Research paper

Prediction of rTMS treatment response in major depressive disorder using machine learning techniques and nonlinear features of EEG signal

Fatemeh Hasanzadeh a, Maryam Mohebbi a,b,⁎, Reza Rostami b

a Department of Biomedical Engineering, Faculty of Electrical Engineering, K.N. Toosi University of Technology, Tehran, Iran
b Department of Psychology, University of Tehran, Tehran, Iran

ARTICLE INFO

Keywords:
Classification
EEG
Major depressive disorder
Prediction treatment response
Transcranial magnetic stimulation

ABSTRACT

Background: Prediction of therapeutic outcome of repetitive transcranial magnetic stimulation (rTMS) treatment is an important purpose that eliminates financial and psychological consequences of applying inefficient therapy. To achieve this goal we proposed a method based on machine learning to classify responders (R) and non-responders (NR) to rTMS treatment for major depression disorder (MDD) patients.

Methods: 19 electrodes resting state EEG was recorded from 46 MDD patients before treatment. Then patients underwent 7 weeks of rTMS, and 23 of them responded to treatment. Features extracted from EEG include Lempel-Ziv complexity (LZC), Katz fractal dimension (KFD), correlation dimension (CD), the power spectral density, features based on bispectrum, frontal and prefrontal cordance and combination of them. The most relevant features were selected by the minimal-redundancy-maximal-relevance (mRMR) feature selection algorithm. For classifying two groups of R and NR, k-nearest neighbors (KNN) were applied. The performance of the proposed method was evaluated by leave-1-out cross-validation. For further study, the capability of features in differentiating R and NR was investigated by a statistical test.

Results: Effective EEG features for prediction of rTMS treatment response were found. EEG beta power, the sum of bispectrum diagonal elements in delta and beta bands and CD were the most discriminative features. Power of beta classified R and NR with the high performance of 91.3% accuracy, 91.3% specificity, and 91.3% sensitivity.

Limitations: Lack of large sample size restricted our method for using in clinical applications.

Conclusion: This considerable high accuracy indicates that our proposed method with power and some of the nonlinear and bispectral features can lead to promising results in predicting treatment outcome of rTMS for MDD patients only by one session pretreatment EEG recording.

1. Introduction

Major depressive disorder (MDD) is a remarkable cause of disability in the world that meanwhile more than 300 million people suffer from it in the world (World Health Organization, 2018). Even though this prevalent mental disorder is treatable but for many MDD patients, the initial treatments do not lead to remission (Simon et al., 2006). Thus they experience trial and error periods with different treatments before finding the optimal one. This trial and error approach is associated with increased psychiatric and medical comorbidities, experiencing more severe symptoms of MDD, reduction of the probability of responding to new treatments and even increment of the risk of suicide for the patient (McIntyre and O'Donovan, 2004). So the current incorrect treatment approach heightens the risk of stopping treatment procedure by patients. Accordingly developing an indicator to predict the treatment outcome is desirable to avoid mentioned problems by eliminating or reducing the lengthy and non-effective therapeutic duration.

Neurophysiological modalities have been used for predicting treatment response in MDD patients includes fMRI (Patel et al., 2015; Redlich et al., 2016; Wade et al., 2016) and EEG. Since EEG is relatively more wide-available and cost-effective than fMRI, it is a good option for developing treatment predicting biomarker. One of the useful methods for this purpose is applying machine learning techniques. In recent years several studies have been done to find an indicator for predicting MDD treatment response, and some researchers used EEG based machine learning techniques for classifying responders from non-responders. Khodayari et al. applied a mixture of factor analysis (MFA) model to features derived from pretreatment EEG to predict the efficacy of selective serotonin reuptake inhibitors (SSRI) for MDD treatment. They obtained a classification accuracy of 87.9% (Khodayari-
Rostamabad et al., 2013). Prediction of therapeutic outcome of SSRI medication has been done in another study with an accuracy of 87.5% by applying logistic regression (LR) on wavelet features of baseline EEG (Mumtaz et al., 2017). The effectiveness of transcranial direct-current stimulation (tDCS) on the improvement of mood and cognition of MDD patients was predicted by classification of EEG features using linear support vector machine (LSVM), linear discriminant analysis (LDA) and neural networks. Reported prediction accuracy of mood labels and cognition labels is 76% and 92% respectively (Al-Kaysi et al., 2017). In the mentioned studies responding to SSRI and tDCS treatments have been investigated.

Repetitive transcranial magnetic stimulation (rTMS) is the other therapy for depression. This effective treatment is non-invasive, safe and pain-free with minimal side effects and does not require anesthesia (Iznak et al., 2015; Wasserman, 2011). However a meta-analysis reported relatively low response rates (40.9%) and remission rates (16.4%) for rTMS (Cao et al., 2018). Considering the merits of rTMS, to reduce the costly ineffective treatments, it is desirable to predict the clinical response to rTMS. Bailey et al. predicted responses to rTMS treatment for depression by classification on EEG data recorded during working memory and they achieved a classification accuracy of 91% (Bailey et al., 2018). They also predicted responses to rTMS by resting EEG recorded at baseline and end of the first week of treatment duration. The classification accuracy of the combination of mood and EEG features by LSVM classifier was 86.6% (Bailey et al., 2019). The features they used were EEG power and weighted phase lag index (wPLI) only in alpha, theta (Bailey et al., 2019) and gamma (Bailey et al., 2018) frequency bands, alpha peak frequency (APF) and frontal theta cordance (Bailey et al., 2019). The other EEG features have been ever used for MDD treatment response prediction by machine learning techniques are power spectral features (Al-Kaysi et al., 2017; Khodayari-Rostamabad et al., 2013), coherence (Khodayari-Rostamabad et al., 2013; Mumtaz et al., 2017), mutual information (MI) (Khodayari-Rostamabad et al., 2012) and wavelet coefficients (Mumtaz et al., 2017). The Power of EEG in the different frequency bands is one of the most frequently studied EEG feature in treatment prediction of depression. It is also reported that power of frequency bands and their combinations including cordance (Leuchter et al., 1994) and ATR index (Josifescu et al., 2009) are related to treatment outcome (Arns et al., 2012; Bruder et al., 2008; Carvalho et al., 2011; Cook and Leuchter, 2001; Knott et al., 2000; Pellicciari et al., 2013; Spronk et al., 2011; Suffin and Emory, 1995; Tenke et al., 2011; Ulrich et al., 1984; Wade et al., 2016). As the EEG is a non-Gaussian signal, higher order statistics of it can provide supplementary information not revealed in the power spectrum (Mohabbi and Ghassemian, 2012). For example, bispectrum which is based on third-order statistics indicates the phase coupling of different frequency components (Mohabbi and Ghassemian, 2012).

To the best of our knowledge, the bispectrum of EEG has not been ever investigated for treatment outcome prediction and generally in depression studies. Furthermore, since EEG is a signal with nonlinear dynamics, applying nonlinear features of EEG may be informative for predicting responding to treatment. However as mentioned before, most of the previous studies in the prediction of MDD treatment outcome only have focused on linear and spectral features, and just a single study has analyzed nonlinear features of EEG including Lempel-Ziv complexity (LZC), largest Lyapunov exponent and false nearest neighbors. This study which was limited to statistical testing reported no differences between responders and non-responders in the analyzed features and only changes of LZC from minute 1 to 2 were different in two groups (Arns et al., 2014).

Given this background, in this paper, to find a predictor for depression treatment, we studied the capability of a more comprehensive set of EEG measures including spectral, bispectral and nonlinear features in predicting rTMS treatment response by machine learning techniques. These measures that are extracted from pretreatment resting EEG includes three bispectrum features, LZC, correlation dimension (CD), Katz fractal dimension (KFD), EEG power in all frequency bands (delta, theta, alpha, and beta) and frontal and prefrontal cordance in the theta band. In addition to classification, to perform a complete analysis, the extracted features were also compared from the statistical viewpoint. So a statistical test was also applied to evaluate differences of features between two groups of responders (R) and non-responders (NR). It is notable that as far as we know, this is the first time that the capability of selected bispectral, nonlinear and spectral EEG measures is investigated for predicting treatment response by machine learning techniques simultaneously.

The rest of paper is organized as follow: in the “Materials and methods” section, firstly the information of participants of the study is explained, then our procedure for acquiring EEG data of MDD patients is described. After EEG data description, the extracted features, classification technique, and statistical test are explained. In the “Results” section we will evaluate the prediction ability of feature sets by classification analysis. The features also compared between two groups of R and NR by statistical analysis. The results are discussed in the "Discussion" section. Some limitation of the current study and suggestion for future work are also presented in this section. Finally, the paper is concluded in the “Conclusion” section.

2. Materials and methods

2.1. Participants

In this study, 46 patients with MDD in the age range of 17–65 years participated. The patients were referred to Atieh Clinical Neuroscience Centre, Tehran Iran and MDD diagnosis for them was made by a psychiatrist based on Diagnostic and Statistical Manual-IV diagnostic (DSM-IV) criteria (Association, 2000). Participants also assessed by Hamilton Rating Scale for Depression (HRSD), and Beck Depression Inventory (BDI-II) and all have the HRSD score ≥12 and BDI-II score ≥15. The demographic and clinical information of participants has been summarized in Table 1. The groups of R and NR were compared by Wilcoxon rank sum and Friedman tests, and the results of tests are shown in column 4 of Table 1. There are no significant differences (p ≤ 0.05) between responders and non-responders in age, gender, rTMS treatment characteristics, pretreatment BDI-II and HRSD score, medications, duration of illness, length of current depression episode, number of previous medications and anxiety comorbidity. Only post-treatment BDI-II and 24-item HRSD scores are significantly different

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Demographic and clinical information.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Non-responders</td>
<td>Responders</td>
</tr>
<tr>
<td>N</td>
<td>23</td>
</tr>
<tr>
<td>Age</td>
<td>39(± 14.6)</td>
</tr>
<tr>
<td>Gender (M/F)</td>
<td>8/15</td>
</tr>
<tr>
<td>Treatment (L/R/Bilateral)</td>
<td>17/3/9</td>
</tr>
<tr>
<td>Pre-treatment BDI-II</td>
<td>28.1(± 9.4)</td>
</tr>
<tr>
<td>Post-treatment BDI-II</td>
<td>23.1(± 8.4)</td>
</tr>
<tr>
<td>Post-treatment HRSD</td>
<td>28.9(± 12.1)</td>
</tr>
<tr>
<td>Post-treatment HRSD</td>
<td>22.4(± 10.6)</td>
</tr>
<tr>
<td>Medications (AD/AD ± MS / AD ± MS ± AP)</td>
<td>5/8/1</td>
</tr>
<tr>
<td>Illness duration (years)</td>
<td>7.9(± 7.8)</td>
</tr>
<tr>
<td>Current episode length (years)</td>
<td>3.2(± 2.2)</td>
</tr>
<tr>
<td>Number of previous medications</td>
<td>2.7(± 1.5)</td>
</tr>
<tr>
<td>Anxiety (Y/N)</td>
<td>14/9</td>
</tr>
</tbody>
</table>

* Protocol of rTMS includes left (L), right (R) or bilateral.

** AD (Antidepressant), MS (mood stabilizer), AP (antipsychotic).
(p ≤ 0.05) between responders and non-responders which are significantly lower in responders. It is notable that as pretreatment BDI-II and HRSD of responders and non-responders are not statistically different; the outcome of rTMS cannot be linked to differences in depression severity of the two groups. The exclusion criteria for this study includes the presence of Axis I or II disorders, substance abuse, suicidal risk, unstable medical conditions, implanting devices, cardiac arrhythmia, and pregnancy. Participants with present or history of head injury, seizures, epilepsy, and neurological disorders were also excluded. The participants also either should not be on antidepressant medication or their medication usage must be unchanged during the rTMS. In this study, 23 patients were under medication treatment (antidepressant, mood stabilizer and antipsychotic) fixedly from more than four weeks before the treatment. This study was approved by the Iran University of Medical Sciences’ ethics committee, and all of the patients gave informed consent for treatment.

2.2. Procedure

The procedure for rTMS treatment was so that the patients underwent five weeks (3 sessions per week) of left DLPFC 10 Hz rTMS treatment. Then the patients that showed response were continued on two weeks of the same rTMS treatment and non-responders were randomly continued to either left 10 Hz, right 1 Hz or bilateral rTMS for two weeks (3 sessions per week). Depression severity was assessed by BDI-II and HRSD at the baseline and the end of treatment. The HRSD was also repeated during the treatment after every five sessions of rTMS. Responding to treatment is defined as more than 50% decrease in BDI-II scores or HRSD or by BDI ≤ 8 (HRSD ≤ 9) which indicates remission. Resting EEG was recorded form the participants at baseline and the end of the treatment. In every session, EEG was acquired for five minutes while they were eye closed and sat on a comfortable chair in a shielded room. The participants were instructed to avoid falling asleep.

2.3. EEG recording and pre-processing

EEG was recorded by a Mitsar-EEG 201 with 19 Ag/AgCl electrodes and a sampling frequency of 500 Hz. Channel location that was based on International 10–20 system includes Fp1, Fp2, F7, F3, Fz, F4, F8, T3, C3, Cz, C4, T4, T5, P3, Pz, P4, T6, O1, and O2. Location of EEG channels is shown in Fig. 1. For preprocessing of recorded EEG data, EEGLAB Toolbox (Delorme and Makeig, 2004) have been used. To remove artifacts including eye blinks, movements and environmental noise the recorded EEG data was preprocessed. In the first step, a bandpass FIR filter with low and high cutoff frequencies of 1 and 42 Hz respectively has been used. Then to remove noisy data, the ICA algorithm has been applied. A multiple artifact rejection algorithm (MARA) that was an open source EEGLAB plug-in (Winkler et al., 2011) was used to label noisy Independent components (IC). By inspection of the power spectrum of ICs and considering the labels, the noisy ICs have been removed. After that, the pruned EEG data was reconstructed. In the next step, the EEG data were visually inspected to eliminate remained artifacts. Finally, all data were re-referenced against the average reference. Considering removed parts of EEG data, for equalization of the length of data for all participants, we retained 220 s of EEG signal for every data.

2.4. Feature extraction

After EEG data preparation, the next step in the prediction of treatment response to rTMS is extracting features. In this paper, we studied a total number of 21 features categorized into four groups including nonlinear, spectral, bispectral and cordance measures. These measures are extracted from the baseline EEG of both groups of R and NR, and each feature (except cordance measures) was computed for all EEG channels. So for each measure, we have a feature set contained 19 feature vectors corresponding to 19 channels. The studied measures are described in the following parts.

2.4.1. Nonlinear features

Nonlinear measures that are applied to the EEG signal includes LZC, CD, and KFD. LZC is a complexity measure of a time series that is an estimator for evaluating stochastic and chaotic behavior of time series (Aboy et al., 2006). KFD is the other nonlinear measure used here. KFD is an algorithm for computing fractal dimension of EEG signal. Fractal dimension is a measure that shows the self-similarity of a time series based on the number of repetitions of a pattern in time series. It is reported that KFD is more robust against noise in comparison with similar algorithms Petrosian and Higuchi fractal dimensions (Ahmadlou et al., 2012). CD is a different way of computing the fractal dimension. It is a measure of the complexity of the system in a state space that is based on chaos and time delay reconstruction theory (Baker et al., 1996). The computational details of these measures can be found in Appendix A. All nonlinear measures were calculated for epochs of EEG data with the length of 3000 samples and then have been averaged over all epochs. Thus for every EEG channel, we have three values that are corresponding to each nonlinear measure for every participant.

2.4.2. Power spectrum features

The power spectrum indicates the power of the signal in frequency components of it. In this study power of EEG signals was estimated in delta (1–4 Hz), theta (4–8 Hz), alpha (8–12 Hz) and beta (12–30 Hz) frequency bands by Welch method with a non-overlapped window with 500 samples length. For every channel, the average power of frequencies contained in each frequency band was computed. Thus for every participant, four feature sets corresponding to each frequency band was obtained.

2.4.3. Bispectrum features

In addition to power features, to obtain information not revealed in the spectrum we derived some features from bispectrum. The bispectrum is defined as the two-dimensional Fourier transform of the third order cumulant. So the bispectrum of a signal is a 2D mapping of the interaction level between all frequency pairs in the desired band of the signal. Therefore to apply bispectrum in analysis, some quantitative features related to moments and entropy of it are extracted in previous researches frequently (Assi et al., 2018; Mohebbi and Ghassemian, 2012). Here we derived three measures from the bispectrum of EEG signal of every channel in the delta, theta, alpha, and beta frequency bands. Two of these measures were related to moments of bispectrum.
that include the sum of logarithmic amplitudes of diagonal elements in the bispectrum and the second-order spectral moment of the amplitudes of diagonal elements in the bispectrum. The other one was based on the entropy of bispectrum that is normalized bispectral entropy. The mathematical equations of bispectrum and extracted measures of it are described in Appendix A. Therefore by this analysis; we obtain 12 sets of feature vectors corresponding to each three measure in four frequency band for every participant.

2.4.4. Cordance features

Cordance is one of the measures that many researchers have been interested in for both discriminating depression and predicting treatment. Some of the previous studies reported that decrement of prefrontal cordance in theta frequency band after the first week of treatment is associated with responding to treatment (Bares et al., 2008, 2015, 2010; Cook and Leuchter, 2001; Cook et al., 2002, 2005). Therefore we are going to investigate whether baseline prefrontal theta cordance by itself can predict rTMS treatment response. A previous study used the three prefrontal electrodes (i.e. Fz, Fp1 and Fp2) in the theta frequency band in computing prefrontal theta cordance for predicting treatment response (Baskaran, 2016), so we calculated prefrontal theta cordance by same electrodes. Cordance was calculated by the formula described in Leuchter et al. (1994). Moreover since the frontal brain has been reported as one of the areas that its activity is highly influenced by depression (Bench et al., 1992; Davidson et al., 1999; Henriques and Davidson, 1991), we also examined whether frontal cordance (Fp1, Fp2, F7, F3, Fz, F4, F8 electrodes) in theta frequency band can differentiate participants who responded to rTMS treatment from non-responders. So for every participant, we have two feature sets containing three and seven features that are related to prefrontal and frontal theta cordance respectively.

2.5. Classification

To achieve the aim of this paper which is predicting the outcome of rTMS treatment, we proposed a method based on classification analysis to distinguish responders from non-responders. In this way, we performed our analysis on a different combination of extracted features. At first, we compared the treatment response prediction ability of studied features by applying classification independently for every measure. Then a combination of features in every four categories of measures (i.e., nonlinear, spectral, bispectral and cordance measures) was used for classification of R and NR in separate classification analysis. Finally, we evaluate the performance of the combination of the total studied features in classifying R and NR. The procedure of classification that is described at the following was similar for all analysis. After feature extraction, for the first step, the feature sets were standardized by Z-score normalization independently to eliminate amplitude variations due to differences of participants and electrode placements. Since irrelevant features may increase the complexity of classification and deteriorate its accuracy, a promising step before classification is feature selection. Therefore after EEG feature extraction and standardization, we used minimal redundancy-maximal-relevance (mRMR) algorithm to select the most informative features for applying to the classifier. The mRMR algorithm is a feature selection method that selects features that have maximum relevancy to the target classes and minimum redundancy with each other simultaneously (Ding and Peng, 2005; Peng et al., 2005). In this study, we applied mRMR based on mutual information quotient scheme by the toolbox available at Matlab source codes exchange site (Peng, 2007). In all analysis, the number of selected features is chosen in a trial and error procedure to find a value which maximizes the accuracy and simultaneously prevents overfitting of the classifier. In our analysis, this optimal number of features was matched to the square root of the total number of features in the corresponding analysis (exceptionally for prefrontal cordance analysis, that due to the low number of features we used all features for classification).

After feature selection, final features were applied to k-nearest neighbors (KNN) classifier that $k = 2$ is the number of nearest neighbors. The number of neighbors was selected by trial and error to minimize classification error. We used leave-one-out cross-validation method to validate the performance of the classifier, method that is for 46 available subjects in this study, 45 subjects used for training and the remaining one for testing the classifier. This procedure repeated 46 times through all subjects so that each subject was the test subject once. Then the classification results were averaged. The criterions used for evaluating response prediction were specificity, sensitivity, and accuracy. Specificity indicates the percentage of non-responders who were predicted as non-responders, sensitivity shows the percentage of responders who were predicted as responders and accuracy shows the percentage of how often the classifier predicts these correct responses overall.

2.6. Statistical comparison

The goal of statistical testing is to understand how the system behaves while the machine learning techniques aim to predict future behavior (unobserved outcome) (Bzdok et al., 2018). Therefore in the current study, they can be complementary to achieve a more comprehensive assessment of the ability of studied measures in the detection of R and NR to rTMS treatment. Additionally to compare the results of our data with previous studies that used statistical analysis, we have performed a statistical test. Although each measure is influenced by two factors including two groups (responders and non-responders) and 19 EEG electrodes (channels), our focus is on group differences. Since some of the data did not have a normal distribution, we applied the Friedman test ($p<0.05$) to compare every measure between two groups of non-responders and responders. The compared measures of this part were the same as the single measures of classification section.

3. Results

3.1. Classification analysis

As mentioned in Section 2.5, to evaluate the capability of selected measures in the prediction of responding to rTMS treatment, we applied feature sets to KNN classifier in the following forms: every measure independently, the combination of measures in each group of non-linear, spectral, bispectral and cordance measures and the combination of all features. The criterions applied for evaluating our proposed classification method include classification accuracy, specificity, and sensitivity. Table 2 shows these criteria for classification by every single measure which was calculated for 19 channel (i.e., 19 features) and reduced to 4 features by mRMR. The evaluation of classification by the combination of studied measures are also given in Table 3. For simplification in the rest of this paper, bispectrum features that are the sum of logarithmic amplitudes of diagonal elements in the bispectrum, the second-order spectral moment of the amplitudes of diagonal elements in the bispectrum and normalized bispectral entropy are shown by BispSL, Bisp2M, and BispEn respectively.

To represent the differences of measures in classifying responders from non-responders more clearly, the reported accuracy of classification by all measures is depicted in Fig. 2. As it is evident from Tables 2, 3 and Fig. 2, beta power of baseline EEG is the most predictive measure for treatment response with an accuracy of 91.3%. Power of EEG in other frequency bands predict responding with lower accuracy of 84.8% for both theta and alpha bands and 80.4% for delta band. Among the other features, the sum of logarithmic amplitudes of diagonal elements of bispectrum in delta and beta bands and CD were the most discriminative features. In each one of delta, alpha and beta frequency band, the sum of diagonal elements of bispectrum is also the highest discriminative measure in comparison to the other two bispectrum measures. While the second moment of bispectrum in theta band gives
features in delta bands are significantly different in responders and non-responders. None of the cordance features had significant differences between the two groups. Since the results of the statistical test on power features indicate differences between responders and non-responders, to display differences more clear, the average power of the two groups in all frequency bands are illustrated in Fig. 3. From Fig. 3(a) it is evident that the delta power of occipital and frontal especially left frontal in non-responders is higher than responders. The Fig. 3(c) also indicates that obtained higher alpha power of R than NR is related to the occipital region. From Fig. 3(d) it is obvious that an extensive area in NR has a higher activity than R which is confirmed by results of the statistical test in Table 4.

It can be seen in Table 4 that the first and second bispectrum features in the delta frequency band and all of them in beta are significantly different between two groups of responders and non-responders. As an example, the average bispectrum plot for two groups of R and NR in all frequency bands is depicted in Fig. 4. The Fig. 4 shows that beta and delta frequency bands, phase coupling exists between a broader range of frequencies in responders that means the phase of EEG signal of responders in different frequencies has higher coupling than non-responders.

4. Discussion

To develop a method for predicting the therapeutic outcome of rTMS by EEG signals, we derived several measures from pretreatment resting state EEG of MDD patients who were received rTMS treatment (23 non-responders, 23 responders). However, most of the previous studies including (Bailey et al., 2018, 2019) focused on the power of alpha and theta frequency bands, our classification analysis shows that beta power can be an excellent candidate to predict treatment response with high accuracy. Also from the statistical point of view, beta power is significantly lower in responses than non-responders. In some studies, it is demonstrated that higher beta activity is related to depression level (Knott et al., 2001; Lieber and Prichep, 1988). Therefore based on our results, it can be speculated that the high beta power of EEG may indicates a lower probability of responding to treatment. About the power of other frequency bands despite equivalent and relatively high classification accuracy of power of theta and alpha bands, only alpha and delta power were statistically different between the two groups. The results indicate higher pretreatment alpha power in responders than non-responders. This increased alpha power seen in responders is observed in many studies (Bruder et al., 2008, 2001; Knott et al., 2000; Suffin and Emory, 1995; Ulrich et al., 1984). To explain the meaning of these changes in alpha power (Bruder et al., 2008) declared that elevated pretreatment alpha power might be an indicator of the correspondence between low arousal and low serotonergic activity. It has been shown that the activity of 5-HT mediates arousal and serotonergic activity may indicate activity of the mesencephalic raphe nuclei and cortical afferents (Bruder et al., 2008). Furthermore, depression can be related to dysfunction of temporoparietal mechanisms that may interfere low arousal (Heller et al., 1995). Based on the results, delta power is lower in rTMS responders in comparison with non-responders that was also observed in other studies (Knott et al., 2000, 1996). The frontal lobe is one of the highly attended brain regions in researches on depression. Significant differences in frontal delta power of responders and non-responders can also be seen in Fig. 3. About the power of theta frequency band, previous studies did not provide consistent result in responding to rTMS treatment since higher theta power is reported for both responders (Woźniak-Kwaśniewska et al., 2015) and non-responders (Arns et al., 2012). In our study similar to another study (Cook et al., 1999), no significant differences between responders and non-responders was observed for theta power. About the nonlinear features, it is observed that many researchers neglect these features in the prediction of depression treatment response, while based on our results these features perform well in differentiating R/NR groups. CD

### Table 3
The performance of the proposed classification method for single measures.

<table>
<thead>
<tr>
<th>Feature set</th>
<th>Accuracy (%)</th>
<th>Specificity (%)</th>
<th>Sensitivity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Nonlinear (LZC,KFD)</td>
<td>80.4</td>
<td>87</td>
<td>73.9</td>
</tr>
<tr>
<td>2 Power (D,T,A,B)</td>
<td>91.3</td>
<td>95.7</td>
<td>87</td>
</tr>
<tr>
<td>3 Bispectrum (BispSL, Bisp2M and BispEn in all bands)</td>
<td>84.8</td>
<td>82.6</td>
<td>87</td>
</tr>
<tr>
<td>4 Cordance (Fr,PreFr)</td>
<td>76.1</td>
<td>82.6</td>
<td>69.6</td>
</tr>
<tr>
<td>5 All</td>
<td>87</td>
<td>91.3</td>
<td>82.6</td>
</tr>
</tbody>
</table>

D, T, A, and B denote delta, theta, alpha and beta frequency band respectively.

### Table 2
The performance of the proposed classification method for single measures.

<table>
<thead>
<tr>
<th>Feature set</th>
<th>Accuracy (%)</th>
<th>Specificity (%)</th>
<th>Sensitivity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 LZC</td>
<td>82.6</td>
<td>87</td>
<td>78.3</td>
</tr>
<tr>
<td>2 KFD</td>
<td>82.6</td>
<td>82.6</td>
<td>82.6</td>
</tr>
<tr>
<td>3 CD</td>
<td>87</td>
<td>91.3</td>
<td>82.6</td>
</tr>
<tr>
<td>4 Power-D</td>
<td>80.4</td>
<td>82.6</td>
<td>78.3</td>
</tr>
<tr>
<td>5 Power-T</td>
<td>84.8</td>
<td>78.3</td>
<td>91.3</td>
</tr>
<tr>
<td>6 Power-A</td>
<td>84.8</td>
<td>87</td>
<td>82.6</td>
</tr>
<tr>
<td>7 Power-B</td>
<td>91.3</td>
<td>91.3</td>
<td>91.3</td>
</tr>
<tr>
<td>8 BispSL-D</td>
<td>89.1</td>
<td>91.3</td>
<td>87</td>
</tr>
<tr>
<td>9 Bisp2M-D</td>
<td>84.8</td>
<td>91.3</td>
<td>78.3</td>
</tr>
<tr>
<td>10 BispEn-D</td>
<td>82.6</td>
<td>87</td>
<td>82.6</td>
</tr>
<tr>
<td>11 BispSL-T</td>
<td>80.4</td>
<td>78.3</td>
<td>82.6</td>
</tr>
<tr>
<td>12 Bisp2M-T</td>
<td>76.1</td>
<td>69.6</td>
<td>82.6</td>
</tr>
<tr>
<td>13 BispEn-T</td>
<td>82.6</td>
<td>73.9</td>
<td>91.3</td>
</tr>
<tr>
<td>14 BispSL-A</td>
<td>82.6</td>
<td>82.6</td>
<td>82.6</td>
</tr>
<tr>
<td>15 Bisp2M-A</td>
<td>78.3</td>
<td>73.9</td>
<td>82.6</td>
</tr>
<tr>
<td>16 BispEn-A</td>
<td>80.4</td>
<td>78.3</td>
<td>82.6</td>
</tr>
<tr>
<td>17 BispSL-B</td>
<td>87</td>
<td>87</td>
<td>87</td>
</tr>
<tr>
<td>18 Bisp2M-B</td>
<td>82.6</td>
<td>78.3</td>
<td>82.6</td>
</tr>
<tr>
<td>19 BispEn-B</td>
<td>80.4</td>
<td>82.6</td>
<td>78.3</td>
</tr>
<tr>
<td>20 FrCord-T</td>
<td>80.4</td>
<td>87</td>
<td>73.9</td>
</tr>
<tr>
<td>21 PreFrCord-T</td>
<td>78.3</td>
<td>73.9</td>
<td>82.6</td>
</tr>
</tbody>
</table>

The results show that between nonlinear measures CD of non-responders (Arns et al., 2012). In our study similar to another study (Woźniak-Kwaśniewska et al., 2015) and non-responders can also be seen in Fig. 3. About the power of theta frequency band, previous studies did not provide consistent result in responding to rTMS treatment since higher theta power is reported for both responders (Woźniak-Kwaśniewska et al., 2015) and non-responders (Arns et al., 2012). In our study similar to another study (Cook et al., 1999), no significant differences between responders and non-responders was observed for theta power. About the nonlinear features, it is observed that many researchers neglect these features in the prediction of depression treatment response, while based on our results these features perform well in differentiating R/NR groups. CD

### 3.2. Statistical analysis

The results of the comparison of features extracted from the baseline EEG of non-responders and responders by Friedman test have been shown in Table 4. Mean, and standard deviation of every feature set on individuals of each group is also listed in columns 6 to 9 of Table 4.

The results show that between nonlinear measures CD of non-responders is significantly higher than responders. Among the EEG power, power of all frequency bands except theta are statistically different between the two groups. In delta and beta bands responders have lower power than non-responders while alpha power of responders is significantly higher than non-responders. It also can be seen that all bispectrum features in the beta band and first and second bispectrum

---


F. Hasanzadeh, et al.
of EEG which is significantly lower in responders provides high classification accuracy of 87%. The significant higher CD in non-responders indicates higher dimensionality that is considered as a reflection of a wider distribution of oscillation of neural nodes at lower synchrony (Ibáñez-Molina et al., 2018). This can be an indicator of isolation of brain nodes (Friston et al., 1995) and disorganized spiking activity (Takahashi et al., 2010). So the observed higher CD in NR can be a sign of more abnormal brain activity in this group in comparison to the R group. Our classification analysis also illustrates that other nonlinear measures, i.e., LZC and KFD both with providing 82.6% classification accuracy, represent relatively high discriminative power. In spite of the high classification accuracy, LZC and KFD were not statistically different between responders and non-responders. Analogous to our results, another study that compared LZC between treatment responders and non-responders, didn’t observe differences between LZC of two groups and only reported a significant decrease in LZC from minute 1 to

Fig. 2. Classification accuracy of every measure in the prediction of treatment response. D, T, A, and B denote delta, theta, alpha, and beta frequency band respectively.

### Table 4

<table>
<thead>
<tr>
<th>Feature set</th>
<th>N</th>
<th>Chi-sq</th>
<th>p-value</th>
<th>NR Mean</th>
<th>NR SD</th>
<th>R Mean</th>
<th>R SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LZC</td>
<td>19</td>
<td>3.43</td>
<td>0.06</td>
<td>180</td>
<td>31.76</td>
<td>173.69</td>
</tr>
<tr>
<td>2</td>
<td>KFD</td>
<td>19</td>
<td>3.37</td>
<td>0.07</td>
<td>1.04</td>
<td>0.02</td>
<td>1.03</td>
</tr>
<tr>
<td>3</td>
<td>CD</td>
<td>19</td>
<td>4.8</td>
<td>0.03</td>
<td>1.952</td>
<td>0.007</td>
<td>1.94</td>
</tr>
<tr>
<td>4</td>
<td>Power-D</td>
<td>19</td>
<td>11.15</td>
<td>0</td>
<td>2.93</td>
<td>2.27</td>
<td>2.18</td>
</tr>
<tr>
<td>5</td>
<td>Power-T</td>
<td>19</td>
<td>3.25</td>
<td>0.07</td>
<td>2.37</td>
<td>2.63</td>
<td>1.79</td>
</tr>
<tr>
<td>6</td>
<td>Power-A</td>
<td>19</td>
<td>5.04</td>
<td>0.02</td>
<td>5.86</td>
<td>5.58</td>
<td>6.74</td>
</tr>
<tr>
<td>7</td>
<td>Power-B</td>
<td>19</td>
<td>16.12</td>
<td>0</td>
<td>0.47</td>
<td>0.29</td>
<td>0.43</td>
</tr>
<tr>
<td>8</td>
<td>BisplSL-D</td>
<td>19</td>
<td>9.1</td>
<td>0</td>
<td>−7941.14</td>
<td>2464.52</td>
<td>−8592.68</td>
</tr>
<tr>
<td>9</td>
<td>Bispl2M-D</td>
<td>19</td>
<td>8.86</td>
<td>0</td>
<td>−2.77</td>
<td>2</td>
<td>−3.08</td>
</tr>
<tr>
<td>10</td>
<td>BisplEn-D</td>
<td>19</td>
<td>0.37</td>
<td>0.54</td>
<td>2.27</td>
<td>0.13</td>
<td>2.28</td>
</tr>
<tr>
<td>11</td>
<td>BisplSL-T</td>
<td>19</td>
<td>0.78</td>
<td>0.38</td>
<td>−9967.7</td>
<td>2728.47</td>
<td>−10179</td>
</tr>
<tr>
<td>12</td>
<td>Bispl2M-T</td>
<td>19</td>
<td>0.38</td>
<td>0.54</td>
<td>−4.76</td>
<td>3.48</td>
<td>−4.64</td>
</tr>
<tr>
<td>13</td>
<td>BisplEn-T</td>
<td>19</td>
<td>0.28</td>
<td>0.6</td>
<td>1.83</td>
<td>0.06</td>
<td>1.84</td>
</tr>
<tr>
<td>14</td>
<td>BisplSL-A</td>
<td>19</td>
<td>0.97</td>
<td>0.32</td>
<td>−8453.67</td>
<td>3005.21</td>
<td>−7971.04</td>
</tr>
<tr>
<td>15</td>
<td>Bispl2M-A</td>
<td>19</td>
<td>0.65</td>
<td>0.42</td>
<td>−3.53</td>
<td>3.87</td>
<td>−2.68</td>
</tr>
<tr>
<td>16</td>
<td>BisplEn-A</td>
<td>19</td>
<td>0.03</td>
<td>0.86</td>
<td>0.94</td>
<td>0.08</td>
<td>0.95</td>
</tr>
<tr>
<td>17</td>
<td>BisplSL-B</td>
<td>19</td>
<td>28.78</td>
<td>0</td>
<td>−8599.04</td>
<td>1766.88</td>
<td>−9604.58</td>
</tr>
<tr>
<td>18</td>
<td>Bispl2M-B</td>
<td>19</td>
<td>22.85</td>
<td>0</td>
<td>−2.98</td>
<td>1.9</td>
<td>−3.91</td>
</tr>
<tr>
<td>19</td>
<td>BisplEn-B</td>
<td>19</td>
<td>11.25</td>
<td>0</td>
<td>2.12</td>
<td>0.09</td>
<td>2.08</td>
</tr>
<tr>
<td>20</td>
<td>FrCord-T</td>
<td>7</td>
<td>0.24</td>
<td>0.63</td>
<td>0.41</td>
<td>0.13</td>
<td>0.42</td>
</tr>
<tr>
<td>21</td>
<td>PreFrCord-T</td>
<td>3</td>
<td>0.14</td>
<td>0.71</td>
<td>0.38</td>
<td>0.14</td>
<td>0.38</td>
</tr>
</tbody>
</table>

* p < 0.05.
** p < 0.001.
This difference between the results of KNN classification and statistical analysis may be referred to the entirely different theoretical basis of these two techniques. As the first one seeks to find a model that be able to predict the outcome of a new observation, but the aim of the latter is evaluating whether there is a relationship between studied parameters (Bzdok et al., 2018). The results of both classification and statistical analysis show that among the bispectrum measures, the sum of bispectrum diagonal elements in the delta and especially in the beta band are highly informative about the treatment outcome. Considering classification accuracy of all studied bispectrum measures in delta and beta band it can be seen that bispectrum of these two bands outperform than theta and alpha bands in classifying R/NR. While based on statistical analysis, only all bispectrum measures extracted from the beta band are significantly different between the two groups. One of the other measures investigated in this study was theta cordance. The classification accuracy of frontal and prefrontal theta

Fig. 3. Average pretreatment power of (a) delta, (b) theta, (c) alpha and (d) beta frequency bands for non-responders and responders to rTMS treatment. The color bar is the same for both plots of every band.

Fig. 4. Average bispectrum of one EEG channel for non-responders and responders in (a) delta, (b) theta, (c) alpha, and (d) beta frequency bands.

Table 5
Comparison of classification results of studies that applied machine learning techniques for prediction of MDD treatment response with the current study.

<table>
<thead>
<tr>
<th>Features</th>
<th>Accuracy (%)</th>
<th>Specificity (%)</th>
<th>Sensitivity (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MADRS*, PSD, wPLI, IAPF, theta cordance (Bailey et al., 2019)</td>
<td>86.60</td>
<td>89</td>
<td>84</td>
</tr>
<tr>
<td>MADRS*, Working Memory accuracy, PSD, wPLI (Bailey et al., 2018)</td>
<td>91</td>
<td>92</td>
<td>91</td>
</tr>
<tr>
<td>PSD, PSD ratio, Coherence, MI (Khodayari-Rostamabad et al., 2013)</td>
<td>87.9</td>
<td>80.9</td>
<td>94.9</td>
</tr>
<tr>
<td>PSD, Alpha asymmetry (Al-Kaysi et al., 2017)</td>
<td>76 (Mood) 92 (Cognition)</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Wavelet coefficients (Muntaz et al., 2017)</td>
<td>87.5</td>
<td>95</td>
<td>80</td>
</tr>
<tr>
<td>Our proposed method based on beta power features</td>
<td>91.3</td>
<td>91.3</td>
<td>91.3</td>
</tr>
</tbody>
</table>

* MADRS: Montgomery and Asberg Depression Rating Scale.

2 (Arns et al., 2014). This difference between the results of KNN classification and statistical analysis may be referred to the entirely different theoretical basis of these two techniques. As the first one seeks to find a model that be able to predict the outcome of a new observation, but the aim of the latter is evaluating whether there is a relationship between studied parameters (Bzdok et al., 2018). The results of both classification and statistical analysis show that among the bispectrum measures, the sum of bispectrum diagonal elements in the delta and especially in the beta band are highly informative about the treatment outcome. Considering classification accuracy of all studied bispectrum measures in delta and beta band it can be seen that bispectrum of these two bands outperform than theta and alpha bands in classifying R/NR. While based on statistical analysis, only all bispectrum measures extracted from the beta band are significantly different between the two groups. One of the other measures investigated in this study was theta cordance. The classification accuracy of frontal and prefrontal theta
cordance was 80.4% and 78.3% respectively. Statistical testing indicates no significant differences between responders and non-responders for these measures that are consistent with some studies that reported no differences in prefrontal theta cordance between R and NR to rTMS (Arns et al., 2012; Bailey et al., 2019). As classification results illustrated, our proposed method lead to high classification accuracy for most of the studied measures. To compare our results with previous studies that applied machine learning techniques for prediction of MDD treatment response, the features and classification results of these analysis are shown in Table 5.

As seen in Table 5, our proposed method outperformed the previous analysis except the cognition output in Al-Kaysi et al. (2017). Al-Kaysi et al. predicted mood and cognition output separately. They assigned 0 or 1 values to mood (and cognition) based on decrease or increase of mood (and cognition) scores of the middle of the treatment from the baseline score. Then these assigned values which are considered as outputs were predicted from EEG features by machine learning techniques. It is notable the defined cognition output only shows changes of Symbol Digit Modalities Test (SDMT) after applying the treatment. Therefore since SDMT is not a depression severity rating scale, the results of cognition output must not be considered as the output of depression treatment outcome prediction. However, the mood output which shows changes of MADRS has a low accuracy of 76%. It should be mentioned that the other study (Bailey et al., 2018) which has high accuracy, applied mood measures (MADRS) in addition to EEG measures. The mood measures can be affected by expert rating, i.e., can be different from one to another. Thus since our method is only based on EEG activity of patients, it is more accurate than the mentioned study. Moreover, Bailey et al. differentiated responders and non-responders to rTMS treatment by using EEG of two sessions (pretreatment and after one week of treatment). While our proposed method presented high classification result (accuracy=91.3%, specificity=91.3%, sensitivity=91.3%) by using only one session pretreatment EEG recording. The requirement of only one session EEG indicates the efficiency of our method. Moreover it removes the economic and psychological burden of undergoing one week of a potentially ineffective treatment.

The current study has some limitations. The first one is related to sample size. As the number of participants in this study was not very high, we cannot predict treatment response by these results confidently in real applications. So to make the proposed method applicable to clinical using, it should be tested on a broader set of patients. The second one is about the patients who were on antidepressant medications; hence their brain activity, in addition to rTMS may have also been affected by medications. Thus if all patients only underwent rTMS treatment, it would be easier to interpret the results. The last one is referred to the clinical information that the number of previous depressive episodes was not reported as most of the patients were not sure about the number of experienced depressive episodes, so we reported the number of times that patients used medications.

For future research, to provide a simple, cheap and portable monitoring system for evaluating treatment responding, we can use only EEG electrodes which are highly informative. For this purpose, the selected features for classification can be investigated to find the EEG channels that are most helpful for the prediction of treatment response. Further research also can be done on evaluating the proposed method on the other therapies of MDD rather than rTMS.

5. Conclusion

In this study, we proposed a method for prediction of rTMS treatment outcome by applying KNN classifier to several measures of pretreatment EEG including nonlinear, power spectrum, bispectrum, cordance measures and combination of them. The classification accuracy of 91.3% by EEG power of beta and power of all frequency bands shows that EEG power, especially in beta band, is an appropriate measure to be used in discriminating responders and non-responders. Correlation dimension and some of the measures based on bispectrum amplitude also perform well in the prediction of treatment outcome. The significant results of the proposed classification suggest the potential of applying this method in clinical applications.

Contributors

Author FH contributed in designing the study and acquisition of the data. She wrote the protocol, and was the member of analysis team. She also drafted the manuscript. Author MM conducted the analysis group, revised the manuscript and contributed in designing the study. Author RR contributed in acquisition of data. All authors have contributed in this study and have approved the final manuscript.

Conflict of interest

All the authors declare that they have no conflict of interest.

Role of the funding source

This study has no funding source.

Acknowledgment

We are grateful to Atieh Psychiatry Centre for their kind help in collecting data. Also we appreciate all participants who took part in this study.

Appendix A. Computation of nonlinear and bispectrum measures

A.1. Lempel-Ziv complexity

Lempel-Ziv complexity (LZC) must be applied on symbol sequences with finite length (Aboy et al., 2006; Li et al., 2008). In this way, EEG signals are converted to a binary sequence by comparing with a threshold that is mostly median. For EEG signal {x(n)} a binary sequence [s(n)] = [s(1), s(2), ..., s(n)] will be obtained where

\[ s(i) = \begin{cases} 0, & x(i) < \text{median} \\ 1, & \text{otherwise} \end{cases} \]  

(1)

Then by scanning the EEG sequences [s(n)] from the beginning, different patterns appear in the signal are counted by c(n). It has been proved that the upper limit of complexity measure c(n) with a median threshold is:

\[ \lim_{n \to \infty} c(n) = b(n) = \frac{n}{\log_2(n)} \]  

(2)

The complexity measure c(n) is normalized to b(n), to make the measurement independent of the length of data and LZC will obtain:

\[ \text{LZC} = \frac{c(n)}{b(n)} \]  

(3)
A.2. Katz fractal dimension

One of the known algorithms for computing fractal dimension is KFD (Ahmadlou et al., 2012). This algorithm is based on the assumption that irregularity in a time series appears as changes in the distance of consecutive points. For time series \( x = [x(1), x(2), ..., x(N)] \), the average distance of consecutive points is:

\[
a = \frac{L}{N-1}
\]

where \( N \) is the length of time series and \( L \) is the sum of distances between all successive points that obtained by:

\[
L = \sum_{i=2}^{N} \text{distance}(x(i), x(i-1))
\]

If \( d \) is the maximum distance between \( x(1) \) and other points of time series, \( KFD \) is defined as:

\[
KFD = \frac{\log\left(\frac{1}{a}\right)}{\log\left(\frac{1}{d}\right)}
\]

A.3. Correlation dimension

Correlation dimension (CD) is computed by an algorithm which is based on state space reconstruction by time delay embedding theory (Hosseinifard et al., 2013; Theiler, 1987). Consider the time series \( x = [x(1), x(2), ..., x(N)] \). By applying time delay embedding theory, the state \( i \) of the time series in another space with \( d \) dimensions is defined as:

\[
X_i = (x(i), x(i+\tau), ..., x(i+(d-1)\tau))
\]

where \( \tau \) is the time delay. Then the average probability that the states of the system be closer than a threshold is computed. This average probability which has been named correlation integral is given by:

\[
C(r) = \frac{2}{N(N-1)} \sum_{i\neq j} \theta(r - |X_i - X_j|)
\]

where \( r \) is the threshold or radius of similarity and \( \theta(X) \) is Heaviside step function that is defined as:

\[
\theta(X) = \begin{cases} 
0, & X < 0 \\
1, & X > 0 
\end{cases}
\]

By estimation of the slope of the log-log plot of \( C(r) \) versus \( r \) in the linear region, correlation dimension, \( d \) is obtained by:

\[
d = \lim_{r \to 0} \left[ \frac{\log C(r)}{\log r} \right]
\]

A.4. Bispectrum

For non-Gaussian signal \( x(t) \), the third order cumulant \( \kappa_3(m, n) \) is given by (Mohebbi and Ghassemian, 2012):

\[
\kappa_3(m, n) = E[x(k)x(k+m)x(k+n)]
\]

where \( E[.] \) indicates the expectation operator. If \( X(f) \) is the Fourier transform of \( x(t) \), the bispectrum of the signal \( B(f_1, f_2) \) is defined as the Fourier transform of third order cumulant:

\[
B[f_1, f_2] = E[X(f_1)X^*(f_2)f_1 + f_2]
\]

where * indicates complex conjugate.

Bispectrum is symmetrical in the frequency range of calculation. The non-redundant region of bispectrum plot is where \( f_1 \geq f_2 \geq 0 \) that is depicted in Fig. 5 for a frequency range of \([0, h]\).

Due to the symmetry of bispectrum, the triangular non-redundant region of bispectrum can characterize the whole bispectrum. Thus the measures that are extracted from bispectrum are computed in this region. The formula of measures that are extracted from bispectrum in this study.

![Fig. 5. The non-redundant region of bispectrum plot.](image)
are as follow (The measures are abbreviated as described for Table 2). In all formula the non-redundant region is indicated by Ω.

- The sum of logarithmic amplitudes of diagonal elements in the bispectrum

\[
BispSL = \sum_{k=0}^{n} \log(B(k, f, f))
\]

(13)

where \( B(k, f, f) \) is the bispectrum of the signal in diagonal elements in the region Ω.

- The second-order spectral moment of the amplitudes of diagonal elements in the bispectrum

\[
Bisp2M = \sum_{j=1}^{N} (j - \frac{N}{2}) \log(B(j, f, f))^2 \log(B(f, f))
\]

(14)

where \( N \) is the number of diagonal elements in the region Ω.

- Normalized bispectral entropy

\[
BispEn = -\sum_{n} p_n \log(p_n)
\]

(15)

where

\[
p_n = \frac{B(f, f)}{\sum_{j} B(j, f, f)}
\]

(16)

References


Baskaran, A., 2016. The Comparative Effectiveness of EEG Biomarkers in Antidepressant Response and Illness Prediction in Major Depressive Disorder. Queen’s University (Canada).


Baskaran, A., 2016. The Comparative Effectiveness of EEG Biomarkers in Antidepressant Response and Illness Prediction in Major Depressive Disorder. Queen’s University (Canada).


