2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Combining Psychophysical and EEG Biomarkers for Improved Observation of Altered Nociceptive Processing in Failed Back Surgery Syndrome

Boudewijn van den Berg, Tom Berfelo, Elisa M.H. Verhoeven, Imre P. Krabbenbos, Jan R. Buitenweg

Abstract—**Diagnosis and stratification of chronic pain patients is difficult due to a lack of sensitive biomarkers for altered nociceptive and pain processing. Recent developments enabled to preferentially stimulate epidermal nerve fibers and simultaneously quantify the psychophysical detection probability and neurophysiological EEG responses. In this work, we study whether using one or a combination of both outcome measures could aid in the observation of altered nociceptive processing in chronic pain. A set of features was extracted from data from a total of 66 measurements on 16 failed back surgery syndrome patients and 17 healthy controls. We assessed how well each feature discriminates both groups. Subsequently, we used a random forest classifier to study whether psychophysical features, EEG features or a combination can improve the classification accuracy. It was found that a classification accuracy of 0.77 can be achieved with psychophysical features, while a classification accuracy of 0.65 was achieved using only EEG features.**

Clinical Relevance—**This study shows which combined features of nociceptive detection behavior and evoked EEG responses are most sensitive and specific to altered nociception in failed back surgery syndrome.**

I. Introduction

Clinical assessment, diagnosis and stratification of chronic pain patients is difficult due to a lack of objective pain biomarkers [1]. Current diagnostic guidelines largely rely on the identification of structural abnormalities using imaging, and assessment of the perceived pain using questionnaires. However, it remains unclear to which extent these structural abnormalities contribute to the perceived pain as these abnormalities are often observed in both painful and non-painful conditions [2,3]. Furthermore, if no structural abnormality could be identified, one is left without objective information about the underlying alterations in nociceptive processing that cause the pain. Development of methods to measure a composition of various sensitive biomarkers to observe altered nociceptive processing could aid clinicians with monitoring and diagnosis of chronic pain patients.

Recently, we developed a method with the aim of improving monitoring of chronic pain patients by measuring the sensitivity to electro-nociceptive stimuli and brain evoked potentials in response to those stimuli [4]. The method uses intra-epidermal electric stimulation around the detection threshold to preferentially active nociceptive afferents in the skin [5,6]. An adaptive procedure is used to continuously center stimulation intensities around the detection threshold. The scalp EEG is measured simultaneously to analyze the brain activity evoked by nociceptive stimulation. The results of this procedure is a large heterogeneous collection of stimulus-response and stimulus-EEG pairs which can be used to effectively assess input-output relations of that patients nociceptive system.

Using this heterogeneous dataset, the effects of stimulus amplitude, stimulus types (i.e., single-pulse and double-pulse stimuli) and the number of administered stimuli (i.e., habituation) on the EEG can be quantified using linear (mixed) models [7]. Simultaneously, effects of the same stimulus properties on the detection probability can be quantified using generalized linear (mixed) models. A recent study [8] used this method to show that the various steps of processing a nociceptive stimulus, including peripheral nerve fiber recruitment, central synaptic summation, and habituation to a repeated stimulus are reflected by these effects of stimulus properties on detection probability and the evoked potential.

As a next step, we want to know whether we could use this method to observe altered nociceptive processing in individual chronic pain patients. We recently started combined sensitivity and brain activity measurements in patients suffering from failed back surgery syndrome (FBSS). In this work, we use individual features of detection probability and EEG obtained in this study to determine if a combination of these features could aid future monitoring and diagnosis of FBSS patients.

II. Method

The experimental procedures described in this paper were approved by the local Medical Review and Ethics Committee.

A. Participants

Psychophysical and brain activity features were extracted from a larger dataset of 16 FBSS patients (9 males; age: 50.1 ± 9.1 years; NRS: 7.0 ± 2.1; CSI: 44.6 ± 13.9) and 17 healthy controls (3 males, age: 35.9 ± 11.9 years; NRS: 0.0 ± 0.0; CSI: 14.6 ± 8.8) measured at the St. Antonius Hospital in Nieuwegein.

B. Procedure

Participants were seated in a comfortable chair and instructed to focus their gaze at a fixed point on the wall. Intra-

*Research supported by the Netherlands Organization for Scientific Research (NWO) and the Anesthesiology R&D department of the St. Antonius Hospital Nieuwegein.

B. van den Berg, T. Berfelo, E. Verhoeven and J.R. Buitenweg are with the department of Biomedical Signals and Systems, Technical Medical Centre, University of Twente, Enschede, the Netherlands (e-mail: b.van.den.berg@utwente.nl).

I. P. Krabbenbos is with the Department of Anesthesiology, Intensive Care and Pain Medicine, St. Antonius Hospital, Nieuwegein, the Netherlands.

epidermal electric stimuli were applied to the back of the hand via a custom made electrode with 5 microneedles [9] and centered around the detection threshold using an adaptive psychophysical procedure [10]. Participants were instructed to release a response button whenever they detected a stimulus. Each stimulus was randomly chosen from an equidistant vector of 5 amplitudes. When a stimulus was reported as detected, all amplitudes were decreased by 0.025 mA. When a stimulus remained undetected, all amplitudes were increased by 0.025 mA. This procedure continued until a total of 450 stimuli was applied (Fig. 1). The total procedure had a duration of 35-45 minutes, and was repeated on each hand.

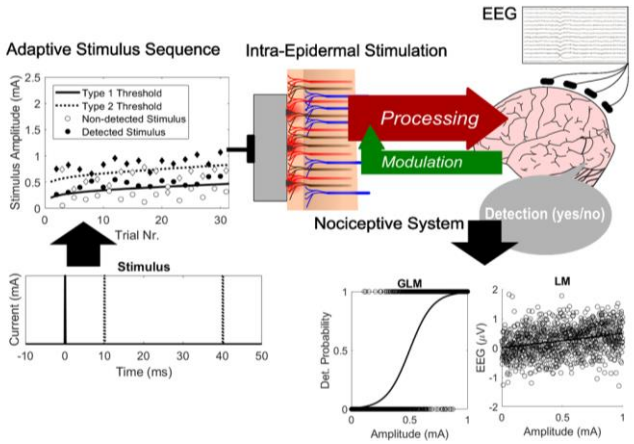

Figure 1. Simultaneous measurement of the nociceptive detection threshold (NDT) and evoked potential (EP), referred to as the NDT-EP method. In this method, the detection probability and threshold of multiple stimulus types (here with one or two pulses, 10 or 40 ms inter-pulse interval) is tracked using an adaptive algorithm while recording EEG. The effect of stimulus properties on the detection probability is quantified using a generalized linear model (GLM) and on the EEG using a linear model (LM). In this way, one might observe altered nociceptive processing (e.g. in FBSS patients) through a change of the relation between stimulus properties, detection probability and brain activity.

D. Psychophysical Features

The 450 stimulus-response pairs obtained during each measurement were used to compute an average detection rate (R_{det}), average response time (RT_{mn}) and the standard deviation of the response time (RT_{std}). The generalized linear model in (1) was fit to the stimulus-response pairs to compute the effects of stimulus properties on the detection probability. The model quantified the effects of amplitude of the first pulse ($PU1$), amplitude of a second pulse with 10 ms inter-pulse interval ($PU2_{10}$), amplitude of a second pulse with 40 ms inter-pulse interval ($PU2_{40}$), trial number (TRL) a model intercept (INT) on the log-odds of stimulus detection. Subsequently, model coefficients were used to compute the average detection thresholds and slopes of single-pulse and double-pulse stimuli with inter-pulse intervals of 10 and 40 ms (T_{SP}, T_{DP10}, T_{DP40}, S_{SP}, S_{DP10} and S_{DP40}).

$$\ln\left(\frac{P}{1-P}\right) \sim 1 + PU1 + PU2_{10} + PU2_{40} + TRL \qquad (1)$$

E. Brain Activity Features

The EEG was recorded at 1000 Hz using a 64-channel Ag/AgCl electrode cap (10-20 system) during the entire experiment. The signal was divided into epochs -0.5 to 1.0 s with respect to stimulus onset and bandpass filtered between 0.1 and 40 Hz using the Fieldtrip toolbox [11] in Matlab. Latencies of the N1 and P2 component of the evoked potential were estimated to be 190 and 440 ms respectively based on the grand average global field power. At both latencies the average and standard deviation of the evoked potential for each stimulus type and overall were computed. The linear model in (2) was fit at both latencies to compute the effects of stimulus properties on the evoked potential. The model quantified the effects of amplitude of the first pulse ($PU1$), amplitude of a second pulse with 10 ms inter-pulse interval ($PU2_{10}$), amplitude of a second pulse with 40 ms inter-pulse interval ($PU2_{40}$), trial number (TRL), stimulus detection (D) and a model intercept (INT) on the evoked potential amplitude.

$$U_{EEG} \sim 1 + PU1 + PU2_{10} + PU2_{40} + TRL * D \qquad (2)$$

F. Random Forest Classification

A random forest classification model was fit separately on all psychophysical features, all brain activity features and all features combined using the 'Scikit-learn' toolbox [12] in Python. Random forests are an accurate classification technique that is able to use complex nonlinear combinations of features for classification and is robust to outliers and noise [13]. As such, this classifier was expected to find an optimal combination of features for classification with an accuracy close to the Bayes rate. The number of estimators for random forest classification was fixed to 1000 estimators. Other parameters were optimized using grid search with 10-fold cross-validation. These optimized parameters included a maximum depth of 3, a minimum number of samples per split of 15 and a minimum number of samples per leaf of 5. Classification performance was evaluated using 10-fold cross-validation in terms of accuracy, sensitivity, specificity, and area under the curve (AUC). We also added a performance metric for consistency, which was defined as the percentage of subjects of which both measurements were assigned the same class. Minimum redundancy maximum relevance (MRMR) feature selection based on the F-test correlation quotient was used to select a subset of all features [14]. Shapley values of the random forest classifier were computed using the 'shap' toolbox [15] to identify the top 6 features of interest in multiple sets of features.

III. Results

A. Psychophysical Features

All psychophysical features extracted from this dataset were found to demonstrate a significant difference between both groups (Fig. 2). The difference between FBSS patients and healthy controls is most clearly demonstrated by the logarithm of psychophysical slopes and detection thresholds (T_{SP}, T_{DP10}, T_{DP40}, S_{SP}, S_{DP10} and S_{DP40}), effects of trial number and amplitude of the second pulse in the psychophysical model (TRL, $PU2_{10}$ and $PU2_{40}$), the detection rate (R_{det}) and standard deviation of the reaction time (RT_{std}) which all were significant with $p<0.001$.

B. EEG Features

Most EEG features did not demonstrate a significant difference between both groups. None of the averages or standard deviations at the N1 or P2 latency were significant. However, fitting a linear mixed model to the EEG data lead to the observation of significant effects by successfully

accounting for the multivariate experimental design (Fig. 3). The difference between FBSS patients and healthy controls is most clearly demonstrated by the effect of the amplitude of the second pulse with either a 10 ms or 40 ms inter-pulse interval ($PU2_{10}$ and $PU2_{40}$) on the P2.

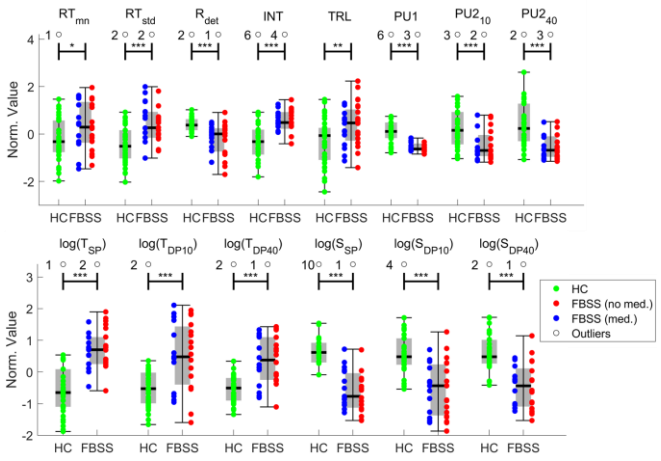

Figure 2. Psychophysical features with a significant difference between FBSS patients and healthy controls (HC). All psychophysical features were found to differ significantly between both groups, independent of used medication.

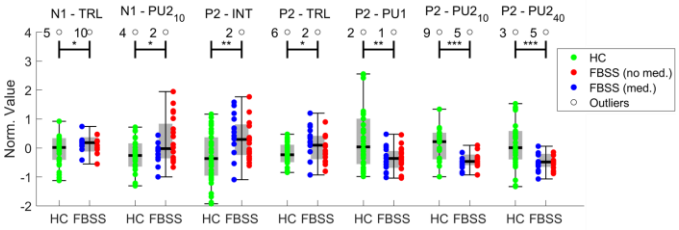

Figure 3. EEG features with a significant difference between FBSS patients and healthy controls (HC). The average and standard deviation of the evoked potential for each stimulus type and overall were not significant, and therefore not included in this figure.

C. *Random Forest Classification*

Assessment of the random forest classification performance using 10-fold cross-validation lead to an overall accuracy of 0.75, with 0.78 sensitivity, 0.72 specificity, 0.82 AUC and 0.72 consistency (Table I). The accuracy improved to 0.78, with 0.78 sensitivity, 0.76 specificity, 0.89 AUC and 0.72 consistency when using only the top 5 features from MRMR feature selection. The accuracy also improved to 0.77, with 0.76 sensitivity, 0.76 specificity, 0.83 AUC and 0.73 consistency when using only the psychophysical features. The accuracy decreased to 0.65, with 0.63 sensitivity, 0.73 specificity, 0.71 AUC and 0.60 consistency when using only EEG features.

An overview of the top 6 Shapley values for each model fit shows the features with the largest contribution to the each predicted class (Fig. 4). The model fit using all features and the model fit using only psychophysical features both use the same psychophysical values as their most important features for classification: the effect of the first pulse amplitude in the psychophysical model ($PU1$), and log-transformed detection thresholds and slopes (T_{SP}, T_{DP10}, T_{DP40}, S_{SP}, S_{DP10} and S_{DP40}). The model fit using only EEG features uses mostly the features obtained by fitting a linear model: the model intercept and effect of the first and second pulse amplitudes on the P2 ($PU1, PU2_{10}, PU2_{40}$), the effect of the second pulse amplitude with 10 ms inter-pulse interval ($PU2_{10}$) on the N1, and the mean P2 amplitude for single-pulse stimuli (μ_{SP}). MRMR feature selection uses a combination of EEG and psychophysical features.

TABLE I. PERFORMANCE OF CLASSIFYING FBSS PATIENTS AND HEALTHY CONTROLS USING ALL, PSYCHOPHYSICAL AND EEG FEATURES. THE LARGEST VALUE OF EACH METRIC IS HIGHLIGHTED IN **BOLD**.

	Features			
	All	***MRMR***	***Psychophysical***	***EEG***
Accuracy	0.75	**0.78**	0.77	0.65
Sensitivity	**0.78**	**0.78**	0.76	0.63
Specificity	0.72	**0.76**	**0.76**	0.73
AUC	0.82	**0.89**	0.83	0.71
Consistency	0.72	0.72	**0.73**	0.60

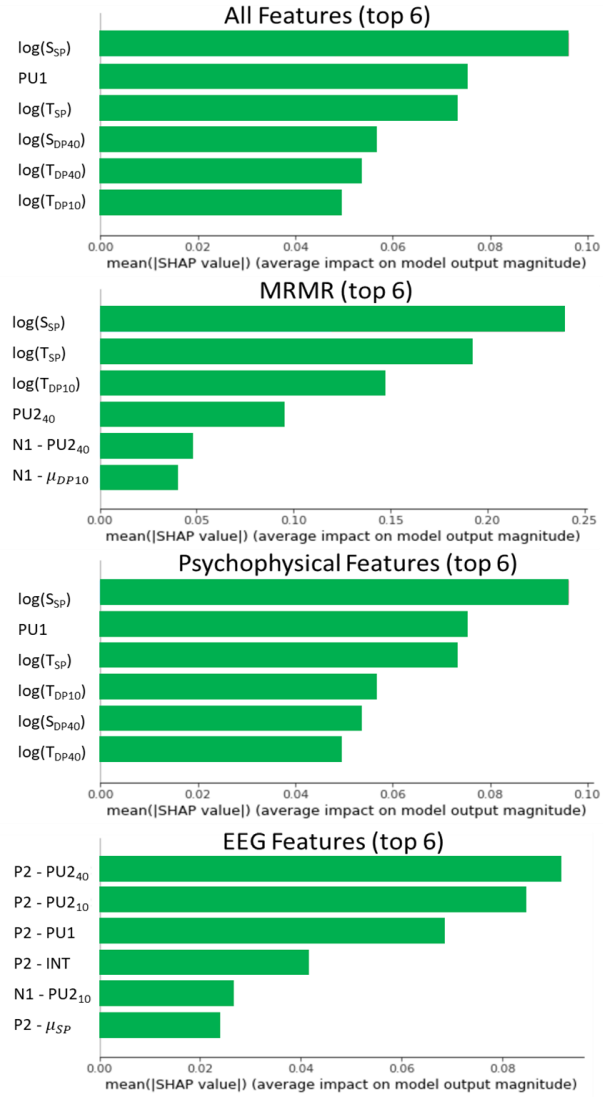

Figure 4. Shapley values of the 6 most important features for random forest classification when using all features (top), the top 6 features based on minimum redundancy maximum relevance (MRMR, middle), only psychophysical features (middle) and only EEG features (bottom).

IV. Discussion

In this work, we extracted individual psychophysical and EEG features of FBSS patients and healthy controls to study whether a combination of these features could aid the observation of altered nociceptive processing in chronic pain patients. Note that the influence of the experimental procedure parameters on outcomes has been discussed previously by Van den Berg & Buitenweg [8], and is therefore not studied here.

It was found that all psychophysical features differed significantly between all FBSS patients, with and without medication, and healthy controls. As such, stimulus detection behavior in FBSS patients, was completely different with respect to healthy controls. Differences included much larger detection thresholds and much lower slopes of the psychometric curve. Furthermore, the patients were characterized by a much lower detection rate and more variation in their reaction times. Each of these results indicates that FBSS patients had more trouble distinguishing nociceptive stimuli. Larger detection thresholds show that a larger stimulus intensity was required to elicit a detectable sensation. Lower psychometric slopes show that when a stimulus was reported as detected, patients were much less certain about their detection. Larger variation in response times might also be explained by the extra difficulty patients experienced in determining whether or not they detected a nociceptive stimulus.

It was also found that some EEG features obtained by fitting a linear model differed significantly between all FBSS patients and healthy controls. However, none of the features obtained by averaging or computing standard deviation of the EEG at a single latency differed significantly between both groups. This demonstrates that the effect of stimulus properties on the response is more important than the average response. As such, it was found that the most significant differences of EEG activity between FBSS patients and healthy controls were at the effect of the pulse amplitudes on the EEG.

Multiple features were combined into one classifier using a random forest model. Using all features resulted in an accuracy of 0.75. Using MRMR feature selection improved the accuracy towards 0.78. Using only psychophysical features resulted in a similar accuracy of 0.77, as well as a similar sensitivity, specificity and consistency. Using only EEG resulted in a much lower accuracy of 0.65, and a lower score on all metrics. This suggests that the information included in the EEG features in this dataset was largely redundant with the information included in the psychophysical features. As such, classification using only the psychophysical features led to a similar accuracy as a classification that also included the informative but noisy EEG features.

V. Conclusion

Combining EEG and psychophysical biomarkers only leads to minor improvements for the classification of FBSS patients with respect to healthy controls. For the sole purpose of monitoring patients, information about the detection thresholds and slopes might be sufficient. Indeed, all psychophysical outcome measures were shown sensitive to FBSS, and some differed between both groups with a high significance ($p<0.001$). We also found in this study that EEG features could be used separately for identification of altered nociceptive processing in the FBSS patients, although with a lower accuracy. As EEG is a measure of brain activity rather than the resulting behavior, this method might prove beneficial in situations where a more objective outcome measure is required. Some of the EEG features differed with a high significance ($p<0.001$) between both groups, and are useful to observe altered nociceptive processing on group level.

References

[1] Mouraux, A., & Iannetti, G. D. “The search for pain biomarkers in the human brain”. *Brain*, 141(12), pp. 3290-3307, 2018. doi: 10.1093/brain/awy281

[2] Jordan, J., Konstantinou, K., & O'Dowd, J. “Herniated lumbar disc”. *BMJ clinical evidence*, 2009, pp. 1118, 2009. Retrieved from https://pubmed.ncbi.nlm.nih.gov/19445754

[3] Van den Bosch, M. A., Hollingworth, W., Kinmonth, A. L., & Dixon, A. K. “Evidence against the use of lumbar spine radiography for low back pain”. *Clin Radiol*, vol. 59, no. 1, pp. 69-76, 2004. doi: 10.1016/j.crad.2003.08.012

[4] Van den Berg, B., Doll, R. J., Mentink, A. L. H., Siebenga, P. S., Groeneveld, G. J., & Buitenweg, J. R. “Simultaneous tracking of psychophysical detection thresholds and evoked potentials to study nociceptive processing”. *Behavior Research Methods*, vol. 52, pp. 1617-1628, 2020. doi: 10.3758/s13428-019-01338-7

[5] Mouraux, A., Iannetti, G. D., & Plaghki, L. Low intensity intra-epidermal electrical stimulation can activate Aδ-nociceptors selectively. *Pain*, vol. 150, no. 1, pp. 199-207, 2010. doi: 10.1016/j.pain.2010.04.026

[6] Poulsen, A. H., Tigerholm, J., Meijs, S., Andersen, O. K., & Mørch, C. D. “Comparison of existing electrode designs for preferential activation of cutaneous nociceptors”. *Journal of Neural Engineering*, vol. 16, no. 3, pp. 36026, 2020. doi: 10.1088/1741-2552/ab85b1

[7] Van den Berg, B., Buitenweg, J. R. "Analysis Of Nociceptive Evoked Potentials During Multi-Stimulus Experiments Using Linear Mixed Models," *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Honolulu, HI, USA, 2018, pp. 3048-3051, doi: 10.1109/EMBC.2018.8513032

[8] Van den Berg, B., Buitenweg, J. R. “Observation of nociceptive processing: effect of intra-epidermal electric stimulus properties on detection probability and evoked potentials”. *Brain Topography*, vol. 34, pp. 139-153, 2021. doi: 10.1007/s10548-020-00816-y

[9] Steenbergen, P., Buitenweg, J. R., Trojan, J., van der Heide, E. M., van den Heuvel, T., Flor, H., & Veltink, P. H. “A system for inducing concurrent tactile and nociceptive sensations at the same site using electrocutaneous stimulation”. *Behavior Research Methods*, vol. 44, no. 4, pp. 924-933, 2012. doi: 10.3758/s13428-012-0216-y

[10] Doll, R. J., Veltink, P. H., & Buitenweg, J. R. “Observation of time-dependent psychophysical functions and accounting for threshold drifts”. *Attention, Perception, and Psychophysics*, vol. 77, no. 4, pp. 1440-1447, 2015. doi: 10.3758/s13414-015-0865-x

[11] Oostenveld, R., Fries, P., Maris, E., & Schoffelen, J. M. “FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data”. *Computational Intelligence and Neuroscience*, 2011. doi: 10.1155/2011/156869

[12] Pedregosa, F., Varoquaux, Gael, Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. “Scikit-learn: Machine learning in Python”. *Journal of Machine Learning Research*, vol. 12, pp. 2825–2830, 2011.

[13] Breiman, L. “Random Forests”. *Machine Learning,* vol. 45, pp. 5–32, 2001. doi: 10.1023/A:1010933404324

[14] Z. Zhao, R. Anand and M. Wang, "Maximum Relevance and Minimum Redundancy Feature Selection Methods for a Marketing Machine Learning Platform," *2019 IEEE International Conference on Data Science and Advanced Analytics (DSAA)*, 2019, pp. 442-452, doi: 10.1109/DSAA.2019.00059.

[15] Lundberg, S.M., Erion, G., Chen, H. et al. “From local explanations to global understanding with explainable AI for trees”. *Nat Mach Intell*, vol. 2, pp. 56–67, 2020. doi: 10.1038/s42256-019-0138-9