



OPEN ACCESS

EDITED BY

Kevin C. Chan,
Grossman School of Medicine,
New York University, United States

REVIEWED BY

Alexandre Rosa Franco,
Nathan Kline Institute for Psychiatric
Research, United States
Xiu-Xia Xing,
Beijing University of Technology, China

*CORRESPONDENCE

Aiqin Peng
Pengaiqinsy@126.com

†These authors have contributed
equally to this work

SPECIALTY SECTION

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

RECEIVED 24 August 2022

ACCEPTED 31 October 2022

PUBLISHED 29 November 2022

CITATION

Wang R, Shen Y, Li G, Du R and
Peng A (2022) Quantitative magnetic
resonance spectroscopy
of depression: The value
of short-term metabolite changes
in predicting treatment response.
Front. Neurosci. 16:1025882.
doi: 10.3389/fnins.2022.1025882

COPYRIGHT

© 2022 Wang, Shen, Li, Du and Peng.
This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Quantitative magnetic resonance spectroscopy of depression: The value of short-term metabolite changes in predicting treatment response

Ranchao Wang^{1†}, Yu Shen^{1†}, Guohai Li², Rui Du¹ and
Aiqin Peng^{3*}

¹Department of Radiology, Affiliated Hospital of Jiangsu University, Zhenjiang, China, ²Department of Clinical Psychology, Zhenjiang Mental Health Center, Zhenjiang, China, ³Department of Radiology, Affiliated Shuyang Hospital of Xuzhou Medical University, Shuyang, China

Background: Although various prediction models of the antidepressant response have been established, the results have not been effectively applied to heterogeneous depression populations, which has seriously limited their clinical value. This study tried to build a more specific and stable model to predict treatment response in depression based on short-term changes in hippocampal metabolites.

Materials and methods: Seventy-four major depressive disorder (MDD) patients and 20 healthy controls in the test set were prospectively collected and retrospectively analyzed. Subjects underwent magnetic resonance spectroscopy (MRS) once a week during 6 weeks of treatment. Hippocampal regions of interest (ROIs) were extracted by using a voxel iteration scheme combined with standard brain templates. The short-term differences in hippocampal metabolites between and within groups were screened. Then, the association between hippocampal metabolite changes and clinical response was analyzed, and a prediction model based on logistic regression was constructed. In addition, a validation set ($n = 60$) was collected from another medical center to validate the predictive abilities.

Results: After 2–3 weeks of antidepressant treatment, the differences in indicators ($t\text{Cho}_{\text{week}0-2}$, $t\text{Cho}_{\text{week}0-3}$ and $\text{NAA}_{\text{week}0-3}$) were successfully screened. Then, the predictive abilities of these three indicators were revealed in the logistic regression model, and the optimal prediction effect was found in $d(t\text{Cho})_{\text{week}0-3}-d(\text{NAA})_{\text{week}0-3}$ ($\text{AUC} = 0.841$, $95\% \text{CI} = 0.736-0.946$). In addition, their predictive abilities were further confirmed with the validation set.

Limitations: The small sample size and the need for multiple follow-ups limited the statistical ability to detect other findings.

Conclusion: The predictive model in this study presented accurate prediction and strong verification effects, which may provide early guidance for adjusting the treatment regimens of depression and serve as a checkpoint at which the eventual treatment outcome can be predicted.

KEYWORDS

hippocampus, magnetic resonance spectroscopy, depression, prediction, treatment response

Introduction

Unlike other diseases that have definitive diagnoses (Barbatesi et al., 2022; Panza and Lozupone, 2022), the mechanism of depression remains unclear (Li et al., 2022). Thus, a variety of treatment programs have been explored. Treatment response is the most important clinical indicator that determines both the sensitivity and treatment outcomes of antidepressant programs (Drysdale and Patel, 2022). In the clinical practice of depression, at least 6 weeks of observation must be taken to confirm an ineffective treatment response (Thase and Rush, 1997; Souery et al., 1999; Berlim and Turecki, 2007). Therefore, breaking the time window of 6-weeks and assessing the treatment response as early as possible will be of great clinical significance, especially for patients who may progress to refractory depression (RD). The earlier confirmation of treatment response and timely selection of another regimen (e.g., electroconvulsive therapy, magnetic stimulation, or new combination therapies) (Mitchell and Loo, 2006; Subramanian et al., 2022) may be more effective in improving their condition.

Presently, prediction is the practical solution to break the time window of treatment response. Numerous prediction studies from hospitals and laboratories have been carried out with different test or experimental methods, following a similar model: screening out the difference indicators between responding and non-responding groups at endpoints retrospectively or prospectively and predicting the treatment response at baseline (Williams et al., 2016; Emam et al., 2019; He et al., 2019). These studies can be further divided into three indicator categories: 1) gene or molecular indicators, such as inflammatory factors, VEGF (vascular endothelial growth factor), and BDNF (brain-derived neurotrophic factor); 2) tissue

level indicators (mainly functional MRI), such as default mode network and functional connectivity (FC); and 3) behavioral indicators, such as childhood events and early life with parental loss or separation.

These reports have provided positive clinical and academic value for assessing and predicting depression treatment response, which also introduces two questions worth exploring. First, compared with the other two indicators (molecular and behavioral indicators), the tissue level indicator (BOLD, blood oxygenation level dependent) shows a significant advantage in non-invasiveness, convenience, and repeatability but is limited to the measurement of neural activity, which only provides indirect rather than direct indicators (Siegle et al., 2007). Second, one model showing good prediction ability in a certain sample set or laboratory may not be applicable in other depression populations that underwent different treatments or had different patient characteristics because various cofactors cannot be adequately corrected in this model based on baseline prediction. Few reliable predictors indicate which depressed individuals respond to antidepressants, as previously reported by Williams et al. (2016). For example, a history of early life trauma predicts a poorer response to antidepressant therapy, but the results are variable and limited in adults (Dunn et al., 2022). In addition to simple expansion of the sample size to improve the stability of the predictive model, the above deficiencies may also be overcome by optimizing the model.

In this study, we intended to explore the correlation between short-term metabolite changes and subsequent treatment response in depression by detecting the hippocampus (the classical treatment response region of the brain) (Huang et al., 2010) dynamically with quantitative MRS combined with a novel voxel-placement technique. Then, we built a relatively more specific and stable prediction model that provides a reference upon which to break the 6-week window of treatment response.

Materials and methods

Participants and procedure

Test set In total, 74 major depressive disorder (MDD) patients and 20 healthy controls (HCs) were prospectively

Abbreviations: AUC, area under curve; tCho, total choline (phosphocholine and glycerophosphocholine); tCr, total creatine (creatinine and phosphocreatine); CSF, cerebrospinal fluid; PRESS, point resolved spectroscopy; FC, functional connectivity; FOV, field of view; FWHM, full width at half maximum; CRLBs, Cramér-Rao lower bounds; GLM, general linear model; Glx, glutamate and glutamine; GM, gray matter; HC, healthy control; HDRS-17, Hamilton Depression Rating Scale-17; Ins, myo-inositol; MRS, magnetic resonance spectroscopy; MDD, major depressive disorder; NAA, N-acetyl-aspartate; n-RD, non-refractory depression; RD, refractory depression; ROC, receiver operating characteristic; ROI, region of interest; SNR, signal-to-noise ratio; WM, white matter.

recruited from March 2017 to March 2019 at Zhenjiang Mental Health Center. All participants were examined at baseline by an experienced psychiatrist or clinical psychologist using the Screening Interview from the Structured Clinical Interview of the DSM-IV (SCID) to assess depression (Williams, 1988; Maffei et al., 1997) and met the following inclusion criteria: (1) aged 18–60 years; (2) Hamilton Depression Scale-17 (HDRS-17) scores of depression ≥ 17 ; HDRS-17 scores of healthy controls (HCs) < 7 ; (3) no general developmental disorder or mental retardation; (4) educational level above junior high school; (5) Chinese Han nationality and right-handed; and (6) voluntary participation in the 6-week follow-up. In addition, subject with any of the following condition was excluded: (1) treatment from any other psychiatric disorders; (2) pregnant or lactating female; (3) any other neurological disorders; (4) organic disorders or somatic complaints in the brain; (5) a history of alcohol or drug abuse; (6) magnetic resonance contraindications; and (7) received any treatment within 2 weeks before enrollment. After enrollment, all MDD patients were treated with citalopram for 3 weeks to observe the efficacy ($n = 67$, average dose 34.6 mg/d). Treatment responders continued to maintain the medication ($n = 29$, average dose 32.8 mg/d), and treatment non-responders were switched to the next stage for augmented treatment with bupropion ($n = 38$, average dose 237.6 mg/d) (Bech et al., 2012). During the treatment, the drug dose was increased or decreased according to the individual condition of the patients. The antidepressant effect was judged by the HAMD score reduction rate. Treatment responders were those who showed either a partial or complete response to treatment, conventionally defined as a 25–50% or $> 50\%$ reduction in HAMD scores, and treatment non-responders were those who showed a $< 25\%$ reduction in HAMD scores (Drysdale et al., 2017). The subjects were then divided into RD and n-RD groups according to a HAMD score reduction rate of 50% after 6 weeks of individualized treatment. After excluding poor image quality and data of failure to follow-up, the remaining follow-up data were used for analysis. *Validation set* Another 60 depression subjects were recruited from the Affiliated Hospital of Jiangsu University from June 2019 to June 2021 for validation. The inclusion and exclusion criteria were the same as those for the test set. These subjects received corresponding treatment and underwent MRS scans at baseline and during the second and third weeks. For each patient, demographic information, past medical history and medication information were collected. The information helped the psychiatrist make decisions about individualized treatment. The study protocol was approved by the institutional ethics committee (ZJJS-2017017), and informed consent was obtained from all participants. The study protocol conformed to the ethical guidelines of the 1975 Declaration of Helsinki. Participant details and follow-up procedures are shown in **Figure 1** and **Table 1**.

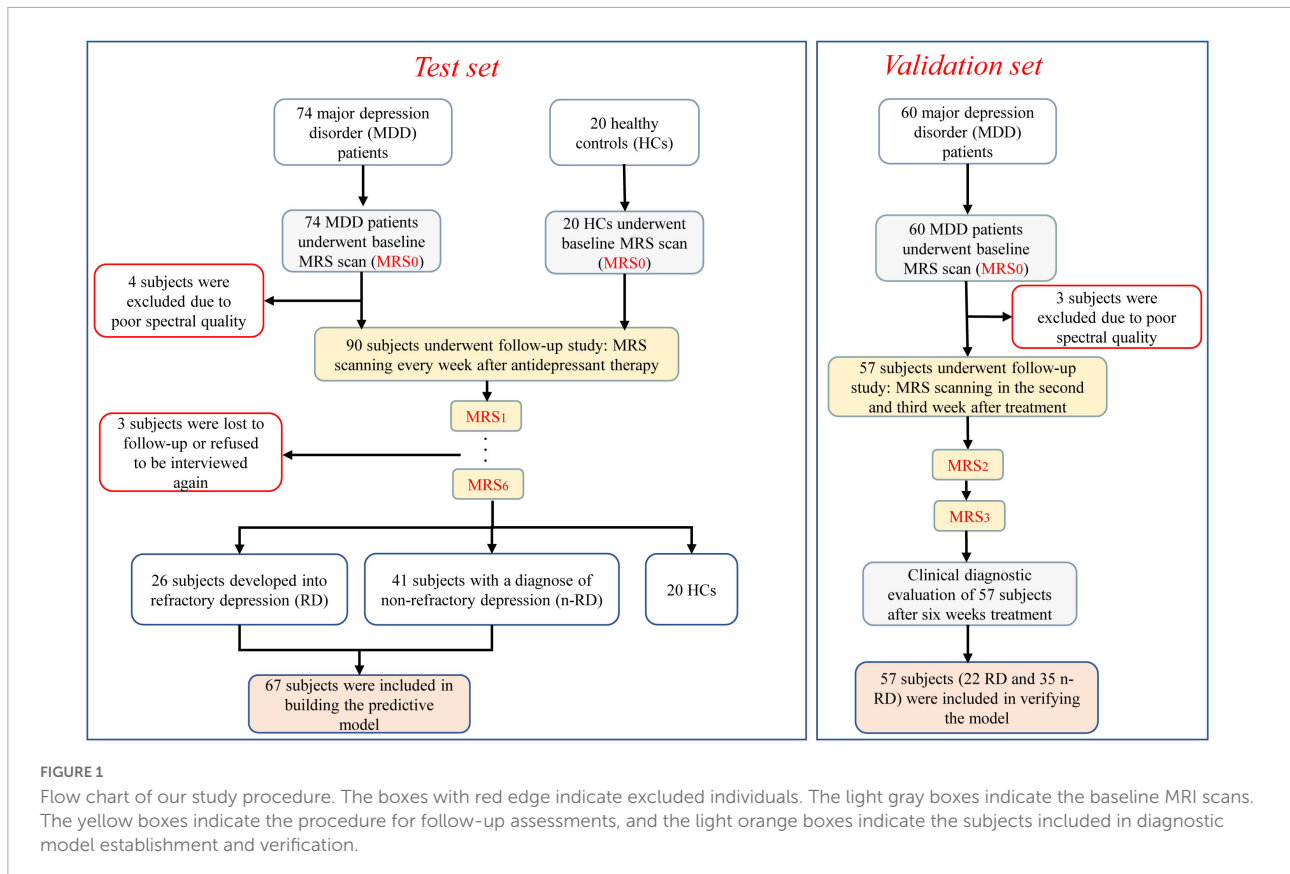
MRI acquisition

MRI acquisitions were performed on a Siemens 3.0T Trio MR scanner using an eight-channel head coil. Subjects were instructed to keep their eyes closed, relax, remain immobile, think of nothing in particular, and avoid falling asleep. Initial image acquisition included a T1-weighted image scan acquired with a Magnetization Prepared Rapid Gradient Echo (MPRAGE) sequence (TR/TE/flip angle = 2,530 ms/2.26 ms/90°; voxel size = $1 \times 1 \times 1$ mm; FOV = 256×256 mm²; matrix = 256×256 ; slice thickness = 1 mm). Additional sequences (such as T2 and FLAIR) were performed to ensure that all studied participants were MRI-negative (no imaging findings of organic lesions). Axial and coronal images were reconstructed based on the sagittal images for MRS localization.

Firstly, the brain template (Montreal Neurological Institute) was registered to the subject's DICOM images to acquire the transformation parameters (Dou et al., 2015). Simultaneously, the subject's standard hippocampus was extracted with the FMRIB software library (FSL)¹ and the ROI (the region of interest) mask was created (size: $10 \times 10 \times 15$ mm³). Secondly, the in-house developed software was designed to achieve maximum overlap between ROI and hippocampus. The in-house developed software was based on the iterative algorithm procedure, formulated in MATLAB (MATLAB 2015b; Mathworks, Natick, MA) and described as: 1. the overlapping volume of ROI placement and hippocampus area was calculated; 2. the spatial position of the ROI was constantly adjusted, and its overlapping volume was recorded; and 3. the position with the largest overlapping volume in all calculation results was compared and determined. Then, Voxel's coordinates, size, angulation about each axis, image orientation ("LAS") were recorded in "Voxel_Location.Txt" file for rapid and accurate voxel placement of future follow-up (Woodcock et al., 2018). Thirdly, the transformation parameters were used to acquire the subject's voxel placement mapped by the voxel mask in atlas space (**Figure 2A**). The voxel overlap was defined as the percentage of the week 0 voxel volume encompassed by the follow-up voxel. The geometric voxel overlap suggested high accuracy of voxel placement across all subjects (mean overlap of each subject's voxel = $93.4\% \pm 4.3\%$ during the 6-week follow-up).

Single-voxel spectra (SVS) were acquired by using a standard point resolved spectroscopy (PRESS) sequence. Water-suppressed SVS was performed with VAPOR water suppression and the following parameters: echo time (TE) of 35 ms, repetition time (TR) of 2,000 ms, nominal voxel size: $10 \times 10 \times 15$ mm³, spectral width of 5,000 Hz, 2048 time points, and 128 averages. Spatial saturation pulses were applied to minimize contamination of the signal from outside the voxel.

¹ www.fmrib.ox.ac.uk/fsl



Linear shims were used to correct the B0 inhomogeneity across the investigated voxel. Water MR spectroscopic spectra were also acquired without water suppression on the same voxel, with TE = 20 ms and all other parameters remaining the same. The acquisition in the oblique coronal plane was perpendicular to the long axis of the hippocampus.

Magnetic resonance spectroscopy data processing

The spectra were pre-processed (including phased and apodized 1 Hz) (Tiwari et al., 2020; Near et al., 2021) and then analyzed with software jMRUI (version 5.2). The time-domain quantification of metabolite signals was conducted using AMARES algorithm with custom prior knowledge. Metabolite concentrations were reported for tCr (creatine plus phosphocreatine), NAA (N-acetyl-aspartate), tCho (phosphocholine and glycerophosphocholine), Ins (myo-inositol) and Glx (glutamate and glutamine). The AMARES prior knowledge model consisted of peaks for NAA, choline (Cho), creatine (Cr), glutamate + glutamine (Glx) and myo-inositol (Ins). The amplitudes of NAA, Cho, Cr, Glx, and Ins peak were estimated by the algorithm. The relative phases of NAA, Cho, Cr, Glx, and Ins peak were fixed at 0. The linewidth

of NAA was estimated by the algorithm, and the linewidths of the remaining peaks were set to be equal to that of NAA. The frequencies of NAA, Cho, Cr, Glx, and Ins peak were estimated by AMARES. All peak shapes were fixed at Lorentzian. Data were subjected to quality control prior to inclusion in the analysis. We required spectra to meet the following criteria to be eligible for inclusion: the signal-to-noise ratio (SNR) ≥ 15 , the full width at half maximum (FWHM) ≤ 16 Hz, and the Cramér-Rao lower bounds (CRLBs) $< 20\%$ (Supplementary Table 3). The spectral analysis window was defined as 0–4.0 ppm (Figure 2B). Absolute concentrations of metabolites were calculated using water signal from the identical voxel as internal reference. The relaxation times (T1 and T2) of water and respective metabolites measured at 3 T were used for relaxation correction. The concentrations were calculated according to Kreis et al. (Kreis et al., 1997) as follows: $C_M = (S_M/S_W) \times C_W \times (n_W/n_M) \times (f_W^{T1}/f_M^{T1}) \times (f_W^{T2}/f_M^{T2})$, where indexes M for metabolite and W for water, C stands for concentration, S for signal intensity, n for the number of chemically equivalent protons (contributing to the signal), f^{T1} for spin-lattice relaxation function ($1 - e^{-TR/T1}$), f^{T2} for spin-spin relaxation function ($e^{-TR/T2}$). C_W stands for concentration of water in white matter, which is 55.51 moles/kg.² To correct the metabolites' concentration of cerebrospinal fluid (CSF) contamination, the CSF, gray matter (GM) and white matter

TABLE 1 Demographic and clinical characteristics.

	Test set MDD (<i>n</i> = 67)					Validation set MDD (<i>n</i> = 57)			
	HC (<i>n</i> = 20)	RD (<i>n</i> = 26)	n-RD (<i>n</i> = 41)	<i>t</i> / <i>x</i> ² -value, df RD vs. n-RD	<i>P</i> -value RD vs. n-RD	RD (<i>n</i> = 22)	n-RD (<i>n</i> = 35)	<i>t</i> / <i>x</i> ² -value, df RD vs. n-RD	<i>P</i> -value RD vs. n-RD
Age, years	30.2 ± 6.8	32 ± 8.4	28 ± 7.8	1.985, 65	0.051	34 ± 8.8	31 ± 8.6	1.271, 55	0.209
Gender, male/female	9/11	10/16	15/26	0.024, 1	0.877	8/14	13/22	0.004, 1	0.953
Education time, years	15.2 ± 2.8	14.8 ± 2.7	13.9 ± 3.1	1.216, 65	0.228	14.5 ± 3.3	13.7 ± 2.8	0.980, 55	0.331
Marital status, married/unmarried	11/9	12/14	18/23	0.033, 1	0.857	10/12	16/19	<0.001, 1	0.985
Age of onset, years	NA	21.8 ± 3.8	24.7 ± 4.1	2.901, 65	0.005	22.6 ± 4.3	25.4 ± 3.4	2.73, 55	0.009
Total duration of illness, years	NA	7.9 ± 3.2	6.2 ± 3.4	2.040, 65	0.045	8.1 ± 4.2	6.1 ± 3.7	1.886, 55	0.065
No medication (%)	20	18(69.23%)	28(68.29%)	0.007, 1	0.936	15(68.18%)	24(68.57%)	<0.001, 1	0.975
Antidepressants (%)		8(30.77%)	13(31.71%)			7(31.82%)	11(31.43%)		
SSRIs	NA	4	8	0.364, 2	0.834	4	7	0.276, 2	0.871
SNRIs	NA	2	3			2	2		
NaSSAs	NA	2	2			1	2		
HDRS-17 _(week0)	5.5 ± 1.2	25.2 ± 4.1	24.6 ± 3.8	0.611	0.543	25.7 ± 3.7	24.8 ± 3.5	0.925	0.359
HDRS-17 _(week6)	5.4 ± 1.3	18.3 ± 5.4	7.6 ± 3.7	9.631	<0.001	16.8 ± 4.8	6.8 ± 3.2	9.450	<0.001

MDD, major depression disorder; HC, healthy control; RD, refractory depression; n-RD, non-refractory depression; HDRS-17, 17-item Hamilton Rating Scale for Depression; Antidepressants are taken at least 3 months ago; SSRIs, selective serotonin reuptake inhibitors; SNRIs, serotonin-norepinephrine reuptake inhibitors; NaSSAs, noradrenergic and specific serotonergic antidepressant; df, degree of freedom; Data are shown as mean ± SD; NA, not available.

(WM) volumes were segmented and calculated from T1-weighted images by FSL. The calibration formula (Near et al., 2021) was as follows: $C_{cor} = C_{raw} \times [V_{total}/V_{total}-V_{CSF}]$, where C_{cor} denoted the corrected value; C_{raw} , the uncorrected value; V_{total} , the voxel volume; V_{CSF} , the CSF volume (Supplementary Table 4). In brief, voxel tissue composition and unsuppressed endogenous water were used to calculate and calibrate absolute concentrations of metabolites. The mean metabolite concentration of bilateral ROI was used in analysis. Test-retest reliability analysis gave ICC of 0.78, 0.75, 0.88, and 0.83 (NAA, Cr, Cho, and Ins) in healthy control group according to the metabolic measurements of week 0 and week 1, indicating excellent reliability, while the Glx (0.64) was not.

Statistical analysis

Statistical analyses were performed using GraphPad Prism 7 (GraphPad Software Inc., La Jolla, CA) and R 3.1.2 software (R Foundation for Statistical Computing, Vienna, Austria). All analyses were two tailed with an alpha level of 5%. Data were tested for normality (Shapiro–Wilk test) and homogeneity (Levene's test). Clinical and demographic data are presented as the means and SD, with appropriate tests for intergroup comparison (t -test for continuous data and X^2 for categorical data). Test-retest reliability was calculated in healthy control group based on metabolites concentrations including NAA, Cr, Cho, Ins and Glx in week 0 and week 1, which was assessed with the intraclass correlation coefficient (ICC). The ICC values > 0.7 indicated good reliability. Trend analyses between groups were performed using ANOVA to detect whether the metabolite levels changed with increasing treatment time. The intragroup [$RD_{week(n)}$ vs. $RD_{baseline}$, $n-RD_{week(n)}$ vs. $n-RD_{baseline}$] and intergroup [$RD_{week(n)}$ vs. $n-RD_{week(n)}$] differences in hippocampal metabolite concentrations were assessed by t -test. All differences were further verified by a general linear model (GLM) correcting covariates (age, gender, age of onset, total duration of illness). All differentiated changes were defined as the week of follow-up minus the week of baseline. The relationship between the changed HDRS scores and changed metabolite concentrations in the short term was assessed using Pearson's correlation coefficients. Stepwise logistic regression was used to screen the changed hippocampal metabolites with predictive capability in the short term to discriminate RD and n-RD with correcting covariates (age, gender, age of onset, total duration of illness). Then, the receiver operating characteristic (ROC) curve was used to analyze the performance of predictors in terms of AUC (area under curve), sensitivity and specificity. The statistical methods used in the validation cohort were consistent with those used in the test cohort.

2 <https://pubchem.ncbi.nlm.nih.gov/compound/Water>

Results

Longitudinal evaluations of hippocampal metabolites in week 0–week 6

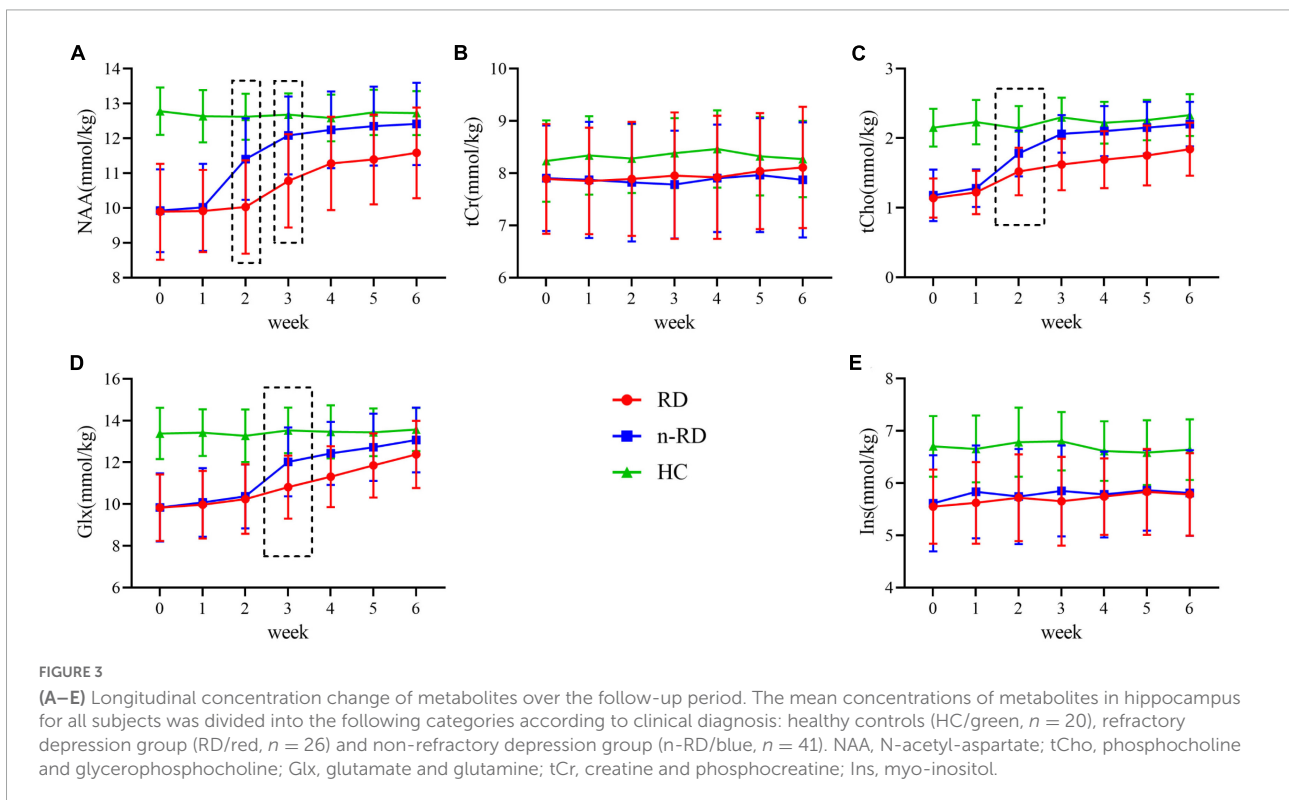
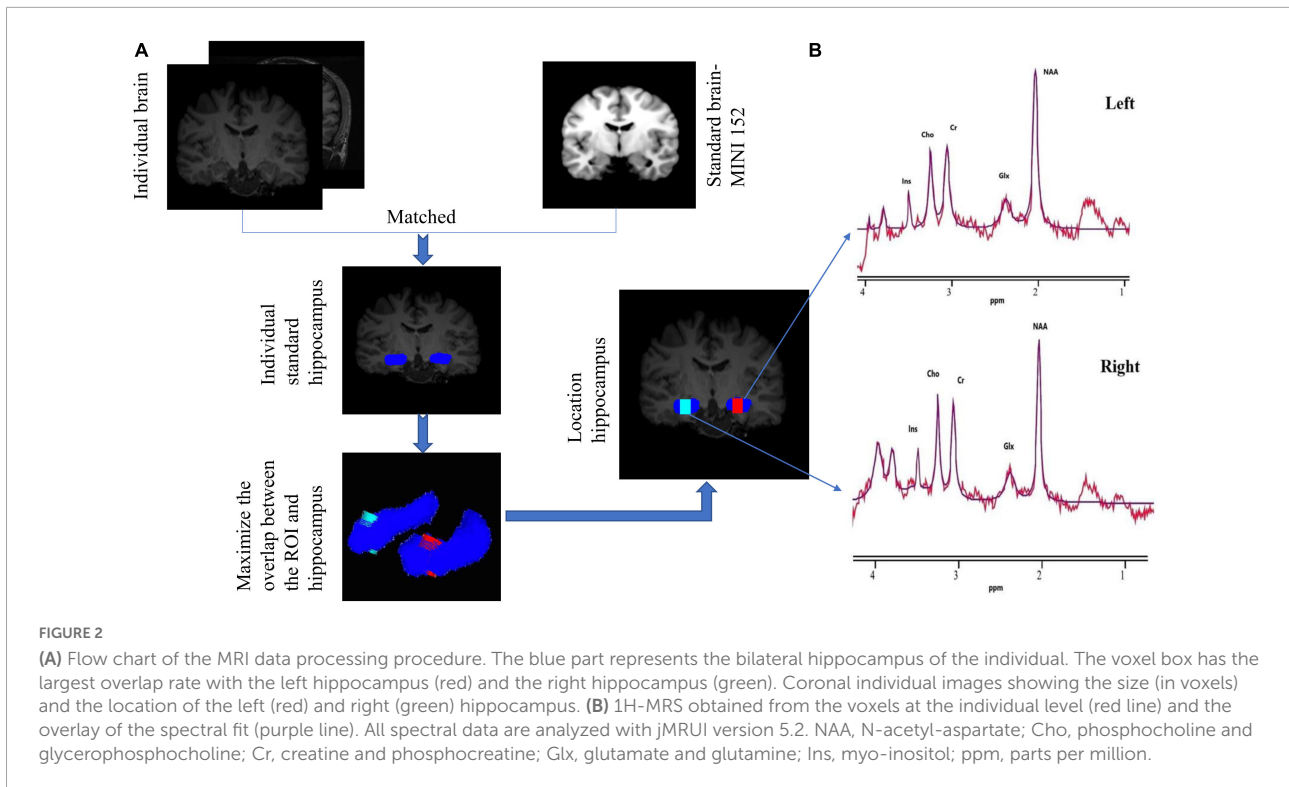
Beginning in the second week, the tCho concentration showed a gradual increase both in the RD (p for trend < 0.001 ; $t_{week0-2} = 4.40$, $p < 0.001$) and n-RD groups (p for trend < 0.001 ; $t_{week0-2} = 7.75$, $p < 0.001$). The Glx concentration showed a similar trend beginning in the third week (RD: p for trend < 0.001 ; $t_{week0-3} = 2.28$, $p = 0.027$; n-RD: p for trend < 0.001 ; $t_{week0-3} = 6.02$, $p < 0.001$). However, there was a slight difference in the trend of NAA concentration between RD and n-RD; the former increased after the third week (p for trend < 0.001 ; $t_{week0-3} = 2.36$, $p = 0.022$), while the latter increased after the second week (p for trend < 0.001 ; $t_{week0-2} = 4.52$, $p < 0.001$). Furthermore, there was no significant longitudinal trend in Ins and tCr (all p for trend > 0.05), which indicated that they might be unrelated to the treatment response (Figure 3). Detailed data are shown in Supplementary Material 1.

Metabolite changes after short-term treatment

Given the above results, indicators of tCho, Glx, and NAA were used for further observation of the treatment response. In the intragroup comparison, after two weeks of antidepressant treatment (week 0–week 2), the concentration of tCho significantly increased in both the RD and n-RD groups, and the NAA concentration only increased in the n-RD group (Figures 4A,C and Supplementary Table 1). After three weeks (week 0–week 3), significant increases in NAA, tCho, and Glx concentrations were found in both the RD and n-RD groups (Figures 4B,D and Supplementary Table 1). In the intergroup comparison, compared with the RD group, the concentrations of tCho and NAA in the n-RD group were significantly higher at the second week (tCho $t = 3.11$, $p = 0.003$; NAA $t = 4.53$, $p < 0.001$) (Figures 4E,F), and the tCho, NAA and Glx concentrations were significantly higher at the third week (tCho $t = 5.62$, $p < 0.001$; NAA $t = 4.29$, $p < 0.001$; Glx $t = 3.02$, $p = 0.004$) (Figures 4G–I).

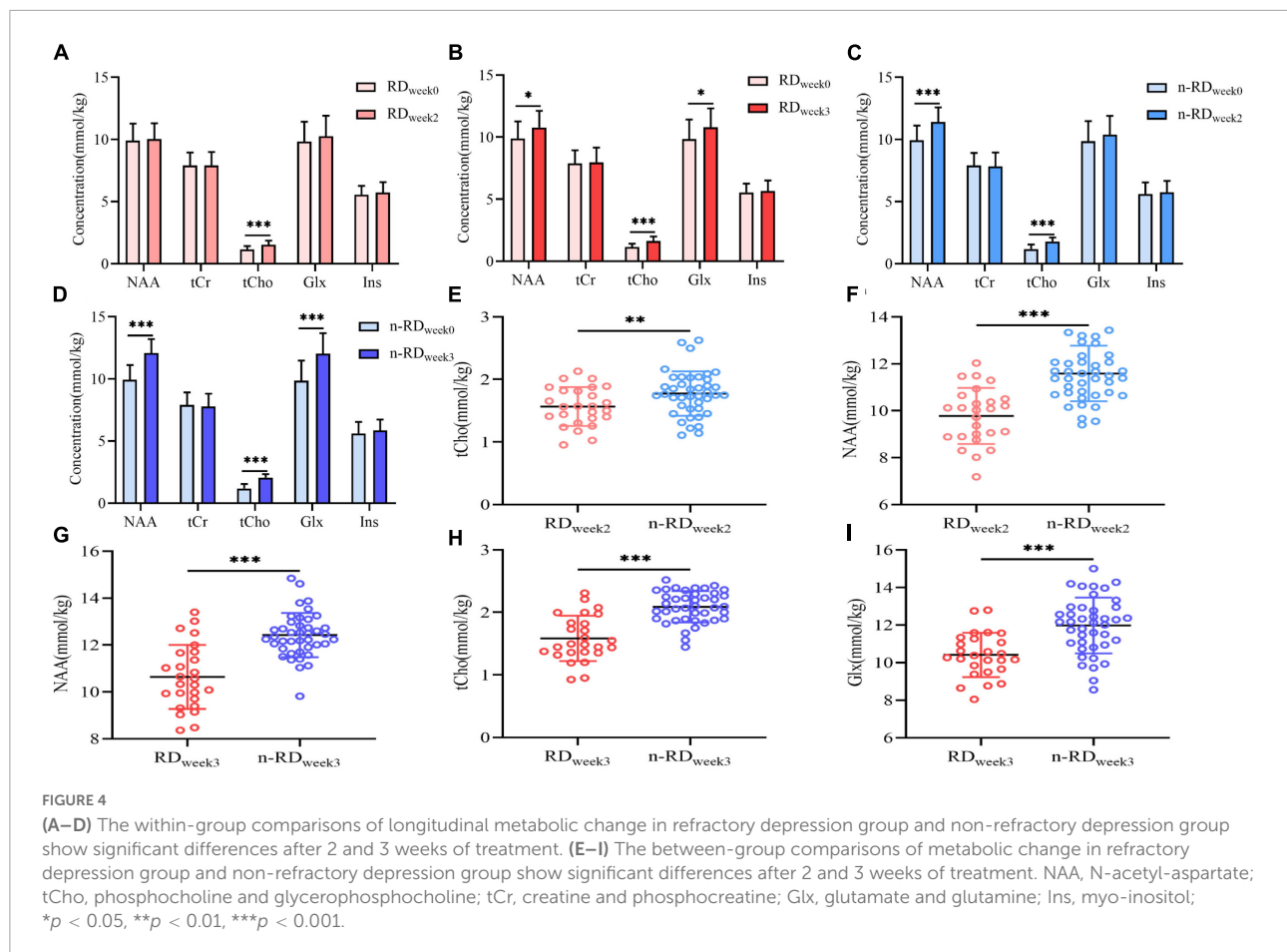
Correlation between metabolite concentration and clinical response

Improved HDRS scores of MDD were found until the third week ($t_{week0-3} = 5.71$, $p < 0.001$), in which a definite decrease in HDRS scores was found in the n-RD group since the second



week ($t_{week0-2} = 2.15, p = 0.037$) and the RD group since the third week ($t_{week0-3} = 2.29, p = 0.024$) (Figure 5A). Subsequent correlation analysis showed that differentiated

HDRS (dHDRS) was in negatively correlated with differentiated tCho [d(tCho)] ($r = -0.639, p < 0.001$) but not with differentiated NAA (dNAA) ($r = -0.169, p = 0.173$) in week



0-week 2 (**Figures 5B,C**), and dHDRS was negatively correlated with $d(tCho)$ ($r = -0.827$, $p < 0.001$) and $dNAA$ ($r = -0.512$, $p < 0.001$), but not with differentiated Glx ($dGlx$) ($r = -0.168$, $p = 0.174$) in week 0-week 3 (**Figures 5D–F**).

Prediction of treatment response

Based on the results of the correlation analysis, $dHDRS_{week0-2} + d(tCho)_{week0-2}$ and $dHDRS_{week0-3} + d(tCho)_{week0-3} + dNAA_{week0-week3}$ were included in further stepwise logistic regression analysis with other factors (age, gender, age of onset, total duration of illness). The results showed that $d(tCho)_{week0-2}$ was an independent predictor for treatment response at the second week (OR = 0.429, $p = 0.01$; AUC = 0.684), but $dHDRS_{week0-2}$ was not (OR = 1.495, $p = 0.371$). At the third week, although $dHDRS_{week0-3}$ presented a non-neglectable predictive value (OR = 3.179, $p = 0.041$; AUC = 0.708), better predictive capabilities were found with $d(tCho)_{week0-3}$ (OR = 0.115, $p < 0.001$; AUC = 0.779) and $dNAA_{week0-3}$ (OR = 0.117, $p = 0.033$; AUC = 0.752). Furthermore, improved capability was obtained with the combined index [$d(tCho)_{week0-3}$ and

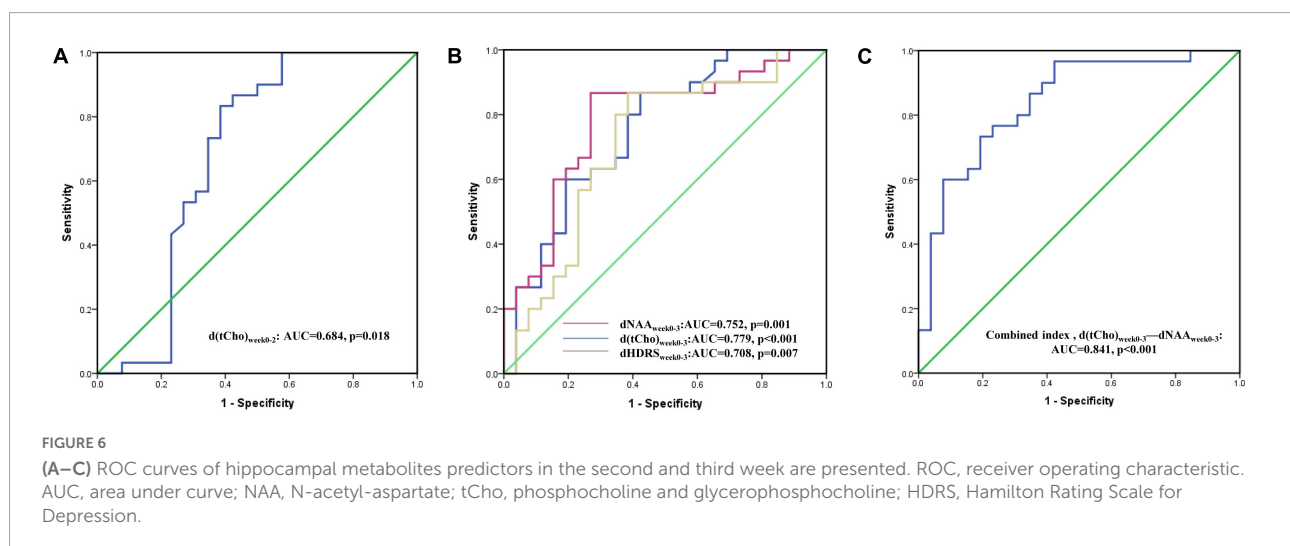
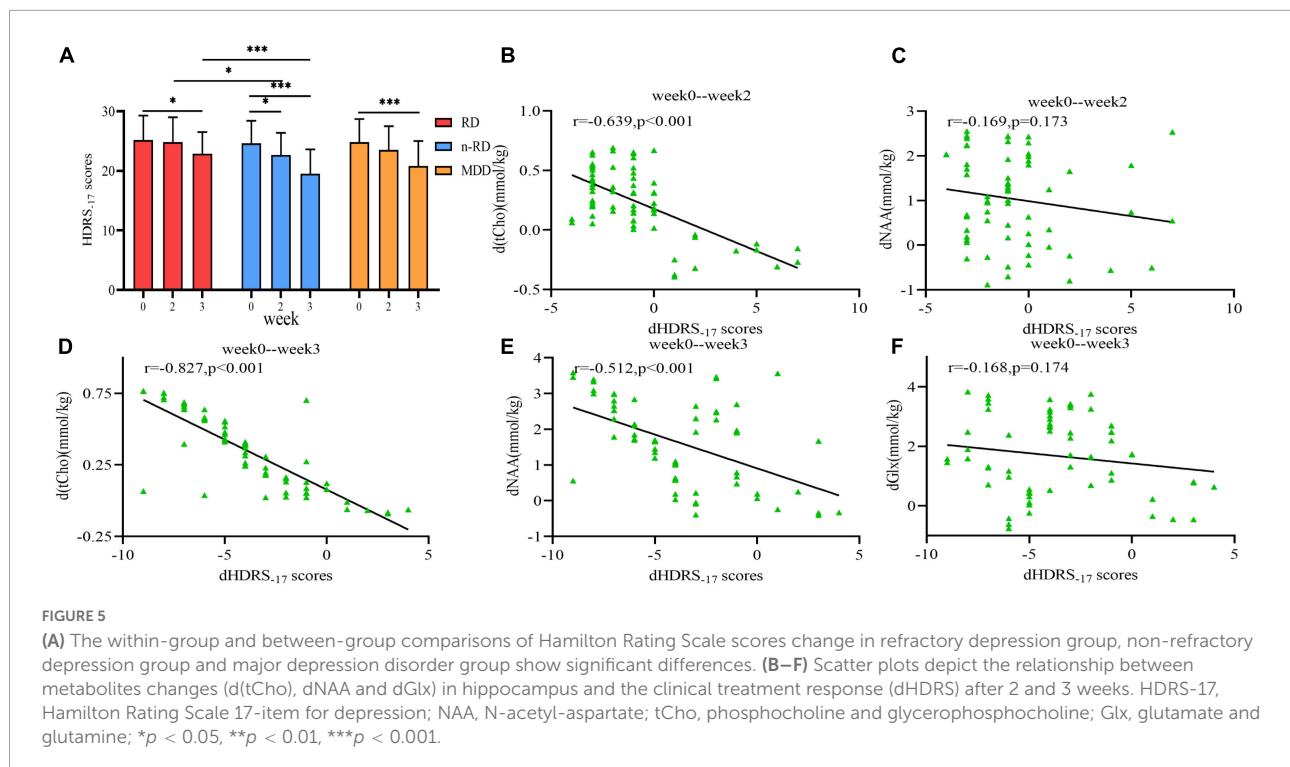
$dNAA_{week0-3}$] (AUC = 0.841, $p < 0.001$) (**Figure 6** and **Table 2**).

Model validation

In the validation set of 57 subjects (22 RD and 35 n-RD patients), good diagnostic value was obtained with $d(tCho)_{week0-2}$ (accuracy = 68.42%, AUC = 0.708) after two weeks of treatment, and better performance was found in $d(tCho)_{week0-3}$ (accuracy = 75.44%, AUC = 0.785), $dNAA_{week0-3}$ (accuracy = 71.93%, AUC = 0.722) and $d(tCho)_{week0-3} - dNAA_{week0-3}$ (accuracy = 85.96%, AUC = 0.837) after 3 weeks of treatment (**Table 3**). Detailed data are shown in **Supplementary Table 2**.

Discussion

In this study, combining quantitative MRS with a new method of ROI positioning, we conducted a longitudinal follow-up from baseline to 6 weeks in MDD and constructed an improved prediction model, in which several interesting



findings were reported: (i) After 2 weeks of treatment, the changed tCho concentration could accurately predict the subsequent treatment response, but the changed HDRS score could not. (ii) After 3 weeks of treatment, although the changed HDRS score could predict treatment response, indicators from MRS (changed tCho and NAA) showed a stronger predictive power. (iii) The new ROI positioning strategy and predictive model presented a more stable verification capability.

Previous studies have indicated that hippocampal metabolites at baseline are positive predictors that suggest functional conditions, which may dominate the treatment

responses (Block et al., 2009). In fact, this predictive model was an ideal model based on standard conditions, in which the individual differences, disease status and treatment plans were ignored, which might be why the predictive results of different laboratories could not be unified or well verified. In our relatively conservative predictive model, the predicted time point was placed 2 or 3 weeks after short-term treatment. In other words, the change in indicators from week 0 to week 2 or week 3 was used to predict the outcome of treatment response, and our results also yielded a positive verification effect in the validation group. In brief, the predictive model of this study was

TABLE 2 Logistic regression analysis in check point week 2 and week 3.

Variables	β	S.E.	Wald	P	OR (95% CI)
dHDRS _{week0-2}	0.402	0.450	0.798	0.371	1.495 (0.619~3.614)
d(tCho) _{week0-2}	-0.846	0.328	6.704	0.01	0.429 (0.251~0.735)
d(tCho) _{week0-3}	-2.167	0.352	37.897	<0.001	0.115 (0.057~0.228)
dNAA _{week0-3}	-2.145	1.002	4.582	0.033	0.117 (0.024~0.608)
dHDRS _{week0-3}	1.156	0.544	3.672	0.041	3.179 (1.836~5.501)

d(tCho)_{week0-2}, differentiated tCho (phosphocholine and glycerophosphocholine) after two weeks of treatment; d(tCho)_{week0-3}, differentiated tCho after three weeks of treatment; dNAA_{week0-3}, differentiated NAA (N-acetyl-aspartate) after three weeks of treatment; dHDRS_{week0-2}, differentiated HDRS (Hamilton Depression Rating Scale) scores after 2 weeks of treatment; dHDRS_{week0-3}, differentiated HDRS scores after 3 weeks of treatment.

reported for the first time in imaging research on depression, which brings the novelty to the field.

In addition to the joint application of programming language and FSL segmentation, this study also applied a novel voxel-placement technique. The realization of ROI localization with accurate overlap in multiple scans is the key and difficult point of MRS studies. The perfect match was considered impossible by hand sketching in the existing studies, especially in the irregular and small gray matter nuclei such as the hippocampus and amygdala. Compared with other existed approaches, the highlight of our technique was the maximum overlap region which was acquired by iterating the ROI and the individual standard hippocampus. We meant to make the voxel include as much hippocampus tissue and little contamination from surrounding tissue as possible, which could make it more representative. Other approaches were developed to improving voxel placement technique for reliable voxel coregistration within- and between-subjects, and we had benefited from these approaches. For example, after the maximum overlap region between ROI and hippocampus was determined, we also used the transformation parameters which were calculated from the registration of the atlas to the skull-stripped subject's brain to acquire the subject's voxel placement mapped by the voxel mask in atlas space (Park et al., 2018). And we also created the file which includes specific individual voxel information to ensure high precision division, maximum overlap and high repeatability in future follow-up (Woodcock et al., 2018). Besides, our technique has a similarity that the brain template space was used for registration. However, other research (Hancu et al., 2005; Dou et al., 2015) used the registration technique based on the first scan session. This precise and stable voxel placement method might deserve further improvement and promotion in sensitive quantitative single-voxel NMR spectroscopy.

In general, studies on hippocampal choline levels and their changes during treatment have given conflicting findings. Some studies have reported increased choline levels at baseline (Milne et al., 2009), while others have shown different results: major depressive disorder (MDD) patients with first episodes had a trend toward lower choline levels, and those with remitted recurrence had higher choline levels than controls (de Diego-Adeliño et al., 2013). MDD patients with a first episode accounted for a large proportion of the subjects in our study. In addition, depression is a heterogeneous disease in which different patients might have significant individual differences (manifested in symptoms, treatment differences, etc.), which might impede the consistency of choline level reports. Above all, the contribution of surrounding tissue (non-hippocampal region in voxel) to choline levels remains unclear. We optimized the placement of voxels in the hippocampus to include more hippocampal tissue and less surrounding tissue to make it more representative, while the quantitative region centered on the hippocampus contained more contributions from surrounding tissues based on traditional voxel placement (Milne et al., 2009). These may be the reasons why the choline level in MDD patients reported by us was lower than that in some other studies. Some studies have reported no change in choline levels during treatment (Wang et al., 2012). However, others have reported that the NAA and choline increase in the hippocampus in association with pharmacological treatment response and that these changes are applicable particularly for patients with low NAA and Cho baseline levels (Block et al., 2009). The different choline levels and their change in treatment reported by different studies may implicate different pathophysiological grounds in MDD.

Cr, a significant marker of material metabolism, is the buffer that maintains the cell energy-dependent system by adjusting adenosine triphosphate (ATP) and adenosine diphosphate (Mountford et al., 2010). Cr concentrations have been considered to be relatively constant under normal conditions; thus, the ratio to Cr is widely used as an internal standard to scale other metabolites in traditional relative quantitative MRS. In fact, emerging evidence has found Cr concentrations in the brains of depressed patients are abnormal because of decreased mitochondrial ATP production and mitochondrial enzyme levels (Baxter et al., 1989; Gardner, 2003), which suggests that evaluating metabolism in depression with relative quantification might not be accurate, highlighting the necessity of absolute quantitative MRS. In this study, we quantified metabolite intensities by referencing internal water, which is the mature quantitative scheme preferred in clinical ¹H-MRS. Additionally, the metabolic concentration of the hippocampus we measured was consistent with existing reports.

Although our study found that the dHDRS score (after 3 weeks) was an independent positive predictor, the data from

TABLE 3 Diagnosis accuracy of the MRS metabolites predictors.

Actual	Predictor							
	d(tCho) _{week0–2}		d(tCho) _{week0–3}		dNAA _{week0–3}		d(tCho) _{week0–3} -dNAA _{week0–3}	
	RD	n-RD	RD	n-RD	RD	n-RD	RD	n-RD
RD	14	8	16	6	16	6	18	4
n-RD	10	25	8	27	10	25	4	31
Accuracy (%)	68.42		75.44		71.93		85.96	

RD, refractory depression; n-RD, non-refractory depression; d(tCho)_{week0–2}, differentiated tCho (phosphocholine and glycerophosphocholine) after 2 weeks of treatment; d(tCho)_{week0–3}, differentiated tCho after three weeks of treatment; dNAA_{week0–3}, differentiated NAA (N-acetyl-aspartate) after three weeks of treatment; d(tCho)_{week0–3}-dNAA_{week0–3}, the combined detection of dNAA and d(tCho) after 3 weeks of treatment.

the spectrum clearly showed a stronger prediction ability, which was reflected in the earlier time [2-week d(tCho)] and higher accuracy (3-week d(tCho), dNAA). NAA is a neuron internal marker that can reflect the functional status and integrity of neurons in the brain. Choline reflects cell membrane transport, which is generally assumed to play a key role in energy metabolism and myelination and has been proposed as a marker in pathological membrane renewal and cell membrane transport (Moffett et al., 2007; Oz et al., 2014). Patients who showed a > 50% reduction in HAMD scores after 6 weeks had higher levels of NAA or Cho and earlier clinical improvement, indicating that an increase in NAA or Cho is associated with treatment response. The differences in tCho and NAA between the RD group and the n-RD group, as well as their excellent predictive abilities, also indicated that the integrity of neurons, energy metabolism and myelination might be treatment targets for refractory depression in the future. In addition, compared with the rating scale of the neuropsychiatric disease evaluation system (HDRS), objective and sensitive brain material metabolism indicators may provide strong supplementary evidence.

Nevertheless, several limitations of this study should be noted. Firstly, refractory depression was defined as having no response to treatment with two antidepressants for 6–8 weeks. Although it is a commonly held view by psychiatrists, the optimal duration for a standard course of treatment has not been fully defined, and the definition of non-response is not completely clear. Therefore, the standard for the diagnosis of refractory depression may not be completely accurate. In addition, it is well demonstrated that glutamate and glutamine at 3.0T MRS are difficult to resolve and that the measurement of glutamate is likely to be affected by glutamine, albeit to a small degree. Thus, we took glutamate and glutamine as the whole Glx into consideration, but the change in glutamate could not be accurately estimated, which possibly affect the reliability. Poor test-retest reliability and difficult measurement of glutamate possibly result in poor correlation between Glx concentration and clinical improvement. We cannot rule out

the effect that the sample size is limited due to difficulties in collecting clinical cases, and the need for multiple follow-ups limits the statistical ability to detect other findings. Moreover, due to the complexity of the pathogenesis of depression, a more sensitive and more specific logistic model and ROC curve should be generated, which requires further exploration. Despite these limitations, this study provides a way to predict the efficacy of antidepressants at an early stage with improved reliability. Once validated, biomarkers and the clinical assessment of patients with major depression could support psychiatrists' diagnostic and treatment decisions and could increase the rationality of treatment.

Conclusion

We developed a robust model to predict antidepressant responses based on short-term treatment changes, which may provide early guidance for adjusting treatment regimens for depression and serve as a checkpoint upon which the eventual outcome of conventional treatments can be predicted, reducing the time and resources wasted on ineffective treatment.

Data availability statement

The data that support the findings of this study are available from Zhenjiang Mental Health Center but restrictions apply to the availability of these data, which were used only under license for the current study, and so are not publicly available. Requests to access the datasets should be directed to Yuefeng Li, jiangdalyf2009@126.com.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of Zhenjiang Mental

Health Center. The patients/participants provided their written informed consent to participate in this study.

Author contributions

RW, YS, and AP made a substantial contribution to the concept and design, analysis, and interpretation of data. RW and YS drafted the article and processed MRI data. AP revised the article critically for important intellectual content. RD organized the study and supported the data analysis. All authors were critically involved in the theoretical discussion and performing of the experiments and read and approved the final version of the manuscript.

Funding

This work was supported by the Jiangsu Provincial Key Research and Development Plan (BE2021692 and BE2017698).

References

- Barbaresi, W., Cacia, J., Friedman, S., Fussell, J., Hansen, R., Hofer, J., et al. (2022). Clinician Diagnostic Certainty and the Role of the Autism Diagnostic Observation Schedule in Autism Spectrum Disorder Diagnosis in Young Children. *JAMA Pediatr.* [Epub ahead of print]. doi: 10.1001/jamapediatrics.2022.3605
- Baxter, L. R., Schwartz, J. M., Phelps, M. E., Mazziotta, J. C., Guze, B. H., Selin, C. E., et al. (1989). Reduction of prefrontal cortex glucose metabolism common to three types of depression. *Arch. Gen. Psychiatry* 46, 243–250. doi: 10.1001/archpsyc.1989.01810030049007
- Bech, P., Fava, M., Trivedi, M. H., Wisniewski, S. R., and Rush, A. J. (2012). Outcomes on the pharmacopsychometric triangle in bupropion-SR vs. bupropion augmentation of citalopram in the STAR*D trial. *Acta Psychiatr. Scand.* 125, 342–348. doi: 10.1111/j.1600-0447.2011.01791.x
- Berlim, M. T., and Turecki, G. (2007). Definition, assessment, and staging of treatment-resistant refractory major depression: A review of current concepts and methods. *Can. J. Psychiatry* 52, 46–54. doi: 10.1177/070674370705200108
- Block, W., Träber, F., von Widdern, O., Metten, M., Schild, H., Maier, W., et al. (2009). Proton MR spectroscopy of the hippocampus at 3 T in patients with unipolar major depressive disorder: Correlates and predictors of treatment response. *Int. J. Neuropsychopharmacol.* 12, 415–422. doi: 10.1017/S1461145708009516
- de Diego-Adelino, J., Portella, M. J., Gómez-Ansón, B., López-Moruelo, O., Serra-Blasco, M., Vives, Y., et al. (2013). Hippocampal abnormalities of glutamate/glutamine, N-acetylaspartate and choline in patients with depression are related to past illness burden. *J. Psychiatry Neurosci.* 38, 107–116. doi: 10.1503/jpn.110185
- Dou, W., Speck, O., Benner, T., Kaufmann, J., Li, M., Zhong, K., et al. (2015). Automatic voxel positioning for MRS at 7 T. *Magma* 28, 259–270. doi: 10.1007/s10334-014-0469-9
- Drysdale, A. T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., et al. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat. Med.* 23, 28–38. doi: 10.1038/nm.4246
- Drysdale, A. T., and Patel, G. H. (2022). A Dynamic Approach to Depression Treatment Prediction. *Biol. Psychiatry* 92, 526–527. doi: 10.1016/j.biopsych.2022.06.028
- Dunn, K. E., Turner, G. M., and Oswald, L. M. (2022). Effects of early life trauma on risks for adult opioid use disorder are mediated by stress and occur independent of depression and anxiety. *J. Addict. Med.* [Epub ahead of print]. doi: 10.1097/ADM.0000000000001011
- Emam, H., Steffens, D. C., Pearlson, G. D., and Wang, L. (2019). Increased ventromedial prefrontal cortex activity and connectivity predict poor sertraline treatment outcome in late-life depression. *Int. J. Geriatr. Psychiatry* 34, 730–737. doi: 10.1002/gps.5079
- Gardner, A. (2003). Alterations of mitochondrial function and correlations with personality traits in selected major depressive disorder patients. *J. Affect. Disord.* 76, 55–68. doi: 10.1016/s0165-0327(02)00067-8
- Hancu, I., Blezek, D. J., and Dumoulin, M. C. (2005). Automatic repositioning of single voxels in longitudinal 1H MRS studies. *NMR Biomed.* 18, 352–361. doi: 10.1002/nbm.965
- He, S., Zeng, D., Xu, F., Zhang, J., Zhao, N., Wang, Q., et al. (2019). Baseline Serum Levels of Beclin-1, but Not Inflammatory Factors, May Predict Antidepressant Treatment Response in Chinese Han Patients With MDD: A Preliminary Study. *Front. Psychiatry* 10:378. doi: 10.3389/fpsy.2019.00378
- Huang, Y., Chen, W., Li, Y., Wu, X., Shi, X., and Geng, D. (2010). Effects of antidepressant treatment on N-acetyl aspartate and choline levels in the hippocampus and thalami of post-stroke depression patients: A study using (1)H magnetic resonance spectroscopy. *Psychiatry Res.* 182, 48–52. doi: 10.1016/j.psychres.2009.11.009
- Kreis, R., Koster, M., Kamber, M., Hoppeler, H., and Boesch, C. (1997). Peak assignment in localized 1H MR spectra of human muscle based on oral creatine supplementation. *Magn. Reson. Med.* 37, 159–163. doi: 10.1002/mrm.1910370202
- Li, Y., Xie, Y., Xu, Y., Xian, X., Wang, R., Cai, L., et al. (2022). Interleukin-6-white matter network differences explained the susceptibility to depression after stressful life events. *J. Affect. Disord.* 305, 122–132. doi: 10.1016/j.jad.2022.03.003
- Maffei, C., Fossati, A., Agostoni, I., Barraco, A., Bagnato, M., Deborah, D., et al. (1997). Interrater reliability and internal consistency of the structured clinical interview for DSM-IV axis II personality disorders (SCID-II), version 2.0. *J. Pers. Disord.* 11, 279–284. doi: 10.1521/pedi.1997.11.3.279

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2022.1025882/full#supplementary-material>

- Milne, A., MacQueen, G. M., Yucel, K., Soreni, N., and Hall, G. B. (2009). Hippocampal metabolic abnormalities at first onset and with recurrent episodes of a major depressive disorder: A proton magnetic resonance spectroscopy study. *Neuroimage* 47, 36–41. doi: 10.1016/j.neuroimage.2009.03.031
- Mitchell, P. B., and Loo, C. K. (2006). Transcranial magnetic stimulation for depression. *Aust. N. Z. J. Psychiatry* 40, 406–413. doi: 10.1080/j.1440-1614.2006.01816.x
- Moffett, J. R., Ross, B., Arun, P., Madhavarao, C. N., and Nambodiri, A. M. (2007). N-Acetylaspartate in the CNS: From neurodiagnostics to neurobiology. *Prog. Neurobiol.* 81, 89–131. doi: 10.1016/j.pneurobio.2006.12.003
- Mountford, C. E., Stanwell, P., Lin, A., Ramadan, S., and Ross, B. (2010). Neurospectroscopy: The past, present and future. *Chem. Rev.* 110, 3060–3086. doi: 10.1021/cr900250y
- Near, J., Harris, A. D., Juchem, C., Kreis, R., Marjańska, M., Öz, G., et al. (2021). Preprocessing, analysis and quantification in single-voxel magnetic resonance spectroscopy: Experts' consensus recommendations. *NMR Biomed.* 34:e4257. doi: 10.1002/nbm.4257
- Oz, G., Alger, J. R., Barker, P. B., Bartha, R., Bizzi, A., Boesch, C., et al. (2014). Clinical proton MR spectroscopy in central nervous system disorders. *Radiology* 270, 658–679. doi: 10.1148/radiol.13130531
- Panza, F., and Lozupone, M. (2022). The challenges of anti-tau therapeutics in Alzheimer disease. *Nat. Rev. Neurol.* 18, 577–578. doi: 10.1038/s41582-022-00702-0
- Park, Y. W., Deelchand, D. K., Joers, J. M., Hanna, B., Berrington, A., Gillen, J. S., et al. (2018). AutoVOI: Real-time automatic prescription of volume-of-interest for single voxel spectroscopy. *Magn. Reson. Med.* 80, 1787–1798. doi: 10.1002/mrm.27203
- Siegle, G. J., Thompson, W., Carter, C. S., Steinhauer, S. R., and Thase, M. E. (2007). Increased amygdala and decreased dorsolateral prefrontal BOLD responses in unipolar depression: Related and independent features. *Biol. Psychiatry* 61, 198–209.
- Souery, D., Amsterdam, J., de Montigny, C., Lecrubier, Y., Montgomery, S., Lipp, O., et al. (1999). Treatment resistant depression: Methodological overview and operational criteria. *Eur. Neuropsychopharmacol.* 9, 83–91.
- Subramanian, S., Lopez, R., Zorumski, C. F., and Cristancho, P. (2022). Electroconvulsive therapy in treatment resistant depression. *J. Neurol. Sci.* 434:120095.
- Thase, M. E., and Rush, A. J. (1997). When at first you don't succeed: Sequential strategies for antidepressant nonresponders. *J. Clin. Psychiatry* 58, 23–29.
- Tiwari, V., Daoud, E. V., Hatanpaa, K. J., Gao, A., Zhang, S., An, Z., et al. (2020). Glycine by MR spectroscopy is an imaging biomarker of glioma aggressiveness. *Neuro Oncol.* 22, 1018–1029.
- Wang, Y., Jia, Y., Chen, X., Ling, X., Liu, S., Xu, G., et al. (2012). Hippocampal N-acetylaspartate and morning cortisol levels in drug-naive, first-episode patients with major depressive disorder: Effects of treatment. *J. Psychopharmacol.* 26, 1463–1470. doi: 10.1177/0269881112450781
- Williams, J. B. (1988). A structured interview guide for the Hamilton Depression Rating Scale. *Arch. Gen. Psychiatry* 45, 742–747. doi: 10.1001/archpsyc.1988.01800320058007
- Williams, L. M., Debattista, C., Duchemin, A. M., Schatzberg, A. F., and Nemeroff, C. B. (2016). Childhood trauma predicts antidepressant response in adults with major depression: Data from the randomized international study to predict optimized treatment for depression. *Transl. Psychiatry* 6:e799. doi: 10.1038/tp.2016.61
- Woodcock, E. A., Arshad, M., Khatib, D., and Stanley, J. A. (2018). Automated Voxel Placement: A Linux-based Suite of Tools for Accurate and Reliable Single Voxel Coregistration. *J. Neuroimaging Psychiatry Neurol.* 3, 1–8. doi: 10.17756/jnpn.2018-020