADRS and vice versa with equipercentile linking. *J Affect Disord*. 2018;226:326–31.
- Furukawa TA, Reijnders M, Kishimoto S, Sakata M, DeRubeis RJ, Dimidjian S, et al. Translating the BDI and BDI-II into the HAMD and vice versa with equipercentile linking. *Epidemiol Psychiatr Sci*. 2019 ;29:e24.
- Takwoingi YDN, Schiller I, Rücker G, Jones HE, Partlett C, Macaskill P. Chapter 10: Undertaking meta-analysis. Draft version (4 October 2022) for inclusion in: Deeks JJ, Bossuyt PM, Leeflang MM, Takwoingi Y, editor(s). *Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 2*. London: Cochrane.
- Hopman HJ, Chan SMS, Chu WCV, Lu H, Tse CY, Chau SWH, et al. Personalized prediction of transcranial magnetic stimulation clinical response in patients with treatment-refractory depression using neuroimaging biomarkers and machine learning. *J Affect Disord*. 2021;290:261–71.
- Nakamura T, Tomita M, Horikawa N, Ishibashi M, Uematsu K, Hiraki T, et al. Functional connectivity between the amygdala and subgenual cingulate gyrus predicts the antidepressant effects of ketamine in patients with treatment-resistant depression. *Neuropsychopharmacol Rep*. 2021;41:168–78.
- van Waarde JA, Scholte HS, van Oudheusden LJ, Verwey B, Denys D, van Wingen GA. A functional MRI marker may predict the outcome of electroconvulsive therapy in severe and treatment-resistant depression. *Mol Psychiatry*. 2015;20:609–14.
- Goldstein-Piekarski AN, Staveland BR, Ball TM, Yesavage J, Korgaonkar MS, Williams LM. Intrinsic functional connectivity predicts remission on antidepressants: a randomized controlled trial to identify clinically applicable imaging biomarkers. *Transl Psychiatry*. 2018;8:57.
- Braund TA, Breukelaar IA, Griffiths K, Tillman G, Palmer DM, Bryant R, et al. Intrinsic Functional Connectomes Characterize Neuroticism in Major Depressive Disorder and Predict Antidepressant Treatment Outcomes. *Biol Psychiatry Cogn Neurosci Neuroimaging*. 2021;7:276–84.
- Korgaonkar MS, Goldstein-Piekarski AN, Fornito A, Williams LM. Intrinsic connectomes are a predictive biomarker of remission in major depressive disorder. *Mol Psychiatry*. 2020;25:1537–49.
- Siegle GJ, Thompson WK, Collier A, Berman SR, Feldmiller J, Thase ME, et al. Toward clinically useful neuroimaging in depression treatment: prognostic utility of subgenual cingulate activity for determining depression outcome in cognitive therapy across studies, scanners, and patient characteristics. *Arch Gen Psychiatry*. 2012;69:913–24.
- Crane NA, Jenkins LM, Bhaumik R, Dion C, Gowins JR, Mickey BJ, et al. Multi-dimensional prediction of treatment response to antidepressants with cognitive control and functional MRI. *Brain*. 2017;140:472–86.
- Marquand AF, Mourão-Miranda J, Brammer MJ, Cleare AJ, Fu CH. Neuroanatomy of verbal working memory as a diagnostic biomarker for depression. *Neuroreport*. 2008;19:1507–11.

46. Costafreda SG, Khanna A, Mourao-Miranda J, Fu CH. Neural correlates of sad faces predict clinical remission to cognitive behavioural therapy in depression. *Neuroreport*. 2009;20:637–41.
47. Cao B, Luo Q, Fu Y, Du L, Qiu T, Yang X, et al. Predicting individual responses to the electroconvulsive therapy with hippocampal subfield volumes in major depressive disorder. *Sci Rep*. 2018;8:5434.
48. Costafreda SG, Chu C, Ashburner J, Fu CH. Prognostic and diagnostic potential of the structural neuroanatomy of depression. *PLoS One*. 2009;4:e6353.
49. Wade BSC, Sui J, Njau S, Leaver AM, Vasvada M, Gutman BA, et al. DATA-DRIVEN CLUSTER SELECTION FOR SUBCORTICAL SHAPE AND CORTICAL THICKNESS PREDICTS RECOVERY FROM DEPRESSIVE SYMPTOMS. *Proc IEEE Int Symp Biomed Imaging* 2017;2017:502–06.
50. Zhang F, Wang C, Lan X, Li W, Ye Y, Liu H, et al. Ketamine-induced hippocampal functional connectivity alterations associated with clinical remission in major depression. *J Affect Disord*. 2023;325:534–41.
51. Zhang F, Wang C, Lan X, Li W, Fu L, Ye Y, et al. The functional connectivity of the middle frontal cortex predicts ketamine's outcome in major depressive disorder. *Front Neurosci*. 2022;16:956056.
52. Redlich R, Opel N, Grotegerd D, Dohm K, Zaremba D, Bürger C, et al. Prediction of Individual Response to Electroconvulsive Therapy via Machine Learning on Structural Magnetic Resonance Imaging Data. *JAMA Psychiatry*. 2016;73:557–64.
53. Widge AS, Bilge MT, Montana R, Chang W, Rodriguez CI, Deckersbach T, et al. Electroencephalographic Biomarkers for Treatment Response Prediction in Major Depressive Illness: A Meta-Analysis. *Am J Psychiatry*. 2019;176:44–56.
54. Wagner S, Tadic A, Roll SC, Engel A, Dreimüller N, Engelmann J, et al. A combined marker of early non-improvement and the occurrence of melancholic features improve the treatment prediction in patients with Major Depressive Disorders. *J Affect Disord*. 2017;221:184–91.
55. Lin E, Kuo P-H, Liu Y-L, Yu WY, Yang AC, Tsai S-J. A Deep Learning Approach for Predicting Antidepressant Response in Major Depression Using Clinical and Genetic Biomarkers. *Front Psychiatry*. 2018;9:290.
56. Joyce JB, Grant CW, Liu D, MahmoudianDehkordi S, Kaddurah-Daouk R, Skime M, et al. Multi-omics driven predictions of response to acute phase combination antidepressant therapy: a machine learning approach with cross-trial replication. *Transl Psychiatry*. 2021;11:513.
57. Whitfield-Gabrieli S, Ford JM. Default mode network activity and connectivity in psychopathology. *Annu Rev Clin Psychol*. 2012;8:49–76.
58. Rolls ET. The cingulate cortex and limbic systems for emotion, action, and memory. *Brain Struct Funct*. 2019;224:3001–18.
59. Höflich A, Michenthaler P, Kasper S, Lanzenberger R. Circuit Mechanisms of Reward, Anhedonia, and Depression. *Int J Neuropsychopharmacol*. 2019;22:105–18.
60. Wu QZ, Li DM, Kuang WH, Zhang TJ, Lui S, Huang XQ, et al. Abnormal regional spontaneous neural activity in treatment-refractory depression revealed by resting-state fMRI. *Hum Brain Mapp*. 2011;32:1290–9.
61. Guo W, Liu F, Xue Z, Gao K, Liu Z, Xiao C, et al. Abnormal resting-state cerebellar-cerebral functional connectivity in treatment-resistant depression and treatment sensitive depression. *Prog Neuropsychopharmacol Biol Psychiatry*. 2013;44:51–7.
62. Ye Y, Wang C, Lan X, Li W, Fu L, Zhang F, et al. Baseline patterns of resting functional connectivity within posterior default-mode intranetwork associated with remission to antidepressants in major depressive disorder. *Neuroimage Clin*. 2022;36:103230.
63. Xiao H, Yuan M, Li H, Li S, Du Y, Wang M, et al. Functional connectivity of the hippocampus in predicting early antidepressant efficacy in patients with major depressive disorder. *J Affect Disord*. 2021;291:315–21.
64. Yamamura T, Okamoto Y, Okada G, Takaishi Y, Takamura M, Mantani A, et al. Association of thalamic hyperactivity with treatment-resistant depression and poor response in early treatment for major depression: a resting-state fMRI study using fractional amplitude of low-frequency fluctuations. *Transl Psychiatry*. 2016;6:e754.
65. Hou Z, Gong L, Zhi M, Yin Y, Zhang Y, Xie C, et al. Distinctive pretreatment features of bilateral nucleus accumbens networks predict early response to antidepressants in major depressive disorder. *Brain Imaging Behav*. 2018;12:1042–52.
66. Lemke H, Romankiewicz L, Förster K, Meinert S, Waltemate L, Fingas SM, et al. Association of disease course and brain structural alterations in major depressive disorder. *Depress Anxiety*. 2022;39:441–51.
67. Sheng W, Cui Q, Jiang K, Chen Y, Tang Q, Wang C, et al. Individual variation in brain network topology is linked to course of illness in major depressive disorder. *Cereb Cortex*. 2022;32:5301–10.
68. Fried EI. Moving forward: how depression heterogeneity hinders progress in treatment and research. *Expert Rev Neurother*. 2017;17:423–25.
69. Runge VM, Heverhagen JT. The Clinical Utility of Magnetic Resonance Imaging According to Field Strength, Specifically Addressing the Breadth of Current State-of-the-Art Systems, Which Include 0.55 T, 1.5 T, 3 T, and 7 T. *Invest Radio*. 2022;57:1–12.
70. Li Q, Zhao Y, Hu Y, Liu Y, Wang Y, Zhang Q, et al. Linked patterns of symptoms and cognitive covariation with functional brain controllability in major depressive disorder. *EBioMedicine*. 2024;106:105255.
71. Ai H, Opmeer EM, Marsman JC, Veltman DJ, van der Wee NJA, Aleman A, et al. Longitudinal brain changes in MDD during emotional encoding: effects of presence and persistence of symptomatology. *Psychol Med*. 2020;50:1316–26.
72. Li Q, Yao L, You W, Liu J, Deng S, Li B, et al. Controllability of Functional Brain Networks and Its Clinical Significance in First-Episode Schizophrenia. *Schizophr Bull*. 2023;49:659–68.
73. Ai Y, Li F, Hou Y, Li X, Li W, Qin K, et al. Differential cortical gray matter changes in early- and late-onset patients with amyotrophic lateral sclerosis. *Cereb Cortex*. 2024;34:bhad426.
74. Feng Y, Murphy MC, Hojo E, Li F, Roberts N. Magnetic Resonance Elastography in the Study of Neurodegenerative Diseases. *J Magn Reson Imaging*. 2024;59:82–96.
75. Wang Y, Li Q, Yao L, He N, Tang Y, Chen L, et al. Shared and differing functional connectivity abnormalities of the default mode network in mild cognitive impairment and Alzheimer's disease. *Cereb Cortex*. 2024;34:bhae094.
76. You W, Li Q, Chen L, He N, Li Y, Long F, et al. Common and distinct cortical thickness alterations in youth with autism spectrum disorder and attention-deficit/hyperactivity disorder. *BMC Med*. 2024;22:92.
77. Hahn A, Wadsak W, Windischberger C, Baldinger P, Höflich AS, Losak J, et al. Differential modulation of the default mode network via serotonin-1A receptors. *Proc Natl Acad Sci USA*. 2012;109:2619–24.
78. Chen F, Madsen TM, Wegener G, Nyengaard JR. Repeated electroconvulsive seizures increase the total number of synapses in adult male rat hippocampus. *Eur Neuropsychopharmacol*. 2009;19:329–38.
79. Rose D, Fleischmann P, Wykes T, Leese M, Bindman J. Patients' perspectives on electroconvulsive therapy: systematic review. *Bmj*. 2003;326:1363.
80. Wade BS, Joshi SH, Njau S, Leaver AM, Vasavada M, Woods RP, et al. Effect of Electroconvulsive Therapy on Striatal Morphometry in Major Depressive Disorder. *Neuropsychopharmacology*. 2016;41:2481–91.
81. Fisher PM, Ozenne B, Ganz M, Frokjaer VG, Dam VN, Penninx BW, et al. Emotional faces processing in major depressive disorder and prediction of antidepressant treatment response: A NeuroPharm study. *J Psychopharmacol*. 2022;36:626–36.
82. Brakemeier E-L, Luborzewski A, Danker-Hopfe H, Kathmann N, Bajbouj M. Positive predictors for antidepressive response to prefrontal repetitive transcranial magnetic stimulation (rTMS). *J Psychiatr Res*. 2007;41:395–403.
83. Brakemeier E-L, Wilbertz G, Rodax S, Danker-Hopfe H, Zinka B, Zwanzger P, et al. Patterns of response to repetitive transcranial magnetic stimulation (rTMS) in major depression: Replication study in drug-free patients. *J Affect Disord*. 2008;108:59–70.
84. Lin H-S, Lin C-H. Early improvement in HAMD-17 and HAMD-6 scores predicts ultimate response and remission for depressed patients treated with fluoxetine or ECT. *J Affect Disord*. 2019;245:91–97.
85. Mohamed AK, Croarkin PE, Jha MK, Voort JLV. Early reduction in irritability is associated with improved outcomes among youth with depression: Findings from the AMOD study. *J Affect Disord*. 2023;324:77–81.
86. Leaver AM, Wade B, Vasavada M, Helleman G, Joshi SH, Espinoza R, et al. Fronto-Temporal Connectivity Predicts ECT Outcome in Major Depression. *Front Psychiatry*. 2018;9:92.
87. 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.
88. Kautzky A, Baldinger-Melich P, Kranz GS, Vanicek T, Souery D, Montgomery S, et al. A New Prediction Model for Evaluating Treatment-Resistant Depression. *J Clin Psychiatry*. 2017 Feb;78:215–+.
89. Perlis RH. A Clinical Risk Stratification Tool for Predicting Treatment Resistance in Major Depressive Disorder. *Biol Psychiatry*. 2013;74:7–14.
90. Su L, Zhang Y, Jia Y, Sun J, Mellor D, Yuan T-F, et al. Predictors of Electroconvulsive Therapy Outcome in Major Depressive Disorder. *Int J Neuropsychopharmacol*. 2023;26:53–60.
91. Rezaei M, Bagheri MMS, Ahmadi M. Clinical and demographic predictors of response to anodal tDCS treatment in major depression disorder (MDD). *J Psychiatr Res*. 2021;138:68–74.
92. Donse L, Padberg F, Sack AT, Rush AJ, Arns M. Simultaneous rTMS and psychotherapy in major depressive disorder: Clinical outcomes and predictors from a large naturalistic study. *Brain Stimul*. 2018;11:337–45.
93. Xue SW, Kuai C, Xiao Y, Zhao L, Lan Z. Abnormal Dynamic Functional Connectivity of the Left RostrHippocampus in Predicting Antidepressant Efficacy in Major Depressive Disorder. *Psychiatry Investig*. 2022;19:562–69.

94. Wu H, Liu R, Zhou J, Feng L, Wang Y, Chen X, et al. Prediction of remission among patients with a major depressive disorder based on the resting-state functional connectivity of emotion regulation networks. *Transl Psychiatry*. 2022;12:391.
95. Pei C, Sun Y, Zhu J, Wang X, Zhang Y, Zhang S, et al. Ensemble Learning for Early-Response Prediction of Antidepressant Treatment in Major Depressive Disorder. *J Magn Reson Imaging*. 2020;52:161–71.
96. Tian S, Sun Y, Shao J, Zhang S, Mo Z, Liu X, et al. Predicting escitalopram monotherapy response in depression: The role of anterior cingulate cortex. *Hum Brain Mapp*. 2020;41:1249–60.
97. Zhu J, Cai H, Yuan Y, Yue Y, Jiang D, Chen C, et al. Variance of the global signal as a pretreatment predictor of antidepressant treatment response in drug-naïve major depressive disorder. *Brain Imaging Behav*. 2018;12:1768–74.
98. Cash RFH, Cocchi L, Anderson R, Rogachov A, Kucyi A, Barnett AJ, et al. A multivariate neuroimaging biomarker of individual outcome to transcranial magnetic stimulation in depression. *Hum Brain Mapp*. 2019;40:4618–29.
99. Ge R, Downar J, Blumberger DM, Daskalakis ZJ, Vila-Rodriguez F. Functional connectivity of the anterior cingulate cortex predicts treatment outcome for rTMS in treatment-resistant depression at 3-month follow-up. *Brain Stimul*. 2020;13:206–14.
100. Sun H, Jiang R, Qi S, Narr KL, Wade BS, Upston J, et al. Preliminary prediction of individual response to electroconvulsive therapy using whole-brain functional magnetic resonance imaging data. *Neuroimage Clin*. 2020;26:102080.
101. Meyer BM, Rabl U, Huemer J, Bartova L, Kalcher K, Provenzano J, et al. Prefrontal networks dynamically related to recovery from major depressive disorder: a longitudinal pharmacological fMRI study. *Transl Psychiatry*. 2019;9:64.
102. Goldstein-Piekarski AN, Korgaonkar MS, Green E, Suppes T, Schatzberg AF, Hastie T, et al. Human amygdala engagement moderated by early life stress exposure is a biobehavioral target for predicting recovery on antidepressants. *Proc Natl Acad Sci USA*. 2016;113:11955–60.
103. Williams LM, Korgaonkar MS, Song YC, Paton R, Eagles S, Goldstein-Piekarski A, et al. Amygdala Reactivity to Emotional Faces in the Prediction of General and Medication-Specific Responses to Antidepressant Treatment in the Randomized iSPOT-D Trial. *Neuropsychopharmacology*. 2015;40:2398–408.
104. Korgaonkar MS, Williams LM, Song YJ, Usherwood T, Grieve SM. Diffusion tensor imaging predictors of treatment outcomes in major depressive disorder. *Br J Psychiatry*. 2014;205:321–8.
105. Queirazza F, Fouragnan E, Steele JD, Cavanagh J, Philiastides MG. Neural correlates of weighted reward prediction error during reinforcement learning classify response to cognitive behavioral therapy in depression. *Sci Adv*. 2019;5:eaav4962.
106. Gong Q, Wu Q, Scarpazza C, Lui S, Jia Z, Marquand A, et al. Prognostic prediction of therapeutic response in depression using high-field MR imaging. *Neuroimage*. 2011;55:1497–503.
107. Nouretdinov I, Costafreda SG, Gammerman A, Chervonenkis A, Vovk V, Vapnik V, et al. Machine learning classification with confidence: application of transductive conformal predictors to MRI-based diagnostic and prognostic markers in depression. *Neuroimage*. 2011;56:809–13.
108. Xu J, Li W, Bai T, Li J, Zhang J, Hu Q, et al. Volume of hippocampus-amygdala transition area predicts outcomes of electroconvulsive therapy in major depressive disorder: high accuracy validated in two independent cohorts. *Psychol Med*. 2022;53:4464–73.

AUTHOR CONTRIBUTIONS

Fei Li conceptualized the study. Fenghua Long, Yufei Chen, Qian Zhang, John A. Sweeney, and Fei Li designed and drafted the manuscript. Fenghua Long, Yufei Chen, Qian Zhang, Youjin Zhao, Qian Li, Yitian Wang, Haoran Li, and Yaxuan Wang contributed to the literature search, data collection and analysis. Fenghua Long, Yufei Chen, Qian Zhang, Robert K. McNamara, Melissa P. DelBello, John A. Sweeney, Qiyong Gong, and Fei Li critically revised the manuscript, and contributed to the data analysis strategy and interpretation. All authors approved the final version of the manuscript.

FUNDING

This study was supported by the National Key R&D Program (2022YFC2009900) and National Natural Science Foundation (Grant Nos. 82001795 to Youjin Zhao, and 82027808 to Qiyong Gong) of China, and Sichuan Science and Technology Program (2024NSFSC0653 to Fei Li).

COMPETING INTERESTS

Melissa P. DelBello receives research support from national institutes of health (NIH), PCORI, Acadia, Allergan, Janssen, Johnson and Johnson, Lundbeck, Otsuka, Pfizer, and Sunovion. She is also a consultant, on the advisory board, or has received honoraria for speaking for Alkermes, Allergan, Assurex, CMEology, Janssen, Johnson and Johnson, Lundbeck, Myriad, Neuronetics, Otsuka, Pfizer, Sunovion, and Supernus. Robert K. McNamara has received research support from Martek Biosciences Inc, Royal DSM Nutritional Products, LLC, Inflammation Research Foundation, Ortho-McNeil Janssen, AstraZeneca, Eli Lilly, NARSAD, and NIH, and previously served on the scientific advisory board of the Inflammation Research Foundation. The remaining authors declare no potential conflicts of interest with regard to this manuscript.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41380-024-02710-6>.

Correspondence and requests for materials should be addressed to Fei Li.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.