



An effective brain connectivity technique to predict repetitive transcranial magnetic stimulation outcome for major depressive disorder patients using EEG signals

Behrouz Nobakhsh¹ · Ahmad Shalbaf¹ · Reza Rostami² · Reza Kazemi³ · Erfan Rezaei¹ · Reza Shalbaf⁴

Received: 11 November 2021 / Accepted: 6 November 2022 / Published online: 29 November 2022
© Australasian College of Physical Scientists and Engineers in Medicine 2022

Abstract

One of the most effective treatments for drug-resistant Major depressive disorder (MDD) patients is repetitive transcranial magnetic stimulation (rTMS). To improve treatment efficacy and reduce health care costs, it is necessary to predict the treatment response. In this study, we intend to predict the rTMS treatment response in MDD patients from electroencephalogram (EEG) signals before starting the treatment using machine learning approaches. Effective brain connectivity of 19-channel EEG data of MDD patients was calculated by the direct directed transfer function (dDTF) method. Then, using three feature selection methods, the best features were selected and patients were classified as responders or non-responders to rTMS treatment by using the support vector machine (SVM). Results on the 34 MDD patients indicated that the Fp2 region in the delta and theta frequency bands has a significant difference between the two groups and can be used as a significant brain biomarker to assess the rTMS treatment response. Also, the highest accuracy (89.6%) using the SVM classifier for the best features of the dDTF method based on the area under the receiver operating characteristic curve (AUC-ROC) criteria was obtained by combining the delta and theta frequency bands. Consequently, the proposed method can accurately detect the rTMS treatment response in MDD patients before starting treatment on the EEG signal to avoid financial and time costs to patients and medical centers.

Keywords EEG · Effective connectivity · Major depressive disorder (MDD) · Repetitive transcranial magnetic stimulation (rTMS).

Introduction

Major depressive disorder (MDD) is a common mental disorder associated with significant personal, social, and economic issues [1]. If a person has at least five symptoms for two weeks or more, such as low mood, decreased enjoyment of formerly pleasurable activities, sleep disturbance, fatigue,

and loss of energy without significant activity, has a change in appetite, pessimism, guilt, and suicidal thoughts, considered as MDD patients and needs favorable treatment [2].

Medication is the first approach to treating MDD patients. However, approximately 30% of patients do not respond to this type of treatment, and these people are considered drug-resistant patients [3]. Non-pharmacological treatments are used for patients who have drug-resistant depression. Electroconvulsive therapy (ECT) is one of the non-pharmacological treatments for MDD patients. However, this method is associated with the risk of anesthesia, memory changes, and affecting cognitive symptoms and it is less commonly used [4]. Repetitive transcranial magnetic stimulation (rTMS) is another effective non-pharmacological treatment for MDD with no side effects compared with ECT and also improves cognitive symptoms [5]. This treatment is based on the principle of electromagnetic induction; a series of magnetic pulses at a specific frequency with an intensity less than the seizure threshold is applied to the cerebral

✉ Ahmad Shalbaf
shalbaf@sbm.ac.ir

¹ Department of Biomedical Engineering and Medical Physics, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

² Department of Psychology, University of Tehran, Tehran, Iran

³ Department of Cognitive Psychology, Institute for Cognitive Science Studies, Tehran, Iran

⁴ Institute for Cognitive Science Studies, Tehran, Iran

cortex for a certain period to regulate the neural activity of the target area [6]. According to the prefrontal cortex asymmetry theory in MDD patients—i.e., left hypoactivity and right hyperactivity of the dorsolateral prefrontal cortex (DLPFC)—rTMS stimulates the hypoactive area and inhibits the hyperactive area [7, 8]. If high-frequency rTMS (usually ≥ 10 Hz) is applied, it increases the brain activity and stimulates the target point, and if low-frequency rTMS (usually ≤ 1 Hz) is used, it reduces the brain activity and inhibits the target point [9]. Studies show that the response rate to rTMS treatment in drug-resistant MDD patients is 50–55% [10, 11]. Given that the total duration of treatment using this method is about 20 sessions, it is necessary to predict the rTMS treatment response. Lack of prediction increases the costs imposed on patients and medical centers, and wastes patients' time in the disease condition [12, 13].

One of the methods to predict the rTMS treatment response in MDD patients is to use demographic and clinical data. A study that used demographic, depressive characteristics, psychiatric and pharmacological history as clinical predictors, showed that patients who are younger and less drug-resistant, respond better to rTMS treatment [14]. In another study, the effects of age, gender, menopausal status, and ovarian hormone levels on the effectiveness of rTMS in drug-resistant MDD patients were investigated. It was observed that there is no significant difference between male and premenopausal female patients in rTMS treatment response, and menopausal status and ovarian steroid levels are the determining factors in the effectiveness of rTMS treatment on women [15]. In other studies that used demographic characteristics, somatic symptoms and cognitive-emotion symptoms for predicting the rTMS treatment response, it was concluded that age is the most critical predictor in all patients. Also, those who respond better to this treatment indicate better cognitive-emotion symptoms than somatic symptoms [16]. The use of demographic and clinical data is not highly discriminant due to differences in these patients' characteristics and brain structure. Therefore, the use of neuroimaging techniques based on EEG in predicting the rTMS treatment response in MDD patients is progressing. EEG is widely used in clinical decisions due to its high temporal resolution, non-invasiveness, cheapness, and availability [17–20].

Various linear and nonlinear features have been proposed to predict the rTMS treatment response in MDD patients using EEG data. In one study, using the absolute power of the alpha frequency band of EEG, patients were classified as responders and non-responders to rTMS treatment [21]. Nonlinear EEG criteria, including the Lempel Ziv complexity and the Lyapunov exponent in the alpha frequency band, were used in another study. The results indicated that the non-responders showed a significant decrease in the Lempel Ziv complexity feature in the first minute than the second

minute. In contrast, the responders showed an increase in the Lempel Ziv complexity feature [22]. Also, other studies have used different approaches such as functional connectivity [23, 24], Katz fractal dimension, and Correlation dimension [25, 26] to predict the treatment response and classify the drug-resistant MDD patients into two groups of responders and non-responders. These methods investigate the complexity of EEG signals but have limitations such as not being suitable for analyzing non-stationary signals and inaccurate estimation of temporal patterns.

With the development of non-invasive neuroimaging techniques, researchers have found that heterogeneous patterns of brain connectivity describe the activity of the brain. A comprehensive map of these patterns leads to better identification of cognitive functions and a wide range of behaviors [27]. Since brain connections' network is involved in psychiatric disorders, single-channel EEG data analysis cannot propose a specific feature of these types of disorders. Therefore, calculating the brain connectivity measures to obtain the intricate brain network patterns to predict the rTMS treatment outcome in drug-resistant MDD patients would greatly help and increase the treatment efficiency. Interactions between different areas of the brain can be analyzed in the form of functional and effective connectivity [28]. Functional connectivity evaluates the statistical dependence of time series, but effective connectivity quantifies the time series' causal and directional impact.

The innovations of the current study are the use of effective brain connectivity methods based on the direct directed transfer function (dDTF) method, which helps to identify the best brain patterns and significant biomarkers between two groups of responder and non-responder MDD patients to rTMS treatment. Also, the other novelty of this study is to find distinctive effective connectivity features from different frequency bands and develop a hierarchical feature selection and classification method to predict the rTMS treatment response in drug-resistant MDD patients by EEG signal before starting the treatment. These will lead to improve the effectiveness of the model and reduce the time and cost of the patients undergoing treatment.

Materials and methods

Participants and clinical assessment

Data were collected from 34 patients (mean age 37.1, standard deviation 13.4, 25 women) who had drug-resistant MDD and were referred to the Atieh clinical neuroscience center for rTMS treatment. One week before rTMS treatment, all patients underwent a baseline clinical evaluation, and an experienced psychiatrist made the diagnosis of MDD by using a structured clinical interview based on the diagnostic

and statistical manual of psychiatric disorders (DSM-IV) [29], and the beck depression inventory (BDI-II) score was recorded for each patient. Then, at the end of the rTMS treatment period (after 20 sessions—3 times a week), the patients were re-evaluated by a psychiatrist, and the BDI-II score was recorded for each patient. If at least 50% of the BDI-II score is reduced, the patient is defined as responding to rTMS treatment. The BDI-II is a 21-item questionnaire that assesses the feelings of a person over the past week. Written consent has been obtained from all participants in this study and has been approved by the ethics committee of Shahid Beheshti University of Medical Sciences. The demographic and clinical characteristics of the patients are summarized in Table 1.

rTMS treatment parameters

In order to choose the best rTMS protocol for the treatment of MDD patients, the influence of the parameters that affect this treatment should be considered. These parameters include the selection of target point, frequency, intensity and number of magnetic pulses and number of treatment sessions. According to the theory of prefrontal cortex asymmetry in MDD patients, i.e. left hyperactivity and right hyperactivity of DLPFC, three types of treatment protocols can be used [7, 8]. One of these protocols stimulates the left DLPFC region with high frequency and increases the activity of the target point and the second protocol inhibits the right DLPFC region with low frequency and reduces the activity. The third protocol is used as a combination of the two previous protocols, namely stimulation of the left DLPFC region and inhibition of the right DLPFC region with high and low frequency, respectively. In this study, according to the existing protocols, we used low frequency to inhibit the right DLPFC region. A meta-analysis indicated that there was no difference between the two protocols of high and low frequency stimulation by rTMS in MDD patients in the left and right DLPFC regions in terms of response rate of MDD patients, respectively, and these two protocols almost have the same effectiveness [30]. Another meta-analysis that examined the acceptability and effectiveness of low frequency rTMS treatment indicated that by increasing the number of magnetic pulses applied in this protocol to more than 1200 pulses, the response rate to rTMS treatment in MDD patients increase [31]. According to the

mentioned items, the selected rTMS protocol is considered a suitable treatment for MDD patients.

rTMS was applied using a Neuro MS device (Neurosoft, Russia) via a 70 mm 8-shaped stimulation coil (air membrane coil) at the Atieh clinical neuroscience center. For obtaining the minimum motor stimulation threshold, the motor area of the abductor pollicis brevis muscle (APB) is stimulated 10 times. If it reacts at least five times, the stimulation intensity is considered the minimum motor threshold. The coil position is 5 cm forward and along the parasagittal line from the optimal stimulation position of the APB muscle. All patients underwent magnetic stimulation for 10 s under a specific protocol with a 120% stimulation threshold in the right DLPFC at a frequency of 1 Hz, and then rested for 2 s and this was repeated. Consequently, for 10 s, 10 magnetic pulses were applied to the patients. In total, this procedure was performed 200 times (200×12 s) with 2000 magnetic pulses in each session and as a result, 40,000 magnetic pulses were applied to patients during a period of 20 sessions.

Pre-treatment EEG acquisition

19-channel EEG data electrodes have been placed according to 10–20 standard (Fp1, Fp2, F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8, O1, and O2). Raw EEG data of patients at resting state with closed eyes before starting rTMS treatment have been recorded at Atieh Clinical Neuroscience center for 5 min at 250 Hz sampling frequency rate with the Mitsar-EEG-201 amplifier and Ag/AgCl electrodes.

EEG preprocessing

Preprocessing of EEG data has been performed using the EEGLAB open-source toolbox [32] to remove environmental and motion artifacts. First, we have used a 1 Hz high-pass filter to remove baseline drift. The scalp potentials' average is independent of the reference location and applied as a re-reference to the signals. The CleanLine [33] open-source plugin has been used in the EEGLAB toolbox to remove the line noise. Independent component analysis (ICA) has been used to remove artifacts such as blinking and head movements. In the end, the EEG data were cleaned visually and we used the reject continuous data by eye due to the existing artifacts, and therefore the length of the data was reduced

Table 1 Demographic and clinical characteristics of participants

	Responder (n = 17)	Non-responder (n = 17)	Total (n = 34)
Gender(Female/Male)	14/3	11/6	25/9
Age	34.8 (\pm 12)	39.3 (\pm 14.7)	37.1 (\pm 13.4)
Pre-treatment BDI-II	31 (\pm 10.3)	31.2 (\pm 10.4)	31.1 (\pm 10.2)
Post-treatment BDI-II	9.4 (\pm 5.5)	23.2 (\pm 11.8)	16.3 (\pm 11.4)

from 300 s. Finally, in order to unify the data, we hold the 150 s of all subjects, continuously.

Feature extraction by effective connectivity

Effective brain connectivity as the extracted feature is a type of directional connectivity that tries to find causal relationships between different brain areas. This brain connectivity presents a new brain network model [34]. The direct directed transfer function (dDTF) is one of the most important effective brain connectivity methods that is widely used. The dDTF method examines the time dependence amount of the two time-series using the model. The dDTF connectivity matrix is asymmetric and determines the directional connectivity and its power. In this study, for feature extraction, the effective brain connectivity between 19 EEG signal channels was calculated using the dDTF method in different frequency bands of the delta, theta, alpha, beta, and gamma. These extracted features are calculated using the SIFT in the EEGLAB toolbox. The EEG signal features are 361 (19×19) per patient. To explain the dDTF method, we first describe the conditions for establishing the Auto-regressive (AR) equation and then explain how to attain this method. Suppose we have an M -channel EEG and $X(t)$ is an M -channel vector at the time t .

$$X(t) = [X_1(t), X_2(t), \dots, X_M(t)]^t \quad (1)$$

We utilize the following equation to state the AR equation. In this equation, $A(k)$ is a matrix of coefficients of order $M \times M$, which indicate the dependence of variables on the delay k . p is the model order, k and $(t - k)$ represent an example of multi-channel data in the past. Also $E(t)$ shows the random noise.

$$X(t) = \sum_{k=1}^p A(k)X(t - k) + E(t) \quad (2)$$

The model order is usually determined by minimizing information criteria such as the Akaike Information Criterion (AIC) [35]. In obtaining the AR equation, two assumptions must be considered; the data must be static (their mean and variance do not change over time), and the equation must be stable. Stability of an equation results in be static, so it is enough to consider only the second condition. In the AR equations, there is a condition for the minimum number of data, which is as follows:

$$M^2p \leq N_r W \quad (3)$$

In this equation, M is the number of variables (number of channels), p is the equation order, N_r is the number of time series, and W is the length of each time window. We need at least M^2p independent samples to calculate the AR equations. Also, with increasing the model degree, more

data is required. In general, there are three different criteria for examining a fitted model: whiteness (if the model is well fitted, the error should be small and uncorrelated), consistency (how much statistical properties (mean, variance, etc.) of the obtained model are consistent with the data), and stability (The obtained model has a limited output for limited input). In neuroscience, the desirable time series are often collected simultaneously and for each of these time series, an AR equation can be defined. Granger causality indicates that if the prediction of the future values of a time series in the presence of a second time series is improved, then the second time series is the cause of the first time series; To obtain the Granger causality in the frequency domain, we write Eq. 2 as follows:

$$\begin{aligned} A(f) &= \sum_{k=0}^p A(k)e^{-2\pi f k} \\ X(f) &= A(f)^{-1}E(f) = H(f)E(f) \end{aligned} \quad (4)$$

In the above equation, $H(f)$ is the system transfer function and calculates the connectivity matrix in the frequency domain. The directed transfer function (DTF) and the partial coherence function (pCoh) can be obtained by using the system transfer function. Finally, the dDTF method is obtained from multiplication (frequency domain) of pCoh in the DTF [36].

Feature selection

Reducing the features' dimension can improve the interpretability and efficiency of the model by using various feature selection methods. Also, we used feature selection methods to decrease the calculations complexity and classification parameters, and increase the computational rate [37]. In the following, we will explain three feature selection methods that have been used in this study. The first method is a forward feature selection algorithm based on the area under the receiver operating characteristic curve (AUC-ROC). AUC-ROC is used for the evaluation of the performance of binary classification algorithms based on given input features [38]. AUC tells how much the input features are capable of distinguishing between two classes and seeing the importance of given input features. The larger AUC-ROC for each feature indicates the higher relationship of that feature to the class label. The AUC-ROC value of features varies from 0 to 1 and a high AUC-ROC value (equal or close to 1) means the ability of the feature to separate classes. Therefore, to select the best features, after calculating the AUC-ROC values of each feature, these features are arranged in descending order of AUC-ROC values [39]. After selecting the best features, the forward feature selection algorithm uses the learning method to evaluate the usefulness of each subset of these features and aims to find a subset of features with the least amount of classification error. First, all features are given

to the classifier one by one, and the best feature is selected. Then the combination of the first selected feature and the other remaining features is given to the classifier, and the best double feature combination is determined. The framework can be continued to identify informative feature groups of various sizes. Another feature selection method that was used in this study is the Relief-F. This method is a supervised method that evaluates the quality of the features. In this method, a sample is selected randomly from the samples in the data set at each step. If the same feature in the selected sample differs from the similar feature in the neighboring sample of the same class, this feature's score reduces. On the other hand, if the same feature in the chosen sample differs from the similar feature in the opposite class's neighboring sample, the score of this feature increases [37]. Finally, after calculating the score values of each feature, these features are arranged in descending order. The third feature selection method is the mRMR. This method evaluates the features based on the maximum relevance and minimum redundancy. For example, features have a maximum relevance that has the maximum amount of mutual information between features and class labels. On the other hand, features with minimum redundancy are identified based on the principle that if two features are interdependent and one of them is removed, the classification performance will not change much [40]. Accordingly, after calculating the maximum relevance and minimum redundancy of each feature, they are arranged in descending order to select the best features.

Classification

Classification methods have been successfully used to analyze complex patterns in neuroimaging data. In this study, two classification methods of linear discriminant analysis (LDA) and support vector machine (SVM) have been used. LDA as a supervised method is used to find a linear hyperplane that best separates data from two classes [41] and SVM obtains a hyperplane with a maximum margin to separate the two classes. In the SVM method, if the data is not linearly separable, the data will be mapped to a larger space to be able to separate them in the new space. First, two boundary planes are created parallel to the classification plane. These two planes are so far apart that they collide with the data, in which case they are called support vectors. So, the best separator is created with the maximum distance from all data [42]. Different kinds of used kernel function for SVM classifier in this study are linear, quadratic, cubic, and Gaussian kernels such as radial basic function (RBF). The correct choice of kernels and parameters greatly affects the performance and final result. We have used the RBF kernel, which is the most common kernel based on the Euclidean Distance. Also, the RBF kernel has good performance due to the consideration of data distribution and finally has less

complexity and less time than other kernel function such as polynomial kernel [42].

Statistical analysis

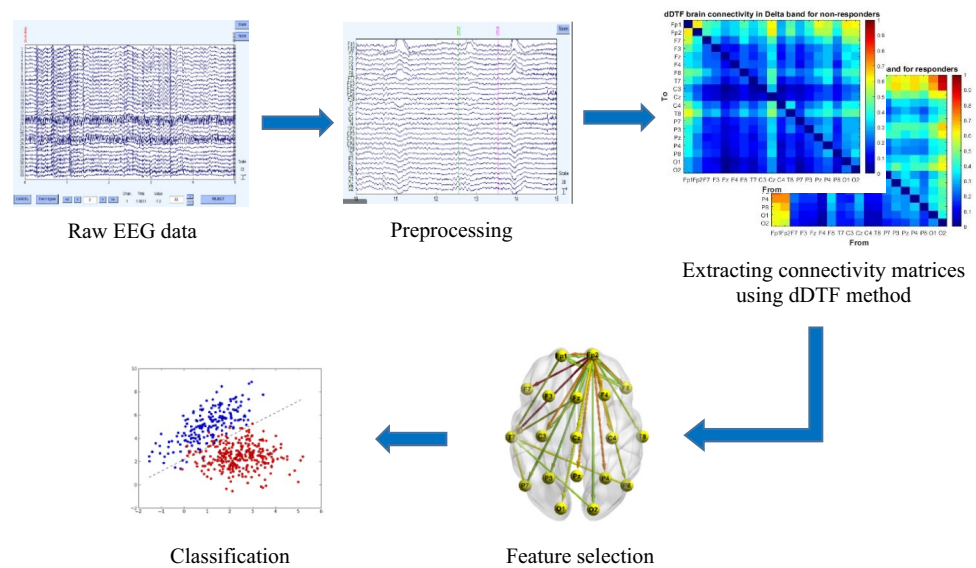
Due to the limited data set, k-fold cross-validation will be used. This method divides the data set into k sections with the same number of samples. The optimal value of k will be selected through trial and error by maximizing classification performance and minimizing error. In each trial, the classification structure is constructed with the k–1 section of the data (for training and validation), and evaluated with the remaining data as test data. This process is repeated k times so that each data is used exactly once as test data. Performance evaluation will be reported with the average k test results. In this study, the 10-fold cross-validation method has been used to evaluate the performance of the classifications. We evaluate our algorithms using four criteria: accuracy, sensitivity, specificity, and F1-score.

In this study, to evaluate the best features, the P-value of the Wilcoxon rank-sum test is used. The Wilcoxon rank-sum test is a non-parametric test for two groups whose samples are independent of each other [43]. Another statistical method used to evaluate the power of feature separation is the use of the AUC-ROC values, which has also been used in this study.

Overview of the proposed method

Figure 1 shows the block diagram of the proposed method. First, the raw data obtained from the EEG were preprocessed with the EEGLAB open-source toolbox. Preprocessing steps include frequency filtering and line noise cancellation, artifacts removing, ICA, and time correction. Then, the effective brain connectivity between 19 EEG signal channels as the extracted feature was calculated using the dDTF method in different frequency bands of the delta, theta, alpha, beta, and gamma. The EEG signal features are 361 (19 × 19) per patient. For extraction of dDTF features, whiteness, consistency, and stability of EEG signals are considered. In this study, the length of the signal window was considered 10 s and considering 150 s for each patient, 15 connectivity 19 × 19 matrix as the extracted feature has been calculated from each subject. These extracted features are calculated using the SIFT in the EEGLAB toolbox [44]. Then we seek to identify effective biomarkers that can be used to obtain rTMS treatment response in MDD patients. In the following, the best features were selected using forward feature selection algorithm based on AUC-ROC, Relief-F, and mRMR feature selection methods. Finally, the selected features were classified by SVM and LDA. All machine learning calculations are performed in MATLAB software.

Fig. 1 Schematic diagram of the proposed method. First, the raw data of the collected EEG signals are pre-processed, and then, using the dDTF method, the brain connectivity matrix is calculated. In the following, by using the feature selection, the best features are obtained, and finally, the classification of the two classes is done



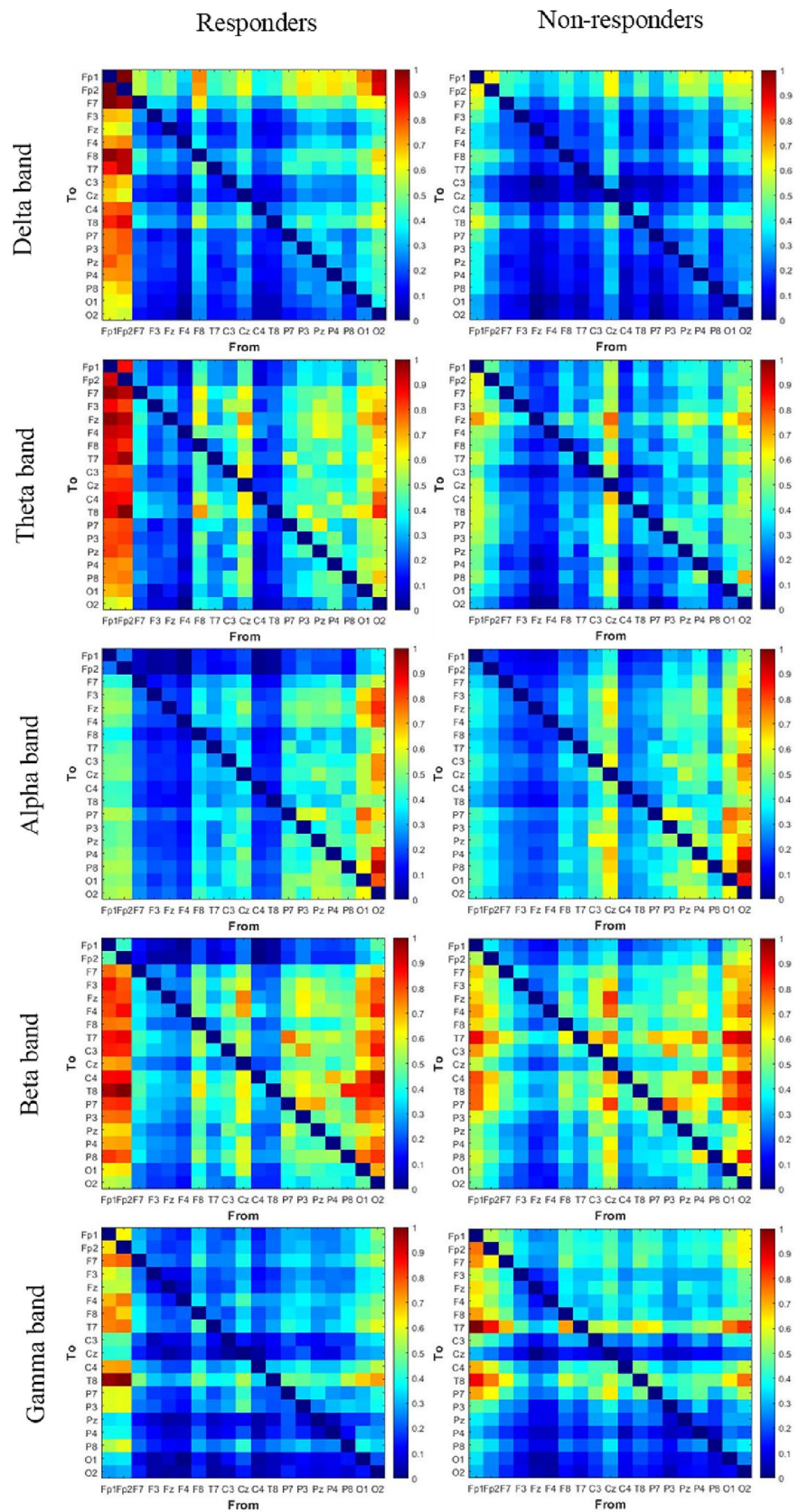
Results

After preprocessing the 19-channel EEG signal, effective brain connectivity has been calculated between different brain areas by the dDTF method in delta, theta, alpha, beta, and gamma frequency bands for all responder and non-responder MDD patients to rTMS treatment, separately. As described in the previous section, the dDTF method examines the dependency degree of the brain signals by using an AR model that determines the power and direction between different brain regions. In this study, the length of the signal window was considered 10 s (The EEG signal length of each patient was 150 s and for each patient, 15 connectivity matrix has been calculated) and two criteria of stability and consistency with a model order of 12 were examined and approved. Effective brain connectivity between 19 channels of each patient was calculated using the dDTF method in different frequency bands as 361 (19×19) features. In Fig. 2, the mean values of the normalized effective brain connectivity matrices calculated by the dDTF method from EEG signals of responder and non-responder MDD patients to rTMS treatment in different frequency bands (delta, theta, alpha, beta, and gamma) have been shown. In these figures, the rows and columns represent 19 EEG signal electrodes, and the color values of the matrix indicate the amount of brain connectivity. Also, to better display these features, brain connectivity using the BrainNet toolbox [45] for responder and non-responder MDD patients is shown in Fig. 3 as graph representation. In this figure, the nodes represent the brain regions or EEG signal electrodes, and normalized mean values of the brain connectivity matrices determined by the dDTF method are defined as edges. We used the representation of the graph to visualize effective brain connectivity and find the best areas of difference between the two groups.

Statistical methods have been used to quantify the brain activation differences between the two groups of responders and non-responders. Thus, for each of the 361 features extracted between the two groups of responders (17 patients) and non-responders (17 patients), the Wilcoxon rank-sum test and the AUC-ROC value have been used. Statistically, the features that lead to $P\text{-value} \leq 0.001$ and have a higher AUC-ROC value can distinguish between the two groups of responders and non-responders and indicate the difference in activity between the two groups. Figure 4 shows 30 features of the best features selected by the AUC-ROC criterion in each frequency band. In this figure, the presence of the edges indicates the membership of the 30 best features in each frequency band, and the color of the edges indicates the AUC-ROC values. In the following, 30 best features of effective brain connectivity between different regions in all frequency bands between the two groups of responder and non-responder to rTMS treatment are ranked by the AUC-ROC criterion and shown in Table 2.

From another perspective, the advanced machine learning techniques' power has been examined to detect the rTMS treatment response for drug-resistant MDD patients from pre-treatment EEG signals. The machine learning process consists of three main steps: feature extraction, feature dimension reduction or feature selection, and classification algorithm. At first, the effective brain connectivity between different brain channels using the dDTF method in the delta and theta frequency bands (other frequency bands were not calculated due to poor results in the previous section) are calculated and 361 different features from each patient are extracted. In the following, the best features combination that can be distinguished between the two groups has been estimated by using three mentioned feature selection methods named mRMR, Relief-F, and forward feature selection

Fig. 2 The normalized values of the brain connectivity matrix calculated by the dDTF method from EEG signals in responder and non-responder MDD patients to rTMS treatment in the delta, theta, alpha, beta, and gamma frequency bands. The color values of the matrix indicate the amount of brain connectivity



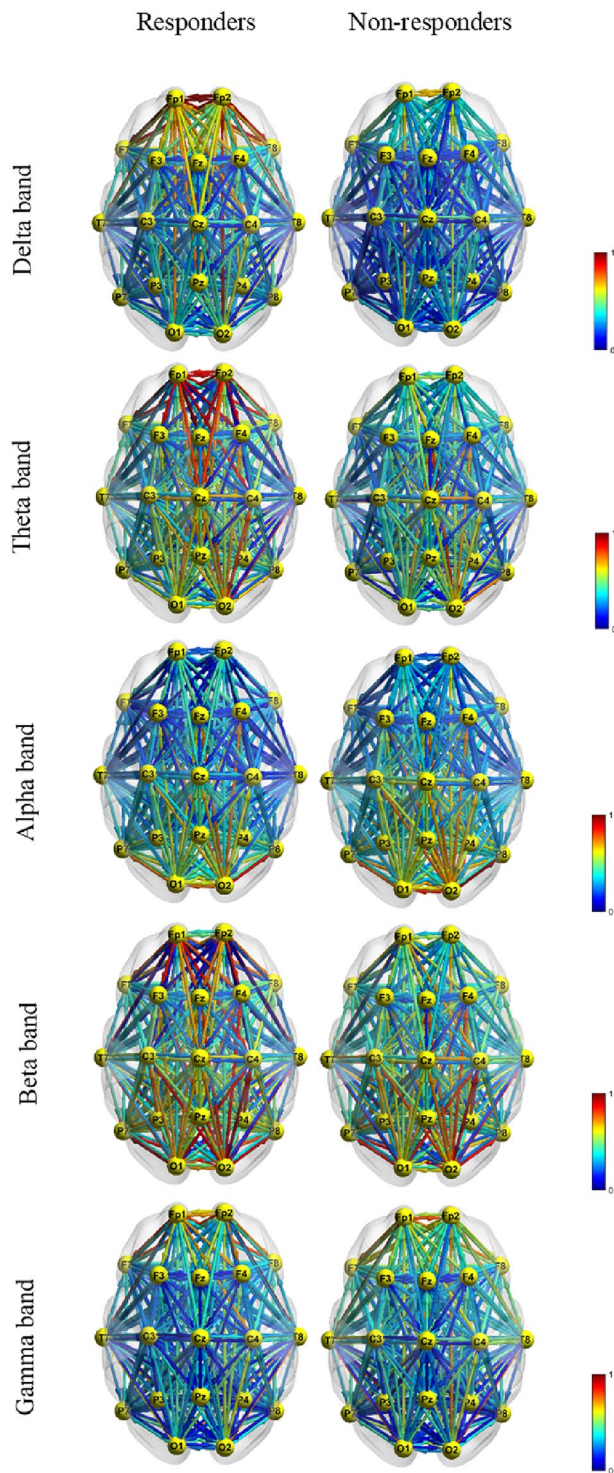


Fig. 3 The values of brain connectivity in responder and non-responder MDD patients as the graph representation in each frequency band. The nodes represent the brain regions or EEG signal electrodes, and the normalized mean values of the brain connectivity matrices are defined as edges

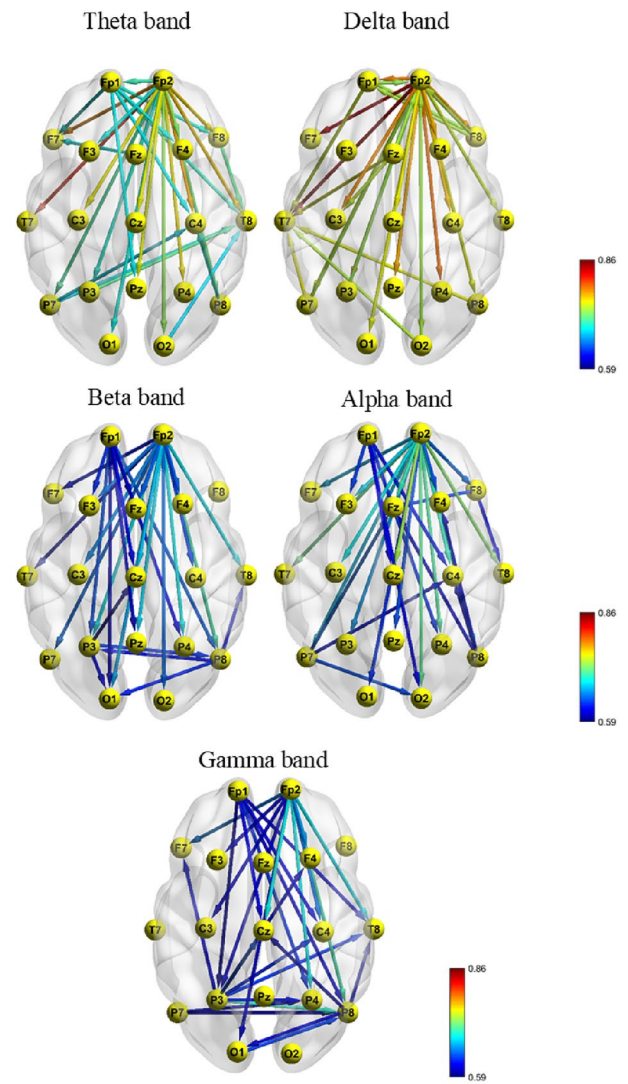


Fig. 4 30 features of the best features selected by the AUC-ROC criterion in each frequency band between two groups of responder and non-responder to rTMS treatment. In this figure, the presence of the edges indicates the membership of the 30 best features in each frequency band, and the color of the edges indicates the AUC-ROC values. The prefrontal regions (Fp1 and Fp2), especially Fp2 in the delta and theta frequency bands, have the highest AUC-ROC values

algorithm based on AUC-ROC. In Tables 2, 3, 4 and 30 best features of effective brain connectivity between the two groups of responder and non-responder to rTMS treatment by the AUC-ROC, Relief-F and mRMR are ranked and shown, respectively. In the end, we have classified the responder and non-responder MDD patients using the selected features and LDA and SVM classifiers. The ‘RBF’ kernel has been considered as kernel function in the SVM classifier. In this study, the 10-fold cross-validation method has been used to evaluate the classifier performance.

In Tables 5 and 6, the classification results of responders and non-responders MDD patients to rTMS treatment

Table 2 Rank of the 30 best features of effective brain connectivity between different regions in all frequency bands between the responder and non-responder MDD patients to rTMS treatment by the AUC-ROC criterion. The prefrontal area (Fp1 and Fp2), especially Fp2 has the highest AUC-ROC values than other areas of the brain. Delta and theta frequency bands also have higher AUC-ROC values than other frequency bands

	EEG electrode names	Frequency band	AUC-values	P-values
1	Fp2 ⇒ T7	Delta	0.850	>0.001
2	Fp2 ⇒ F7	Delta	0.829	>0.001
3	Fp2 ⇒ T7	Theta	0.816	>0.001
4	Fp2 ⇒ P4	Delta	0.797	>0.001
5	Fp2 ⇒ Fp1	Delta	0.790	>0.001
6	Fp2 ⇒ C3	Delta	0.787	>0.001
7	Fp2 ⇒ F7	Theta	0.787	>0.001
8	Fp2 ⇒ Pz	Delta	0.780	>0.001
9	Fp2 ⇒ P3	Delta	0.779	>0.001
10	Fp2 ⇒ F8	Delta	0.778	>0.001
11	Fp2 ⇒ C4	Theta	0.778	>0.001
12	Fp2 ⇒ F4	Delta	0.777	>0.001
13	Fp2 ⇒ C4	Delta	0.775	>0.001
14	Fp2 ⇒ T8	Theta	0.769	>0.001
15	Fp2 ⇒ F8	Theta	0.766	>0.001
16	Fp2 ⇒ P4	Theta	0.761	>0.001
17	Fp2 ⇒ Cz	Delta	0.758	>0.001
18	Fp2 ⇒ P7	Delta	0.758	>0.001
19	Fp2 ⇒ Cz	Theta	0.756	>0.001
20	Fp2 ⇒ C3	Theta	0.756	>0.001
21	Fz ⇒ T7	Delta	0.755	>0.001
22	P7 ⇒ T7	Delta	0.755	>0.001
23	P8 ⇒ T7	Delta	0.753	>0.001
24	Fp2 ⇒ O1	Delta	0.753	>0.001
25	Fp2 ⇒ Pz	Theta	0.752	>0.001
26	Fp2 ⇒ P8	Delta	0.751	>0.001
27	Fp1 ⇒ T7	Delta	0.749	>0.001
28	Fp2 ⇒ T8	Delta	0.746	>0.001
29	Fp2 ⇒ O2	Delta	0.746	>0.001
30	Fz ⇒ C3	Delta	0.743	>0.001

are shown separately by using all 30 best features, feature selection of forward selection based on AUC-ROC ranked, mRMR, and Relief-F and two classification methods in the delta and theta frequency bands in terms of the accuracy, sensitivity, specificity, and F1-score, respectively. The delta and theta frequency bands features are combined to improve the classification performance. In Table 7, the results of combined features in the delta and theta frequency bands have been shown. In this case, the highest accuracy, sensitivity, and specificity values for the forward selection based on AUC-ROC ranked and SVM classifier equal to 89.6%, 84.5%, and 94.7% have reached, respectively. Figure 5 shows

Table 3 Rank of the 30 best features of effective brain connectivity between different regions in all frequency bands between the responder and non-responder MDD patients to rTMS treatment by the Relief-F method

	EEG electrode names	Frequency band
1	P7 ⇒ P8	Gamma
2	P8 ⇒ Fp2	Gamma
3	Cz ⇒ O2	Gamma
4	P4 ⇒ Fp2	Beta
5	C4 ⇒ P8	Gamma
6	P4 ⇒ Fp2	Gamma
7	C4 ⇒ Fp2	Beta
8	Fp2 ⇒ T7	Delta
9	P3 ⇒ Fp2	Gamma
10	F3 ⇒ Fp2	Beta
11	T8 ⇒ Fp2	Gamma
12	Fp2 ⇒ P8	Gamma
13	P7 ⇒ T7	Delta
14	O2 ⇒ P7	Gamma
15	T7 ⇒ Fp2	Gamma
16	Pz ⇒ Fp2	Gamma
17	C3 ⇒ O1	Alpha
18	P8 ⇒ F7	Gamma
19	Cz ⇒ F4	Beta
20	C4 ⇒ T7	Beta
21	Cz ⇒ P8	Gamma
22	C4 ⇒ Fp2	Alpha
23	F7 ⇒ Fp2	Beta
24	C3 ⇒ P8	Theta
25	C3 ⇒ F7	Beta
26	P3 ⇒ O1	Gamma
27	C3 ⇒ O1	Beta
28	O2 ⇒ O1	Gamma
29	T8 ⇒ O2	Gamma
30	Cz ⇒ T8	Gamma

the accuracy diagram for the SVM classifier with forward selection based on AUC-ROC ranked, according to the number of features in the delta, theta, and combination of the two frequency bands. With increasing the number of features, the classification accuracy reaches its maximum value and then decreases. Also, when the delta and theta bands are combined, the classification accuracy is increased. The finally selected features for the best results using forward selection based on AUC-ROC ranked and SVM classifier are: P7 ⇒ Fp2 (Delta), Fp2 ⇒ P3 (Delta), Fp1 ⇒ T7 (Delta), O2 ⇒ T8 (Theta), Fp2 ⇒ F7 (Theta), Fp2 ⇒ T7 (Delta). In Table 8, in accordance with Fig. 5 in combined delta and theta frequency bands, the best two, three, four, five and six selected features with their classification accuracy by forward feature selection algorithm and SVM classification are shown.

Table 4 Rank of the 30 best features of effective brain connectivity between different regions in all frequency bands between the responder and non-responder MDD patients to rTMS treatment by the mRMR method

	EEG electrode names	Frequency band
1	Fp2 ⇒ T7	Delta
2	Fp2 ⇒ T8	Theta
3	Fz ⇒ F8	Delta
4	C3 ⇒ Pz	Gamma
5	Fz ⇒ C4	Delta
6	T8 ⇒ C4	Alpha
7	O1 ⇒ P8	Gamma
8	P8 ⇒ C4	Theta
9	P7 ⇒ Fz	Alpha
10	F8 ⇒ T8	Theta
11	Fp2 ⇒ P4	Alpha
12	P3 ⇒ Fp2	Delta
13	Fp2 ⇒ C4	Delta
14	Cz ⇒ Fp2	Alpha
15	C3 ⇒ Fp2	Gamma
16	Fp2 ⇒ T8	Alpha
17	P3 ⇒ T8	Theta
18	Fp2 ⇒ F8	Theta
19	T8 ⇒ Fp2	Alpha
20	F8 ⇒ Fp2	Delta
21	Fz ⇒ Fp2	Delta
22	Fp2 ⇒ F7	Delta
23	T8 ⇒ P7	Gamma
24	Fp2 ⇒ T7	Theta
25	F8 ⇒ T7	Delta
26	Fz ⇒ Cz	Delta
27	Fp2 ⇒ Pz	Theta
28	Fp2 ⇒ T8	Beta
29	Fz ⇒ F3	Theta
30	Fp2 ⇒ O2	Theta

Discussion

In this study, the effective brain connectivity based on the dDTF method was calculated between two groups of responder and non-responder MDD patients to rTMS treatment. Results indicated that the prefrontal regions, specifically the Fp2 region in the two delta and theta frequency bands have significant differences between the two groups. Also, as results indicate, it can be used as a remarkable brain pattern or valuable brain biomarker to assess the treatment response in MDD patients by EEG signal before starting the treatment and avoid financial and time costs to patients and medical centers. Moreover, the machine learning performance through feature selection methods and classification algorithms has been estimated. The results have shown that the SVM classifier accuracy by combining the delta and theta frequency bands using forward feature selection algorithm based on the AUC-ROC has reached the highest value of 89.6%. In Table 9, a list of existing works to prediction of rTMS treatment response in MDD patients are presented. As it is observed, the accuracy achieved in this study by applying the effective connectivity method with dDTF as the features and selected biomarkers through the AUC-ROC feature selection and SVM classifier is higher than those studies. It proves the preference of the proposed method. Evaluation of the active areas determined that the prefrontal regions (especially the Fp2 region) played the most critical role in selecting the best features to classify the MDD patients in detecting the rTMS treatment response. Finally, our proposed method, compared to deep learning methods, has less processing time. Feature selection and machine learning time is on average less than 2 min, which is less than deep learning methods that take at least a few hours to process.

In the prefrontal cortex, the activity of responder and non-responder MDD patients is different (Figs. 3, 4 and 5). So that responders show more activity in these areas. It happens

Table 5 Results of classification of responder and non-responder MDD patients to rTMS treatment using effective brain connectivity method and all 30 best features, feature selection of forward selection

Classifier	Feature selection	Classification performance			
		Accuracy	Sensitivity	Specificity	F-Measure
SVM	AUC ranked and forward selection	87.8% (± 0.8)	76% (± 1.2)	99.6% (± 0.8)	86% (± 0.2)
	mRMR	81.4% (± 0.7)	73.1% (± 1.8)	89.8% (± 0.8)	80% (± 0.3)
	Relief-F	80.2% (± 1.3)	90.3% (± 1.4)	70.1% (± 2.2)	82% (± 0.2)
	All 30 best features	70.1% (± 1.4)	67.1% (± 1.4)	72% (± 2.3)	69.3% (± 0.5)
LDA	AUC ranked and forward selection	77.7% (± 0.6)	78.4% (± 0)	77% (± 1.3)	77.9% (± 0.1)
	mRMR	78.7% (± 0.9)	77.8% (± 0.9)	79.6% (± 1.6)	78.5% (± 0.1)
	Relief-F	75% (± 1.8)	72.1% (± 1.8)	78% (± 2.8)	74.3% (± 0.6)
	All 30 best features	66.3% (± 1.6)	63.2% (± 1.7)	68% (± 2.5)	65.2% (± 0.6)

based on AUC-ROC ranked, mRMR, and Relief-F by the LDA and SVM classification methods in delta frequency band

Table 6 Results of classification of responder and non-responder MDD patients to rTMS treatment using effective brain connectivity method and all 30 best features, feature selection of forward selection based on AUC-ROC ranked, mRMR, and Relief-F by the LDA and SVM classification methods in theta frequency band

Classifier	Feature selection	Classification performance			
		Accuracy	Sensitivity	Specificity	F-Measure
SVM	AUC ranked and forward selection	86.3% (± 1.4)	83.3% (± 2.1)	89.4% (± 1.3)	85.4% (± 0.4)
	mRMR	78.9% (± 1)	75.2% (± 1.8)	82.5% (± 1.1)	78% (± 0.2)
	Relief-F	81.1% (± 2.1)	80% (± 3)	82.3% (± 3.8)	81.1% (± 0.5)
	All 30 best features	59.2% (± 1.1)	56.4% (± 1.2)	61% (± 2.3)	58.4% (± 0.5)
LDA	AUC ranked and forward selection	83.8% (± 0.6)	84.3% (± 0.9)	83.3% (± 1.3)	83.9% (± 0.2)
	mRMR	77.9% (± 0.6)	81.9% (± 1.2)	73.9% (± 1.3)	79.1% (± 0.4)
	Relief-F	73.8% (± 2)	68.4% (± 2.6)	79.2% (± 3.4)	71.8% (± 1)
	All 30 best features	57.1% (± 1.2)	55.3% (± 1.4)	59.2% (± 2.1)	56.3% (± 0.5)

Table 7 Results of classification of responder and non-responder MDD patients to rTMS treatment using effective brain connectivity method and all 30 best features, feature selection of forward selection based on AUC-ROC ranked, mRMR, and Relief-F by the LDA and SVM classification methods in a combination of delta and theta frequency bands

Classifier	Feature selection	Classification performance			
		Accuracy	Sensitivity	Specificity	F-Measure
SVM	AUC ranked and forward selection	89.6% (± 1)	84.5% (± 1.9)	94.7% (± 0.9)	88.5% (± 0.2)
	mRMR	82.1% (± 1.7)	79.2% (± 2.8)	85% (± 1)	80.9% (± 1.4)
	Relief-F	76.4% (± 1.3)	70.7% (± 1.1)	82.1% (± 1.9)	74.9% (± 0.3)
	All 30 best features	75.2% (± 1.5)	72.1% (± 1.5)	78% (± 2.5)	74.3% (± 0.5)
LDA	AUC ranked and forward selection	86.9% (± 0.6)	88.2% (± 0)	85.6% (± 1.3)	87.1% (± 0.1)
	mRMR	82.5% (± 0.6)	78.2% (± 1.1)	86.8% (± 0.9)	81.5% (± 0.2)
	Relief-F	76.2% (± 1.7)	71.5% (± 1)	80.9% (± 3.2)	74.9% (± 0.4)
	All 30 best features	72.2% (± 1.5)	67.5% (± 0.9)	76.9% (± 3.1)	70.9% (± 0.3)

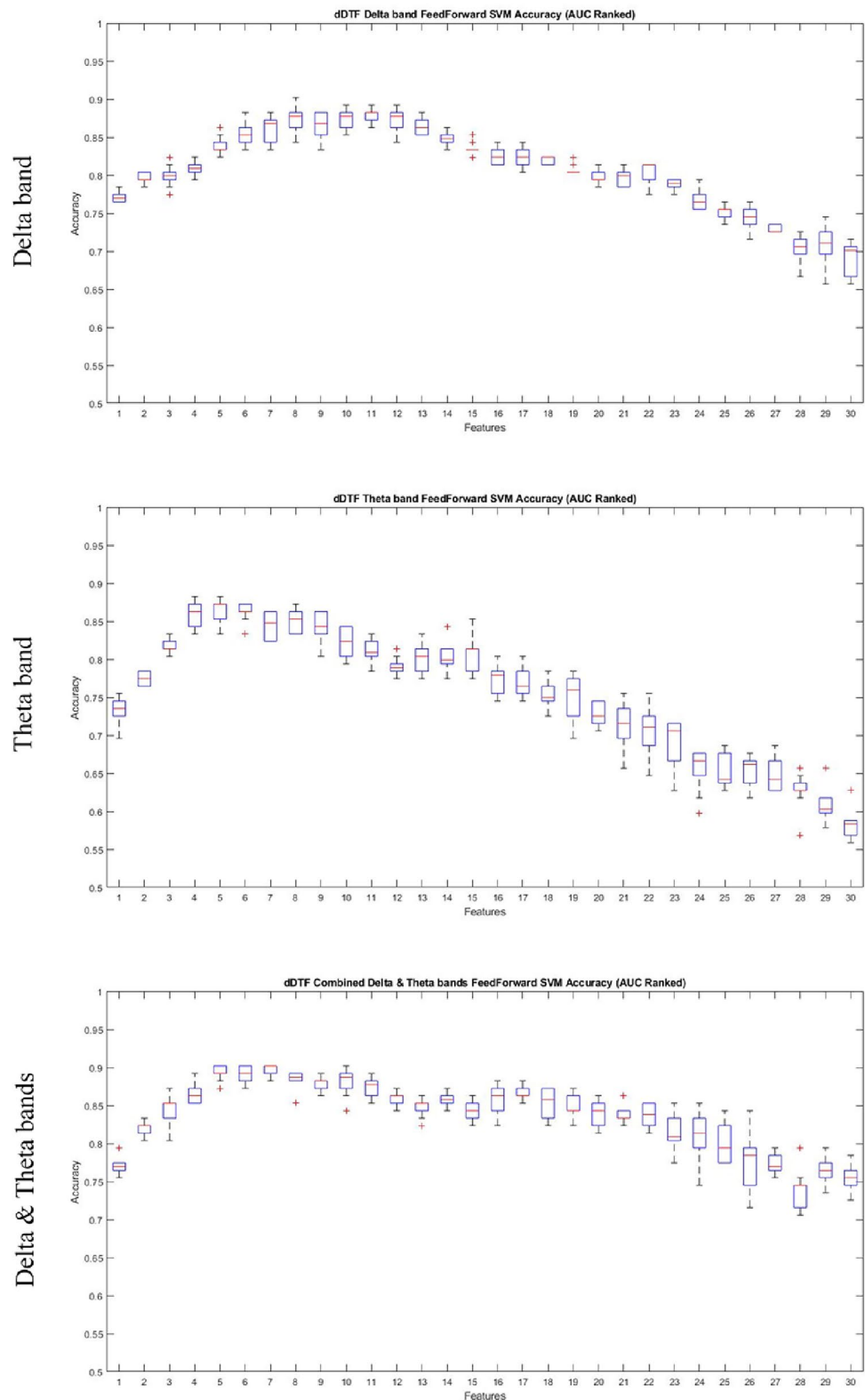
in the prefrontal areas (Fp1 and Fp2), especially in the delta and theta frequency bands. As shown in Table 2 and these Figures, the prefrontal regions (Fp1 and Fp2), especially Fp2, have higher AUC-ROC values than other areas of the brain. Delta and theta frequency bands also have higher AUC-ROC values than other frequency bands. Therefore, after calculating and examining the effective brain connectivity matrices according to the dDTF method between the two groups of responder and non-responder to rTMS treatment, the Fp2 region in the delta and theta frequency bands has the highest activity difference between these two groups and has the highest AUC-ROC values in comparison to other brain regions and other frequency bands. This result can be a significant brain pattern or brain biomarker to classify the responders and non-responders before starting the rTMS treatment. In other words, the Fp2 region in the delta and theta frequency bands play the most critical role in predicting the rTMS treatment response in drug-resistant MDD patients.

The dDTF method provides one of the best effective brain connectivity measures. Unlike the Granger-Geweke

causality (based on Granger causality), only direct directed connectivities calculated, and the indirect and fake connectivities excluded from the connectivity matrix. The dDTF method is one of the multivariate methods based on multi-channel AR models that can identify the causal relationships between signals and determine the direct flow of signals' activations. Frequency dependence is one of the important features of the dDTF method because different rhythms of EEG signals have different roles in information flow processing. Since the DTF method is based on the phase difference between the channels, so it is insensitive to the volume conductance effect and robust to the noise. Due to the mentioned advantages of the dDTF method, it can be concluded that the dDTF method provides the best effective brain connectivity.

We have used three known feature selection algorithms named, mRMR, Relief-F, and forward feature selection algorithm based on AUC-ROC for selecting the best features (Tables 2, 3 and 4). The results displayed that the forward feature selection algorithm based on the AUC-ROC method which uses a classifier during the feature selection phase and also uses its predictive performance to evaluate

Fig. 5 Accuracy of SVM classification as a function of the number of features using forward selection based on AUC-ROC in the delta (top), theta (middle), and combination of delta and theta (bottom) bands



the usefulness of an input feature yielded better classification results (Tables 5, 6 and 7). On the other hand, Relieff and mRMR, only rely on the importance of the general features and consider the relevance of features with dependent

classes using statistical measures. Therefore, the results revealed that the forward feature selection algorithm based on the AUC-ROC selection led to better performance results compared with Relieff and mRMR methods, which can be

Table 8 The selected features by forward feature selection⇒ algorithm and SVM classification in combined delta and theta frequency bands between the responder and non-responder MDD patients to rTMS treatment with their classification accuracy

	EEG electrode names	Classification accuracy
1	Fp2⇒ T7 (Delta)	77.15 ± 1.03
2	Fp2⇒ F7 (Theta), Fp2⇒ T7 (Delta)	81.17 ± 1.20
3	O2⇒ T8 (Theta), Fp2⇒ F7 (Theta), Fp2⇒ T7 (Delta)	84.80 ± 1.40
4	Fp1⇒ T7 (Delta), O2⇒ T8 (Theta), Fp2⇒ F7 (Theta), Fp2⇒ T7 (Delta)	87.05 ± 2.34
5	Fp2⇒ P3 (Delta), Fp1⇒ T7 (Delta), O2⇒ T8 (Theta), Fp2⇒ F7 (Theta), Fp2⇒ T7 (Delta)	89.31 ± 0.72
6	P7⇒ Fp2 (Delta), Fp2⇒ P3 (Delta), Fp1⇒ T7 (Delta), O2⇒ T8 (Theta), Fp2⇒ F7 (Theta), Fp2⇒ T7 (Delta)	89.60 ± 1.05

Table 9 A list of existing work to prediction of rTMS treatment response in MDD patients

Authors	Year	Feature	Classifier	Accuracy
Shalhaf, et al. [12]	2018	Non-linear entropy	LR	80%
He R., et al. [46]	2019	Combined EEG and mood measures	Linear SVM	86%
Hasanzadeh F., et al. [25]	2019	Katz fractal dimension, and Correlation dimension	k-nearest neighbors	91.3%
Ebrahimzadeh E., et al. [47]	2021	Absolute powers, band powers, and theta and beta band entropies	SVM	82.43%
Proposed method	2021	Direct directed transfer function	SVM	89.6%

seen from the results of the classifiers they have not selected appropriate features.

Results of this study using the extracted features from the dDTF method showed that the delta and theta frequency bands have higher efficiency compared with other frequency bands for discrimination of the responder and non-responder MDD patients to rTMS treatment. According to Table 2, based on the value of AUC-ROC, 30 of the best features are plotted in all frequency bands and indicate that the delta and theta bands have a higher separability than other frequency bands. Also, according to Figs. 3 and 4, which indicate normalized connectivity matrices of all five frequency bands of responders and non-responders, the delta and theta bands have significant values compared to the other frequency bands, especially for the Fp1 and Fp2 channels. Finally, these results are consistent with previous studies [39, 48, 49], based on the higher performance of features in the delta and theta bands than in other frequency bands for the two groups classification. Therefore, in machine learning calculation to increase accuracy and speed of classification, only the delta and theta frequency bands are considered and other frequency bands were not calculated due to poor results in the previous section. From a neurobiological point of view, the discrimination of the delta and theta bands in the frontal cortex can be explained using theta current density, localized by LORETA in the rostral anterior cingulate cortex (rACC) in MDD patients [50, 51]. This region is related to the response to different types of antidepressants during depression. The rACC is involved in self-focused processing

and is known as the main hub in the brain's default network. Besides, increasing the resting-state activity in rACC is associated with rumination, remembering, and planning [52]. Rumination is a mechanism for responding to distress and consists of two components: reflective pondering and brooding. Increased rACC activity may lead to treatment response due to adaptive self-referential functions such as mindfulness through reflective pondering and less brooding or less self-focused. Cognitive problem solving is accomplished through reflective pondering. While brooding is like self-focused processing, which is ultimately destructive because it worsens depressive symptoms. Also, the discriminant power of the functional connectivity of the rACC has been demonstrated in MRI data in depressed patients [53].

In addition to the classification accuracy of responders and non-responders to rTMS treatment, sensitivity and specificity are considerable. The more sensitivity, would be deprived the fewer patients of treatment, and on the other hand, the more specificity, would be prevented from wasting time, money and stimulation of the non-responder patients. In general, it can be asserted that both specificity and sensitivity parameters have approximately equal importance.

For future work, it is suggested to calculate the effective brain connectivity in brain source localization of EEG signals with more channels, and then discuss the features and classification methods. In this study, effective brain connectivity features were calculated from 19 channels of EEG signals, while another way to calculate effective brain connectivity is through the brain source localization of the EEG

signals. Also, it is suggested that other feature extraction methods, including functional connectivity and also neural network-based algorithms such as deep learning, utilize to predict the rTMS treatment response in drug-resistant MDD patients.

Conclusion

Results of the effective brain connectivity based on the dDTF method indicated that the prefrontal region and specifically the Fp2 region in the delta and theta frequency bands could be used as a valuable brain biomarker to assess the treatment response in drug-resistant MDD patients by EEG signal before starting the treatment. Also, the results have shown that the accuracy of the SVM classifier in the combination of the delta and theta frequency bands using the forward feature selection algorithm based on AUC-ROC has reached the highest value of 89.6%.

Acknowledgements This research is financially supported by "Research Department of School of Medicine Shahid Beheshti University of Medical Sciences" (Grant No 13731).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Organization WH (2017) Depression. Fact sheet No. 369. Accessed April 2017
2. Association AP (2000) Diagnostic criteria from dsM-iv-tr. American Psychiatric Pub
3. Al-Harbi KS (2012) Treatment-resistant depression: therapeutic trends, challenges, and future directions. *Patient Prefer Adherence* 6:369
4. Fink M (2001) Convulsive therapy: a review of the first 55 years. *J Affect Disord* 63(1–3):1–15
5. O'Reardon JP, Solvason HB, Janicak PG, Sampson S, Isenberg KE, Nahas Z et al (2007) Efficacy and safety of transcranial magnetic stimulation in the acute treatment of major depression: a multisite randomized controlled trial. *Biol Psychiatry* 62(11):1208–1216
6. Beuzon G, Timour Q, Saoud M (2017) Predictors of response to repetitive transcranial magnetic stimulation (rTMS) in the treatment of major depressive disorder. *Encephale* 43(1):3–9
7. Brunoni AR, Chaimani A, Moffa AH, Razza LB, Gattaz WF, Daskalakis ZJ et al (2017) Repetitive transcranial magnetic stimulation for the acute treatment of major depressive episodes: a systematic review with network meta-analysis. *JAMA Psychiatry* 74(2):143–152
8. Mayberg HS, Brannan SK, Tekell JL, Silva JA, Mahurin RK, McGinnis S et al (2000) Regional metabolic effects of fluoxetine in major depression: serial changes and relationship to clinical response. *Biol Psychiatry* 48(8):830–843
9. Milev RV, Giacobbe P, Kennedy SH, Blumberger DM, Daskalakis ZJ, Downar J et al (2016) Canadian network for mood and anxiety treatments (CANMAT) 2016 clinical guidelines for the management of adults with major depressive disorder: sect. 4. Neurostimulation treatments. *Can J Psychiatry* 61(9):561–575
10. Meijs H, Lin B, van Wingen G, Gordon E, Denys D, de Wilde B, van Hecke J, Niemegeers P, van Eijk K, Luykx J (2021) A polygenic-informed approach to a predictive EEG signature empowers stratified antidepressant treatment. *Brain Stimul* 14(6):1694–1695
11. Bailey NW, Krepel N, van Dijk H, Leuchter AF, Vila-Rodriguez F, Blumberger DM, Downar J, Wilson A, Daskalakis ZJ, Carpenter LL (2021) Resting EEG theta connectivity and alpha power to predict repetitive transcranial magnetic stimulation response in depression: a non-replication from the ICON-DB consortium. *Clin Neurophysiol* 132(2):650–659
12. Shalhaf R, Brenner C, Pang C, Blumberger DM, Downar J, Daskalakis ZJ et al (2018) Non-linear entropy analysis in EEG to predict treatment response to repetitive transcranial magnetic stimulation in depression. *Front Pharmacol* 9:1188
13. Silverstein WK, Noda Y, Barr MS, Vila-Rodriguez F, Rajji TK, Fitzgerald PB et al (2015) Neurobiological predictors of response to dorsolateral prefrontal cortex repetitive transcranial magnetic stimulation in depression: a systematic review. *Depress Anxiety* 32(12):871–891
14. Fregni F, Marcolin MA, Myczkowski M, Amiaz R, Hasey G, Rumi DO et al (2006) Predictors of antidepressant response in clinical trials of transcranial magnetic stimulation. *Int J Neuropsychopharmacol* 9(6):641–654
15. Huang CC, Wei IH, Chou YH, Su TP (2008) Effect of age, gender, menopausal status, and ovarian hormonal level on rTMS in treatment-resistant depression. *Psychoneuroendocrinology* 33(6):821–831
16. Rostami R, Kazemi R, Nitsche MA, Gholipour F, Salehinejad M (2017) Clinical and demographic predictors of response to rTMS treatment in unipolar and bipolar depressive disorders. *Clin Neurophysiol* 128(10):1961–1970
17. Shalhaf A, Bagherzadeh S, Maghsoudi A (2020) Transfer learning with deep convolutional neural network for automated detection of schizophrenia from EEG signals. *Phys Eng Sci Med* 2020:1–11
18. Shalhaf A, Shalhaf R, Saffar M, Sleigh J (2020) Monitoring the level of hypnosis using a hierarchical SVM system. *J Clin Monit Comput* 34(2):331–338
19. Shahabi MS, Shalhaf A, Maghsoudi A (2021) Prediction of drug response in major depressive disorder using ensemble of transfer learning with convolutional neural network based on EEG. *BioCybern Biomed Eng* 41(3):946–959
20. Arns M, Meijs H, Lin B, van Wingen G, Gordon E, Denys D, De Wilde B, Van Hecke J, Niemegeers P, van Eijk K (2021) Can polygenic-informed EEG biomarkers predict differential antidepressant treatment response An EEG stratification marker for rTMS and sertraline. *Res Sq*. <https://doi.org/10.21203/rs.3.rs-155886/v1>
21. Micoulaud-Franchi J-A, Richieri R, Cermolacce M, Loundou A, Lancon C, Vion-Dury J (2012) Parieto-temporal alpha EEG band power at baseline as a predictor of antidepressant treatment response with repetitive transcranial magnetic stimulation: a preliminary study. *J Affect Disord* 137(1–3):156–160
22. Arns M, Cerquera A, Gutiérrez RM, Hasselman F, Freund JA (2014) Non-linear EEG analyses predict non-response to rTMS

- treatment in major depressive disorder. *Clin Neurophysiol* 125(7):1392–1399
23. Corlier J, Wilson A, Hunter AM, Vince-Cruz N, Krantz D, Levitt J, Minzenberg MJ, Ginder N, Cook IA, Leuchter AF (2019) Changes in functional connectivity predict outcome of repetitive transcranial magnetic stimulation treatment of major depressive disorder. *Cereb Cortex* 29(12):4958–4967
 24. Bailey N, Hoy K, Rogasch N, Thomson R, McQueen S, Elliot D, Sullivan C, Fulcher B, Daskalakis Z, Fitzgerald P (2018) Responders to rTMS for depression show increased fronto-midline theta and theta connectivity compared to non-responders. *Brain Stimul* 11:190–203
 25. Hasanzadeh F, Mohebbi M, Rostami R (2019) Prediction of rTMS treatment response in major depressive disorder using machine learning techniques and nonlinear features of EEG signal. *J Affect Disord* 256:132–142
 26. Hasanzadeh F, Mohebbi M, Rostami R (2020) Single channel EEG classification: a case study on prediction of major depressive disorder treatment outcome. *IEEE Access* 9:3417–3427
 27. Lynn CW, Bassett DS (2019) The physics of brain network structure, function and control. *Nat Reviews Phys* 1(5):318–332
 28. Friston KJ (2011) Functional and effective connectivity: a review. *Brain Connect* 1(1):13–36
 29. First MB (1997) Structured clinical interview for DSM-IV axis I disorders. Biometrics Research Department
 30. Cao X, Deng C, Su X, Guo Y (2018) Response and remission rates following high-frequency vs. low-frequency repetitive transcranial magnetic stimulation (rTMS) over right DLPFC for treating major depressive disorder (MDD): a meta-analysis of randomized, double-blind trials. *Front Psychiatry* 2018:413
 31. Berlim MT, Van den Eynde F, Jeff Daskalakis Z (2013) Clinically meaningful efficacy and acceptability of low-frequency repetitive transcranial magnetic stimulation (rTMS) for treating primary major depression: a meta-analysis of randomized, double-blind and sham-controlled trials. *Neuropsychopharmacology* 38(4):543–551
 32. Delorme A, Makeig S (2004) EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J Neurosci Methods* 134(1):9–21
 33. Mullen T (2012) CleanLine EEGLAB plugin. Neuroimaging Informatics Tools and Resources Clearinghouse (NITRC), San Diego
 34. Friston KJ (1994) Functional and effective connectivity in neuroimaging: a synthesis. *Hum Brain Mapp* 2(1–2):56–78
 35. Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Autom Control* 19(6):716–723
 36. Korzeniewska A, Mańczak M, Kamiński M, Blinowska KJ, Kasicki S (2003) Determination of information flow direction among brain structures by a modified directed transfer function (dDTF) method. *J Neurosci Methods* 125(1–2):195–207
 37. Roffo G (2016) Feature selection library (MATLAB toolbox). arXiv preprint arXiv:160701327.
 38. Mamitsuka H (2006) Selecting features in microarray classification using ROC curves. *Pattern Recogn* 39(12):2393–2404
 39. Mumtaz W, Xia L, Yasin MAM, Ali SSA, Malik AS (2017) A wavelet-based technique to predict treatment outcome for major depressive disorder. *PloS One*. <https://doi.org/10.1371/journal.pone.0171409>
 40. Peng H, Long F, Ding C (2005) Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans Pattern Anal Mach Intell* 27(8):1226–1238
 41. Tharwat A, Gaber T, Ibrahim A, Hassanien AE (2017) Linear discriminant analysis: a detailed tutorial. *AI Commun* 30(2):169–190
 42. Savas C, Dovis F (2019) The impact of different kernel functions on the performance of scintillation detection based on support vector machines. *Sensors* 19(23):5219
 43. Fay MP, Proschan MA (2010) Wilcoxon-Mann-Whitney or t-test? On assumptions for hypothesis tests and multiple interpretations of decision rules. *Stat Surv* 4:1
 44. Mullen T (2010) Source information flow toolbox (SIFT). Swartz Center for Computational Neuroscience, San Diego
 45. Xia M, Wang J, He Y (2013) BrainNet viewer: a network visualization tool for human brain connectomics. *PloS One*. <https://doi.org/10.1371/journal.pone.0068910>
 46. He R, Fan J, Wang H, Zhong Y, Ma J (2020) Differentiating responders and Non-responders to rTMS treatment for disorder of consciousness using EEG after-effects. *Front Neurol* 2020:11
 47. Ebrahimzadeh E, Asgarinejad M, Saliminia S, Ashoori S, Seraji M (2021) Predicting clinical response to, transcranial magnetic stimulation in major depression using time-frequency EEG signal processing. *Biomed Eng* 2021:2150048
 48. Hunter AM, Cook IA, Leuchter AF (2007) The promise of the quantitative electroencephalogram as a predictor of antidepressant treatment outcomes in major depressive disorder. *Psychiatr Clin North Am* 30(1):105–124
 49. Knott V, Mahoney C, Kennedy S, Evans K (2000) Pre-treatment EEG and it's relationship to depression severity and paroxetine treatment outcome. *Pharmacopsychiatry* 33(06):201–205
 50. Korb AS, Hunter AM, Cook IA, Leuchter AF (2009) Rostral anterior cingulate cortex theta current density and response to antidepressants and placebo in major depression. *Clin Neurophysiol* 120(7):1313–1319
 51. Mulert C, Juckel G, Brunmeier M, Karch S, Leicht G, Mergl R et al (2007) Rostral anterior cingulate cortex activity in the theta band predicts response to antidepressive medication. *Clin EEG Neurosci* 38(2):78–81
 52. Simpson JR, Snyder AZ, Gusnard DA, Raichle ME (2001) Emotion-induced changes in human medial prefrontal cortex: I. During cognitive task performance. *Proc Natl Acad Sci* 98(2):683–687
 53. Zeng LL, Shen H, Liu L, Hu D (2014) Unsupervised classification of major depression using functional connectivity MRI. *Hum Brain Mapp* 35(4):1630–1641

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.