34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

A Machine Learning Approach Using P300 Responses to Investigate Effect of Clozapine Therapy

Maryam Ravan, Duncan MacCrimmon, Gary Hasey, James P. Reilly, and Ahmad Khodayari-Rostamabad

Abstract—**Clozapine (CLZ) is uniquely effective as a treatment for medication resistant schizophrenia. Information regarding its mechanism of action may offer clues to the pathophysiology of the disease and to improved treatment. In this study we employ a machine learning (ML) analysis of P300 evoked potentials obtained from quantitative electroencephalography (QEEG) data to identify changes in the brain induced by CLZ treatment. We employ brain source localization (BSL) on the EEG signals to extract source waveforms from specified regions of the brain. A subset of 8 features is selected from a large set of candidate features (consisting of spectral coherences between all identified source waveforms at multiple frequencies) that discriminate (by means of a classifier) between the pre- and post-treatment data for the schizophrenics (SCZ) most responsive to CLZ. We show these same selected features also discriminate between pre-treatment most responsive SCZ and healthy volunteers (HV), but not after treatment. Of note, these same features discriminate the least responsive SCZ from HV both pre- and post-treatment. This analysis suggests that the net beneficial effects of CLZ in SCZ are reflected in a normalization of P300 brain-source generators.**

Keywords- Brain source localization; EEG signals; P300 evoked potentials; machine learning; schizophrenia; clozapine treatment

I. INTRODUCTION

Schizophrenia is a severe psychotic disorder affecting approximately 1 % of the world's population [1]. Among the many drugs used for treatment, CLZ is recognized to have superior therapeutic effectiveness [2]. Thus, studies investigating differences between CLZ and other agents are of particular interest as they may provide clues to both underlying neuro-pathologies and improved treatments.

EEG abnormalities in schizophrenic subjects and EEG changes due to clozapine therapy have been the focus of clinical studies (see e.g., [3], [4]). Event related potentials (ERPs) are components of the EEG signal in response to an applied stimulus. The P300 is an ERP component of particular interest because it is related to psychological parameters of attention/cognition. This component, a positive ERP with a latency of about 300 ms, occurs when a rare task-relevant stimulus is imbedded in a series of similar irrelevant ones [5].

M. Ravan, J. P. Reilly, and A. Khodayari-Rostamabad are with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, L8S 4K1, Canada. emails: mravan@ece.mcmaster.ca, reillyj@mcmaster.ca, and akhodayari@ieee.org.

D. MacCrimmon and G. Hasey and are with Department of Psychiatry and Behavioural Neurosciences, McMaster University, and also with Mood Disorders Program, St. Joseph Hospital, Hamilton, ON, emails: maccrim@mcmaster.ca and ghasey@sympatico.ca.

In this study, we use a BSL method to isolate the source waveforms which may then allow features with salient information related to clozapine treatment to be extracted. One of the most reliable methods to localize the EEG signal is the source montage approach [6]. Source montage analysis assumes that the electrical activity at the electrode locations on the scalp is a linear transformation of distinct cortical source/current activities (or source derivatives) within the brain, and uses an *a priori* model comprising predefined cortical locations (virtual current sources) within a head topography [6], [7], to effectively derive an inverse transformation from the scalp electrical activity to the source derivative activity. Moreover, this procedure uses a spatial filtering process which suppresses overlapping activity from neighbouring sources [7], including the infra slow activity (frequencies of 0.5 Hz or less) that typically involves several scalp electrodes on conventional montages.

In this paper, ML methods are used to select features from the source waveforms, which, when processed by a classifier, can discriminate between the pre-treatment (*before* treatment, or BT) and post-treatment (*after* treatment, or AT) states of subjects most responsive to CLZ. We show these same features also discriminate between BT SCZ and HV, but NOT between AT most-responsive subjects, and HV. The implications of these findings are discussed in the final section of this paper.

II. METHODS

A. EEG Recordings

P300 evoked potentials were collected from 47 chronically ill, treatment-resistant schizophrenic subjects both BT and AT clozapine therapy, as well as from 66 healthy volunteers. Demographic information of the SCZ participants are: age [years]: avg. = 37.3, std = 9.44, min = 22, max = 57, Gender: 29 male subjects (61.7%) and 18 female subjects (38.3%) at start of treatment. The HV demographics are: age [years]: avg. = 37.1, std=15.8, min = 18, max = 74, Gender: 36 male subjects (54.5%) and 30 female subjects (45.5%). 500 trials of P300 data with duration of 1024 ms were collected with a sampling frequency of 250 Hz using a QSI-9500 system. The stimulus for P300 data is a standard 100 ms tone presented binaurally where the frequencies of 1000 Hz and 2000 Hz were used for "common" and "rare" tones, respectively. On average 80% of the stimuli were "common" and 20% were "rare".

The system then separately recorded the average of the "common" and "rare" trials.

Electrodes were placed in the 10/20 configuration referenced to linked ears with impedances below $5\text{k}\Omega$. The signals were band pass filtered between 0.5 and 80 Hz and notch filtered at 60 Hz by the QSI system during the recording. The signals were then digitally band pass filtered after recording between 0.5 and 20 Hz to partially mitigate the effects of eye movement and muscle artifacts.

Clinical rating scores of symptom severity were obtained both BT and AT. Subjects with reduction in severity rating scores from BT to AT of at least 35% were designated "most responsive". Those with severity score reduction of less than 35% were designated "least responsive". This results in 20 most responsive and 27 least responsive subjects.

B. Brain Source Localization

BSL facilitates detection of focal source processes in the brain by suppressing the overlap from different brain regions seen at the scalp. Each output from a brain source montage can be viewed as a gross virtual electrode placed onto a particular brain region.

The BESA 5.1.8 EEG review and analysis program (MEGIS Software, Gräfelfing Germany) was used for filtering and source montage processing [8]. The BESA source montage method identifies source waveforms (montages) from 15 regions: midline fronto-polar (FpM); frontal left (FL), frontal midline (FM) and frontal right (FR); anterior temporal left (TAL) and anterior temporal right (TAR); central left (CL), central midline (CM) and central right (CR); posterior temporal right (TPR) and posterior temporal left (TPL); parietal left (PL), parietal midline (PM) and parietal right (PR); and midline occipito-polar (OpM) areas. Source activity in each region is modeled as a single regional source (containing of three mutually orthogonal dipoles) in that region. Regional sources provide more stable fits than dipoles and can account for any activity in an extended brain region, independent of the current orientations. Therefore the number of brain source montages was $3 \times 15 = 45$ for 15 regions. These brain source montages were calculated using specific weighted combinations from the scalp EEG signals. The weights were optimized to minimize the total output power under the constraint of fixed power from the desired region.

III. Machine Learning Approach

Our first goal was to determine whether CLZ treatment induced changes in recorded P300 signals that differentiate between BT and AT conditions. To this end we developed a ML algorithm which identified P300 features that discriminated between treatment conditions. A brief summary of the ML process used follows. A more detailed explanation of ML in the clinical context is available in [9], [10].

This procedure was based on a training set, which consisted of the pre-treatment and post-treatment P300 waveforms from the 20 subjects out of 47 SCZ subjects who were the most responsive to the clozapine treatment. The ML procedure was developed from this training set, and is summarized as follows. First a large set of features is extracted from the brain source signals. The dimensionality of the problem is reduced by selecting only the most discriminative features, using a feature selection procedure. These selected features are then fed into a classifier that outputs the BT/AT condition. The procedure is validated using a leave-one-out cross validation procedure [11].

A. Feature selection

For this study, the set of candidate features consisted of the spectral coherence function between all 15 extracted source waveforms at various frequencies. The sampling frequency of 250 Hz and the duration of 1024 ms resulted in 256 samples for the source signal of each region. Therefore the one-sided power spectral coherence function has (256/2+1) = 129 samples. There were 105 spectral coherence functions between the 15 regions in each direction. Hence the total number of candidate features was $N_c = 129 \times 105 \times 3 = 40635$. All candidate features were normalized using their corresponding *z*-score value. After normalization the most relevant features were selected using the supervised greedy method of [12]. This procedure was used to reduce the number of candidate features to a set of only $N_r = 8$ most relevant features.

IV. Classification and Performance Evaluation

Let the set of most relevant features for a particular subject i be expressed in a vector $\mathbf{x}_i \in \Re^{N_r}$ and the corresponding class be denoted by $y_i \in [\text{BT}, \text{AT}]$. The resulting set $\{x_i, y_i\}$, $i = 1, 2, ..., M_t$, represents a training set where M_t is the number of training patterns. In this study the value of M_t is 40 consisting of the pre- and post-treatment (BT and AT) data from 20 subjects.

In our study, we used the fuzzy c-mean (FCM) algorithm to implement the classifiers [13]. This algorithm is an iterative classification method having some advantages with respect to other classifiers, the most prominent of which is its high generalization capacity for a reduced number of training trials.

The proposed methodology used two leave one out (LOO) procedures executed in succession. The first LOO procedure was used to select the best N_r features in order to avoid choosing features that are dominant in just a few patterns. In this process, at each iteration a list of the best $2N_r$ features was determined using the feature selection procedure of [12]. After all iterations are complete, the N_r features with the highest number of repetitions (probability of appearance) among the available lists were selected as the final set of selected features. The second LOO was then used to evaluate the performance of the classifier.

V. RESULTS

The first objective was to discriminate between the BT and AT conditions in regard to treatment response. Table I shows the classification performance of the proposed methodology for discrimination between BT and AT CLZ treatment conditions for the 20 most responsive subjects (same subjects BT and AT) when $N_r = 8$ discriminating features are used. As shown in Table I, the percentage of correct identifications is 85% for both BT and AT EEG samples, which leads to the total classification accuracy (TCA) of 85%. The TCA is calculated by the number of correct identifications in two classes divided by the total number of EEGs.

These discriminating features are shown in Table II. Since these features discriminate BT vs. AT in most-responsive subjects, they are changed as a result of CLZ therapy. These features are sorted based on their pair *t*-statistic absolute values. The means, standard deviations, and *t*- statistic values of the BT and AT treatment classes are given in columns 3, 4, and 5, respectively. The *t*-statistic value was defined as:

$$t(i) = \frac{(M_{\mathrm{BT}}(i) - M_{\mathrm{AT}}(i))}{\sqrt{\frac{S_{\mathrm{BT}}(i)}{n_{\mathrm{BT}}} + \frac{S_{\mathrm{AT}}(i)}{n_{\mathrm{AT}}}}}, \quad i = 1, 2, \ldots, 8, \tag{1}$$

where n_{BT} and n_{AT} are the number of BT and AT SCZ subjects (equal to 20 in this case), $M_{\mathrm{BT}}(i)$ and $M_{\mathrm{AT}}(i)$ are the mean values of feature *i* for these two groups, and $S_{\mathrm{BT}}(i)$ and $S_{\mathrm{AT}}(i)$ are the corresponding variances. It is noted that *t*-statistic value gives a rough indication of the relevance of the feature, and in this study is used only for the purpose of ordering the features in this table. It should be noted that only the *joint* characteristics of the features in N_r dimensional space are discriminating, not necessarily individual features on their own. Thus, statistically non-significant individual feature differences between the classes in Table II do not necessarily imply that the feature fails to discriminate.

Fig. 1 shows the clustering behavior of the feature vectors from the BT and AT EEG groups (for the 20 most responsive subjects). This figure was generated by projecting the 8 dimensional feature space onto the two major nonlinear principal components for the 40 data points using the kernelized principle component analysis (KPCA) method with a Gaussian kernel [14]. The KPCA method was used for visualization purpose only. As the figure shows, the two classes are clearly separated. Note that even though excellent performance is demonstrated with this 2-dimensional representation, better overall performance is obtained in the $N_r = 8$ dimensional feature space.

The same set of features and methodology as in the preceding was then used in an attempt to differentiate SCZ from HV subjects using only BT EEG data. In this analysis all 47 SCZ subjects BT and all 66 HV subjects were used. As shown in Table III, the percentages of correct identifications were 89.4% and 80.3% for SCZ and HV subjects, respectively. The TCA was 84.1%. Thus, the same features that discriminated BT from AT EEG samples in treatment responsive cases also discriminated all BT SCZ from HV.

Finally, the same set of features and methodology as in the preceding was then used in an attempt to differentiate SCZ from HV subjects, this time using only AT EEG data and analyzing most responsive and least responsive cases separately. These results are shown in Tables IV and V respectively. As Table IV shows, the percent correct identification leads to a TCA of only 48.8% indicating that the selected features do not separate most responsive SCZ from HV when AT EEG data is used. In contrast, Table V shows a TCA of 84.9% in separating HV from SCZ, when AT EEG data is used for the least responsive AT subjects.

VI. DISCUSSION AND CONCLUSION

A set of P300 derived features has been identified that can discriminate pre-treatment from post-treatment EEG states in a group of SCZ subjects who respond well to CLZ therapy. These same features are also found to discriminate most responsive subjects from HV using BT EEG data, whereas after CLZ treatment they no longer differentiate these same subjects from HV. As these features discriminate all BT from all AT conditions, it follows that these have been affected by CLZ treatment. Moreover the lack of differentiation between SCZ and HV when AT EEG data is used among subjects who responded well to CLZ suggests that the therapeutic effect of CLZ is associated with a shift of P300 brain source activity to more closely resemble that of a normal state.

Our discriminating features are primarily localized over frontal and right temporal regions. Contemporary neuroanatomical studies implicate frontal involvement in SCZ symptomatology [15], while Bolsche et al. [16] report atypical right anterior hemispheric P300 activity in un-medicated SCZ, all of which supports the above idea that the regions we have identified might be normalized by CLZ. Since the number of subjects is not large, our results should be interpreted with caution pending further replication or supportive evidence from other investigative mythologies.

TABLE I. CLASSIFICATION PERFORMANCE DISCRIMINATING BETWEEN BT AND AT IN 20 MOST RESPONSIVE (MR) SUBJECTS WITH SCZ.

Class	Predicted BT (MR)	Predicted AT (MR)	% correct	TCA
BT (MR)	17	3	85%	85%
AT (MR)	3	17	85%	

TABLE II. LIST OF THE $N_r = 8$ SELECTED FEATURES AND THE AVERAGE , STANDARD DEVIATION (STD), AND *T*- STATISTIC VALUES OF EACH FEATURE FOR THE BT AND AT RESPONSIVE GROUPS.

#	Feature	Average (± std) for BT	Average (± std) for AT	*t*-statistic values
1	*C*(FpM, TAR) at *f*=19 Hz	0.70 (± 1.30)	-0.70 (± 0.92)	3.9228
2	*C*(FpM, PR) at *f*=19 Hz	0.60 (± 1.23)	-0.60 (± 0.88)	3.5427
3	*C*(FpM, FM) at *f*=16 Hz	0.55 (± 0.89)	-0.55 (± 1.28)	3.1650
4	*C*(CM, PR) at *f*=14 Hz	0.45 (± 1.10)	-0.30 (± 1.30)	1.9687
5	*C*(FpM, FM) at *f*=5 Hz	0.15 (± 0.81)	-0.30 (± 1.41)	1.2314
6	*C*(FpM, TAR) at *f*=7 Hz	-0.15 (± 1.42)	0.2 (± 1.15)	0.8545
7	*C*(FpM, CM) at *f*=5 Hz	-0.2 (± 1.36)	0.05 (± 1.10)	0.6391
8	*C*(PL, TAR) at *f*=4 Hz	0 (± 1.08)	0.15 (± 1.18)	0.4197

REFERENCES

[1] D. Kelly, "Treatment considerations in women with schizophrenia," *J. Women's Health*, vol. 15, no. 10, pp. 1132-1140, 2006.

[2] A. Essali, N. Al-Haj-Hasan, C. Li, and J. Rathbone, *Clozapine versus typical neuroleptic medication for schizophrenia*. Cochrane Database Syst. Rev., John Wiley and Sons Ltd, no. CD000059, 2009.

[3] N. N. Boutros, C. Arfken, S. Galderisi, J. Warrick, G. Pratt, and W. Iacono, "The status of spectral EEG abnormality as a diagnostic test for schizophrenia," *Schizophr. Res.*, vol. 99, pp. 225–237, 2008.

[4] D. MacCrimmon, D. Brunet, M. Criollo, H. Galin, and J. S. Lawson, "Clozapine augments delta, theta and right frontal EEG alpha power in schizophrenic patients," *ISRN Psychiatry*, in press.

[5] E. Niedermeyer and F. L. Da Silva, *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*. 4th ed., Williams & Wilkins, 1999.

[6] M. Scherg, N. Ille, H. Bornfleth, and P. Berg, "Advanced tools for digital EEG review: virtual source montages, whole-head mapping, correlation, and phase analysis," *J. Clin. Neurophysiol.,* vol. 19, pp. 91-112, 2002.

[7] N. Ille, P. Berg, and M. Scherg, "Artifact correction of the ongoing EEG using spatial filters based on artifact and brain signal topographies," *J. Clin. Neurophysiol.*, vol. 19, pp. 113-124, 2002.

[8] http://www.besa.de

[9] A. Khodayari-Rostamabad, G. M. Hasey, D. J. MacCrimmon, J. P. Reilly, and H. de Bruin, "A pilot study to determine whether machine learning methodologies using pre-treatment electroencephalography can predict the symptomatic response to clozapine therapy," *J. Clin. Neurophysiol.*, vol. 121, pp. 1998-2006, 2010.

[10] M. Ravan, J. P. Reilly, L. J. Trainor, and A. Khodayari-Rostamabad, "A machine learning approach for distinguishing age of infants using auditory evoked potentials," *J. Clin. Neurophysiol.*, vol. 122, pp. 2139–2150, 2011.

[11] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2nd ed. Springer, 2009.

[12] H. Peng, F. Long, and C. Ding, "Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans. Patt. Anal. Machine Intell.*, vol. 27, no. 8, pp. 1226–1238, 2005.

[13] J. C. Bezdek, *Pattern recognition with fuzzy objective function algorithms*. New York: Plenum Press, 1981.

[14] K. R. Müller, S. Mika, G. Rätsch, K. Tsuda, and B. Schölkopf, "An introduction to kernel-based learning algorithms," *IEEE Trans. Neural Net.*, vol. 12, no. 2, pp. 181–201, 2001.

[15] V. M. Goghari, S. R. Sponheim, and A. W. MacDonald, "The functional neuroanatomy of symptom dimensions in schizophrenia: a qualitative and quantitative review of a persistent question,". *Neurosci. Biobehav. Rev.*, vol. 34, no. 2, pp. 468-486, 2010.

[16] F. Bolsche, D. J. MacCrimmon, and S. Kropf, "The effect of laterality of stimulus presentation on auditory P300 topography in schizophrenia," *J. Psychiatry Neurosci.*, vol. 21, no. 2, pp. 83-88, 1996.

TABLE III. CLASSIFICATION PERFORMANCE DISCRIMINAING BT SCZ FROM HV.

Class	Predicted SCZ	Predicted HV	%correct	TCA
SCZ	42	5	89.4%	84.1%
HV	13	53	80.3%	

TABLE IV. CLASSIFICATION PERFORMANCE DISCRIMINATING AT MOST RESPONSIVE (MR) SCZ FROM HV.

Class	Predicted AT (MR)	Predicted HV	%correct	TCA
AT (MR)	9	11	45%	48.8%
HV	33	33	50.0%	

TABLE V. CLASSIFICATION PERFORMANCE DISCRIMINATING AT LEAST RESPONSIVE (LR) SCZ FROM HV.

Class	Predicted AT (LR)	Predicted HV	%correct	TCA
AT (LR)	26	1	96.3%	84.9%
HV	13	53	80.3%	

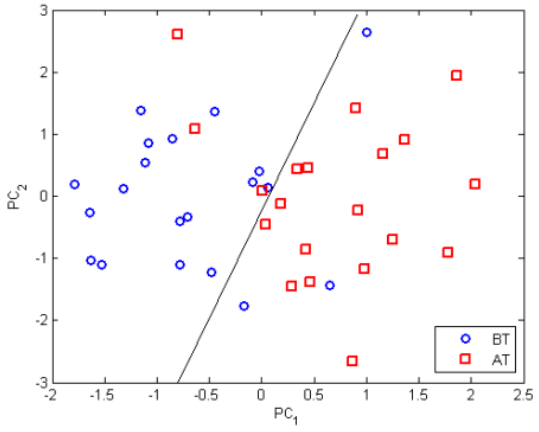

Fig. 1 Subject-wise scatter plot of the N_r dimensional feature space showing BT (blue circles) vs. AT (red squares) of the most-responsive subjects, projected onto the first two major nonlinear principal components.