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

Support vector regression correlates single-sweep evoked brain potentials to gastrointestinal symptoms in diabetes mellitus patients

C. Graversen, J. B. Frøkjær, C. Brock, A. M. Drewes & D. Farina, *Senior Member, IEEE*

Abstract—**Diabetes mellitus (DM) is a multi-factorial and complex disease causing autonomic neuropathy and gastrointestinal symptoms in some patients. The neural mechanisms behind these symptoms are poorly understood, but it is believed that both peripheral and central mechanisms are involved. To gain further knowledge of the central mechanisms, the aim of this study was to identify biomarkers for the altered brain activity in type-1 DM patients compared to healthy volunteers (HV), and to correlate the obtained biomarkers to clinical patient scores. The study included 14 DM patients and 15 HV, with brain activity recorded as multi-channel electroencephalography evoked brain potentials (EPs) elicited by painful electrical stimulations in the esophagus. The single-sweep EPs were decomposed by an optimized discrete wavelet transform (DWT), and averaged for each channel. The DWT features from the DM patients were discriminated from the HV by a support vector machine (SVM) applied in regression mode. For the optimal DWT, the discriminative features were extracted and the SVM regression value representing the overall alteration of the EP was correlated to the clinical scores. A classification performance of 86.2% (*P*=0.01) was obtained by applying a majority voting scheme to the 5 best performing channels. The biomarker was identified as decreased theta band activity. The regression value was correlated to symptoms reported by the patients (*P*=0.04). The methodology is an improvement of the present approach to study central mechanisms in diabetes mellitus, and may provide a future application for a clinical tool to optimize treatment in individual patients.**

I. INTRODUCTION

Diabetes mellitus (DM) is a disease with rising prevalence in the global population, and reduces quality of life for the patients due to autonomic neuropathy [1]. Autonomic neuropathy causes GI dysfunction responsible for GI symptoms such as nausea, bloating, abdominal pain, vomiting, diarrhea, early satiety and constipation. The neural mechanisms behind these symptoms are not completely understood, but it is believed that altered processing of painful afferent input to the central nervous system is involved in the pathogenesis. Consequently, the mechanisms manifest as altered brain activity which may be assessed by electroencephalography (EEG) [2;3].

EEG reflects the neuronal activity in the brain with high temporal resolution, and can be recorded as evoked brain potentials (EPs) to describe the altered central processing to acute pain in patients. The traditional approach with respect to analysis of EPs is to record several sweeps following an external stimulus, and average the sweeps in the time domain to improve the signal-to-noise ratio [4]. However, several studies have demonstrated a limitation in the average process, since important features in the single-sweep EPs have been identified to be time-locked but not phase-locked to the stimulus, and accordingly cancel out during the average procedure [5]. Consequently, an improvement of the present approach would be assessment of single-sweep characteristics, which may identify biomarkers reflecting the clinical symptoms for the patients.

The single-sweep characteristics may be extracted by an optimized discrete wavelet transform (DWT), which has been successfully applied to the average potentials in pain studies [6]. To identify the common alterations in patients diagnosed with DM, the frequency content in all sweeps could be averaged to provide an estimate of the frequency distribution for all recording channels in each subject. These features may be analyzed by a multivariate pattern analysis (MVPA) to classify patients from healthy volunteers (HV). An appropriate choice of classifier would be a support vector machine (SVM), which beside from the categorical output (DM or HV) may be applied in regression mode to assess the overall alteration of the EEG in comparison to the distinct group as illustrated in figure 1 [7]. This regression value could then be correlated to the autonomic and GI symptom scores reported by the patients

To test if MVPA by feature extraction of the single-sweep EP characteristics followed by support vector regression (SVR) could be used to correlate the EEG in the DM patients to the clinical scores, the aims of this study were: 1) to classify 14 type-1 DM patients from 15 age and gender matched HV; 2) to identify the discriminative capacity of the system; and 3) to correlate the EEG alterations described by the regression value to autonomic parameters and GI symptom scores in each individual patient.

Manuscript received March 29, 2012.

C. Graversen is with Mech-Sense, Department of Gastroenterology & Radiology, Aalborg Hospital, DK-9000 Aalborg, Denmark (phone: 0045-26282093; e-mail: cg@mech-sense.com).

J. B. Frøkjær is with Mech-Sense, Department of Radiology, Aalborg Hospital, DK-9000 Aalborg, Denmark (e-mail: jf@mech-sense.com).

C. Brock is with Mech-Sense, Department of Gastroenterology, Aalborg Hospital, DK-9000 Aalborg, Denmark (e-mail: cb@mech-sense.com).

A. M. Drewes is with Mech-Sense, Department of Gastroenterology & Hepatology, Aalborg Hospital, DK-9000 Aalborg, Denmark (e-mail: amd@mech-sense.com).

D. Farina is with Department of Neurorehabilitation Engineering, Bernstein Center for Computational Neuroscience, University Medical Center Göttingen, Georg-August University, Göttingen, Germany (e-mail: dario.farina@bccn.uni-goettingen.de).

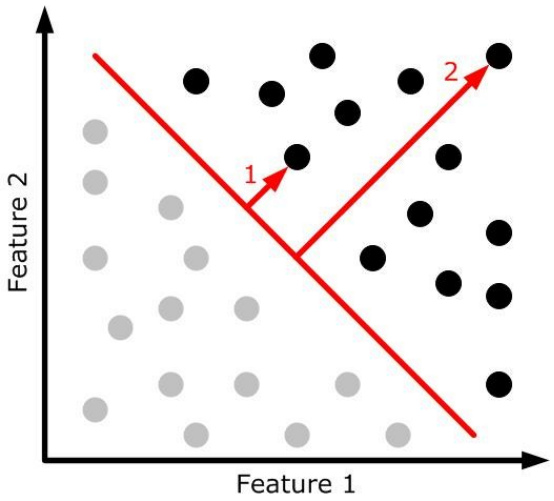

Fig. 1. Support vector regression (SVR) discriminates diabetes mellitus patients (black) from healthy volunteers (grey). Besides from the categorical output, the SVR also outputs the regression value representing the distance to the separating hyperplane as illustrated for two patients.

II. METHODS

A. Study subjects

Fourteen type-1 DM patients (2 males, mean age 34.4, range 20-51 years) and 15 HV (5 males, mean age 33.5, range 21-40 years) completed the study, which conformed to the declaration of Helsinki. Subjects were recruited at Aalborg Hospital, Aarhus University Hospital, Aalborg, Denmark, and Haukeland University Hospital, Bergen, Norway. Patients underwent a general physical examination before study start, and were excluded if they had other pain-related diseases, any prior abdominal surgery or suspicion of psychological abnormalities. All subjects had normal blood samples (except HbA1c level for the patients) including normal creatinine levels.

B. Clinical pain and symptom scores

The clinical pain and symptom scores were obtained as 2 independent values for each patient. Autonomic neuropathy was assessed by scores calculated as a modified Ewing test based on blood pressure and parameters derived from the electrocardiogram [2]. GI symptoms were assessed as the mean of scores (0-4 each) for nausea, bloating, abdominal pain, vomiting, diarrhea, early satiety, and constipation.

C. Experimental protocol

Brain activity was recorded as EPs elicited by painful electrical stimulations in the esophagus. Before the EEG recordings were started, the blood glucose level was adjusted to 6 mmol/l for both patients and HV by a hyperinsulinemic clamp technique [8].

The painful stimulations were conducted by a shielded 70cm long esophageal probe with 2 bipolar platinum ring electrodes. During the recordings, the electrodes were placed in the distal esophagus 32cm from the incisors. Stimulations were delivered at pain threshold with a frequency of 0.2 Hz, and each stimulus consisted of a series of five 1 ms square pulses at 200 Hz, which was felt like one single short pulse in the chest every 5 seconds.

EEG recordings consisted of 50 identical sweeps, recorded with a 62 channel cap (Quick-Cap, Neuroscan, El Paso, TX, USA) and 4 additional electrodes to detect eye movements. Signals were sampled at 1000 Hz and band-pass filtered with cut-off frequencies 0.1 and 200 Hz (SynAmp2, Neuroscan, El Paso, TX, USA). During recordings, the subject rested in a supine position with open eyes in dimmed room light and focused on a fixed point.

The EEG traces were offline notch filtered with cut-off frequencies 49 and 51 Hz and bandpass filtered with cut-off frequencies of 1 and 200 Hz. Data were segmented into epochs starting 50 ms before stimulus onset and ending 500 ms after the stimulus, followed by linear detrending and baseline correction. Channels contaminated by noise or artifacts were interpolated by the neighboring channels, and the 20 best sweeps were selected and saved as one single recording for each subject.

D. Multivariate pattern analysis

For each channel, the single-sweep EPs from 25 to 500 ms post-stimulus were decomposed by an optimized DWT to obtain the power in the delta (0.5 – 4 Hz), theta (4 – 8 Hz), alpha (8 – 16 Hz), beta (16 – 31 Hz) and gamma (31 – 63 Hz) frequency bands as described in detail in a previous study by our group [6]. As the selection of the mother wavelet function is a crucial design parameter for the DWT, coefficients were calculated by 30 different combinations of orthogonal mother wavelet functions in order to determine the optimal solution adapted to the dataset. For each combination, the coefficients for each single-sweep were squared and integrated over time and frequency limits to obtain the power distributions. For each subject, the single-sweep distribution was averaged over all sweeps and normalized to obtain one value in each frequency band.

The normalized band power for all subjects were used as simultaneously input to the SVR in a group analysis, with the classifier trained by leave-one-out (LOU) to optimize the parameters to minimize the probability of error estimated from the training set. System performance for each channel was found by a LOU approach on the optimized system as the ratio of number of correctly classified subjects divided by the total number of subjects, and statistically compared to random classification performance at 50%.

For the best performing channels, a majority voting scheme was applied to obtain the overall system performance, and the biomarkers were identified as the mean of the band power for these channels [9]. Furthermore, the mean of the regression values for the correctly classified patients were correlated to the clinical scores.

E. Statistical analysis

All data are presented as mean±SD unless otherwise indicated. Classification performance was assessed in comparison to random performance at 50% by a Fisher's exact test. The discriminative features were analyzed by a Mann-Whitney rank sum test. The correlation between the regression value and clinical scores was examined by a Pearson product-moment correlation coefficient. All *P* values were two sided, and a *P* value below 0.05 was considered an indication of statistical significance.

III. RESULTS

EEG data were recorded in 14 DM patients and 15 HV. Data were off-line post-processed and the morphology of the average sweeps were validated by comparison to EPs obtained in previous studies of visceral pain inflicted by electrical stimulations. Clinical scores were obtained in 13 out of the 14 DM patients.

The single-sweep frequency power distributions were averaged for each subject and used as independent input to the SVR for each channel. The topographical performance for the optimal wavelet in each channel is displayed in figure 2. The highest performance was 75.9%, which was obtained in the electrodes F4, FC2, FC4, C6, and CP2.

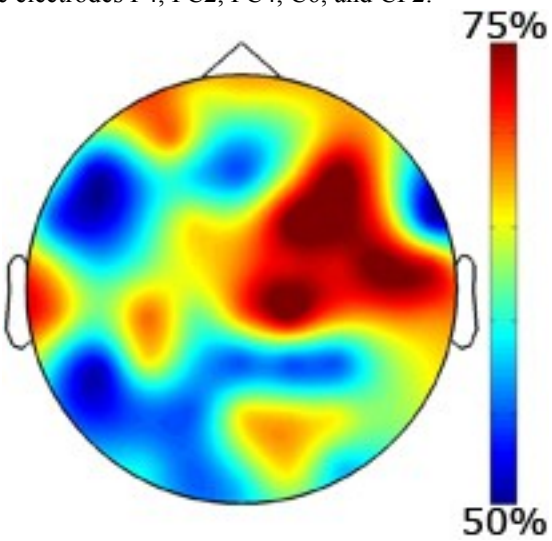

Fig. 2. Topographical classification performance for each of the 62 EEG electrodes. The highest performance of 75.9% was obtained in the F4, FC2, FC4, C6, and CP2 electrodes.

The output verdict for the 5 best performing channels were used as input to the majority voting algorithm, which improved the classification performance to 86.2%, distributed with 12 correctly classified patients and 13 correctly classified HV.

The frequency bands power for each subject were extracted in the 5 best channels for the corresponding optimal wavelet and averaged to one value per band. These features showed a significant decrease in the theta band (*P*=0.006) as presented in figure 3.

For each of the patients classified correctly in the multi-channel scenario, the mean of the corresponding regression values were extracted for the channels providing a correct verdict. The mean regression value was correlated to autonomic and GI symptom scores. As displayed in figure 4, a significant correlation to the symptom score was observed (R=0.715; *P*=0.03) while no correlation was observed with respect to autonomic scores (R=-0.027; *P*=1).

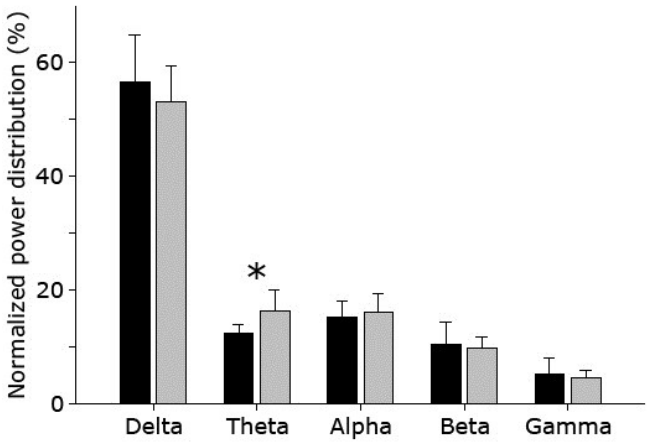

Fig.3. Spectral power distribution for the 5 best performing channels for diabetes mellitus patients (black) and healthy volunteers (grey). An asterisk indicates statistical significant differences between groups.

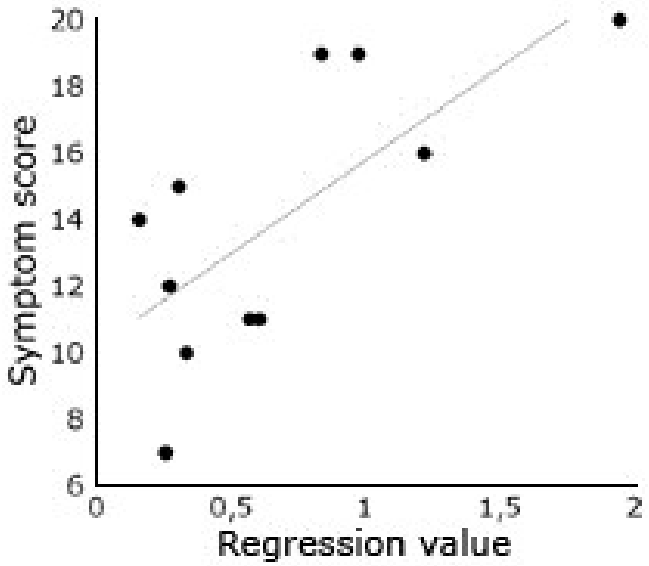

Fig. 4. Correlation between the mean regression value for the best performing channels for correctly classified patients in the multi-channel multivariate pattern analysis and clinical symptom scores reported by each patient (R=0.715; *P*=0.03).

IV. DISCUSSION

In this study we have shown that single-sweep EPs recorded from patients with diabetes mellitus can be discriminated from EPs recorded from healthy volunteers by a support vector machine. Furthermore, by applying the SVM in regression mode, the alterations in frequency band power were assessed as one single score for each patient, which was correlated to symptom scores reported by each patient.

A. Methodological considerations

The multivariate pattern analysis by combined wavelet transform and support vector regression is a novel approach to objectively discriminate DM patients from HV. By this approach, it was possible to extract single-sweep features for both DM patients and HV and correlate findings to clinical scores for the patients. This is an improvement of previous results obtained on the same data, where we have classified the average traces and found differences between patients and HV, but were unable to correlate the biomarkers to clinical scores [10]. Hence, extraction of single-sweep features seems to be superior to feature extraction on average potentials. Furthermore, the regression value from the SVM appear to provide a reliable measure of the overall alteration of the EEG, which we have also demonstrated in other studies, and now for the first time in DM patients [6;11].

B. Interpretation of identified mechanisms

The biomarkers were identified as decreased theta band power in patients compared to HV. This finding is in agreement with previous results obtained from the same data, where the second most discriminative feature was identified as a decrease in the N1-P1 component of the EP, which is primarily described by a theta band oscillation [10]. The best performing channels were in the right temporal region of the scalp. Previous studies have shown that 50% of the EEG amplitudes are generated directly underneath the recording electrode, and up to 95% of the amplitude is caused by generators within a 6 cm radius [12]. Hence, a strong indication of the right insula being responsible for the altered cortical processing in the patients exist, which has been confirmed by an additional study on source localization [13].

C. Clinical implications

The findings obtained in this study have shed new light over the altered central mechanisms in DM patients. The correlation between scalp potentials and the regression value may provide a new automatic and objective tool to assess if central mechanisms contribute to symptoms reported by the patients. Such a tool may in the future provide an application for a clinical tool to optimize individual treatment by selection of pharmaceutical compounds to target the underlying mechanisms.

ACKNOWLEDGEMENTS

The research leading to these results has received funding from the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement no. 223630.

REFERENCES

[1] M. L. Rodrigues and M. E. Motta, "Mechanisms and factors associated with gastrointestinal symptoms in patients with diabetes mellitus" *J. Pediatr. (Rio J.)*, vol. 88, no. 1, pp. 17-24, Jan.2012.

[2] J. B. Frokjaer, E. Softeland, C. Graversen, G. Dimcevski, L. L. Egsgaard, L. Arendt-Nielsen, and A. M. Drewes, "Central processing of gut pain in diabetic patients with gastrointestinal symptoms" *Diabetes Care*, vol. 32, no. 7, pp. 1274-1277, July2009.

[3] J. B. Frokjaer, L. L. Egsgaard, C. Graversen, E. Softeland, G. Dimcevski, R. A. Blauenfeldt, and A. M. Drewes, "Gastrointestinal symptoms in type-1 diabetes: is it all about brain plasticity?" *Eur. J. Pain*, vol. 15, no. 3, pp. 249-257, Mar.2011.

[4] G. Pfurtscheller and F. H. Lopes da Silva, "Event-related EEG/MEG synchronization and desynchronization: basic principles" *Clin. Neurophysiol.*, vol. 110, no. 11, pp. 1842-1857, Nov.1999.

[5] C. Domnick, M. Hauck, K. L. Casey, A. K. Engel, and J. Lorenz, "C-fiber-related EEG-oscillations induced by laser radiant heat stimulation of capsaicin-treated skin" *J. Pain Res.*, vol. 2, pp. 49-56, 2009.

[6] C. Graversen, C. Brock, A. M. Drewes, and D. Farina, "Biomarkers for visceral hypersensitivity identified by classification of electroencephalographic frequency alterations" *J. Neural Eng*, vol. 8, no. 5, p. 056014, Oct.2011.

[7] R. G. Brereton and G. R. Lloyd, "Support vector machines for classification and regression," *Analyst*, vol. 135, no. 2, pp. 230-267, Feb.2010.

[8] R. A. DeFronzo, J. D. Tobin, and R. Andres, "Glucose clamp technique: a method for quantifying insulin secretion and resistance" *Am. J. Physiol*, vol. 237, no. 3, p. E214-E223, Sept.1979.

[9] L. Lam and C. Y. Suen, "Application of majority voting to pattern recognition: an analysis of its behavior and performance" *IEEE Trans. Syst. Man. Cybern.*, pp. 553-568, 1997.

[10] C. Graversen, C. Brock, J. B. Frokjaer, G. Dimcevski, D. Farina, and A. M. Drewes, "Multivariate pattern analysis of evoked brain potentials by temporal matching pursuit and support vector machine" *Submitted. to Neuroimage*, 2012.

[11] C. Graversen, S. S. Olesen, A. E. Olesen, K. Steimle, D. Farina, O. H. Wilder-Smith, S. A. Bouwense, G. H. van, and A. M. Drewes, "The analgesic effect of pregabalin in patients with chronic pain is reflected by changes in pharmaco-EEG spectral indices" *Br. J. Clin. Pharmacol.*, vol. 73, no. 3, pp. 363-372, Mar.2012.

[12] P. Nunez, "Electrical fields of the brain" *Oxford University. Press, New York.*, 1981.

[13] C. Brock, C. Graversen, J. B. Frokjaer, E. Soefteland, M. Valeriani, and A. M. Drewes, "Maladaptive brain changes contribute to gastrointestinal symptoms in diabetes mellitus patients" *Submitted. to Pain*, 2012.