Pre-implant Heart Activity Differs in Responders and Non-responders to Vagal Nerve Stimulation Therapy in Epileptic Patients

F. Plesinger, J. Halamek, J. Chladek, P. Jurak, A. Ivora, I. Dolezalova, E. Koritakova, T. Jurkova,
J. Chrastina, M. Brazdil

Abstract—**Vagal Nerve Stimulation (VNS) is used to treat patients with pharmacoresistant epilepsy. However, generally accepted tools to predict VNS response do not exist. Here we examined two heart activity measures – mean RR and pNN50 and their complex behavior during activation in pre-implant measurements. The ECG recordings of 73 patients (38 responders, 36 non-responders) were examined in a 30-sec floating window before (120 sec), during (2x120 sec), and after (120 sec) the hyperventilation by nose and mouth. The VNS response differentiation by pNN50 was significant (min p=0.01) in the hyperventilation by a nose with a noticeable descendant trend in nominal values. The mean RR was significant (p=0.01) in the rest after the hyperventilation by mouth but after an approximately 40-sec delay.**

Clinical Relevance— **Our study shows that pNN50 and mean RR can be used to distinguish between VNS responders and non-responders. However, details of dynamic behavior showed how this ability varies in tested measurement segments.**

I. Introduction

Vagal nerve stimulation (VNS) is a treatment option in drug-resistant epileptic patients. However, it is less effective than resective surgery - about 50% of patients respond to the therapy [1]. A generally accepted approach to distinguish between responders and non-responders before VNS implantation was not established yet.

Although the research in this field is usually focused on EEG markers [2], [3], several parameters related to heart activity were already shown to have a separating ability. These parameters are heart-rate-variability (HRV) features computed from interbeat (RR) intervals. These include the absolute power of the high-frequency band (0.15-0.4 Hz, HF power), a root mean square of RR intervals differences (RMSSD), a percentage of consecutive RR intervals differing > 50 ms (pNN50), or a standard deviation derived from Poincaré plot (SD1/2) and others [4]–[6]. However, our previous work focused on VNS response [7] shown that even a simple mean RR interval, when corrected by a proper baseline, can significantly distinguish between responders and non-responders. This behavior was found before and after the hyperventilation.

In this paper, we further explored different reactions to hyperventilation in responders and non-responders. More specifically, we examined dynamic changes in pNN50 and mean RR before, during, and after the hyperventilation by a nose and mouth.

II. Data

Data were collected in Brno Epilepsy Center (Brno, Czech Republic) as the VNS pre-implantation assessment for the study [3] conducted by St. Anne's University Hospital. The study was approved by the local ethics committee (1G2019); all patients gave their written consent to use their pre-implantation data. We used the Alien Deymed device for EEG recordings with a standard 10-20 electrode setup; at least one ECG channel was recorded during each measurement. Signals were recorded with a 128 Hz sampling frequency.

After the implantation, patients were followed up, and after two or more years, the VNS response was evaluated. If seizures were reduced by more than 50%, a patient was considered a responder. Otherwise, the patient was considered a non-responder. Patients with unclear outcomes were removed from this study, as well as recordings with corrupted ECG channels. The cohort slightly increased compared to our previous study [7] since we examined only a subgroup of measurement segments. Therefore, fewer patients had to be removed due to inconsistency in the measurement.

A. Patient cohort

The used patient cohort consisted of 73 patients; 32 men (age 31 ± 11 years) and 41 women (age 35 ± 13 years). Eighteen male patients were responders and 14 non-responders to the VNS, respectively. Twenty female patients responded to the VNS, while 21 did not.

B. Examined measurement segments

All patients were measured according to the standardized protocol of a length of approximately 20 minutes. The protocol contained multiple rests and several kinds of excitation: the initial rest 1, open/close eyes event, rest 2, photo-stimulation, rest 3 (Rest 3), hyperventilation by a nose (HVN), hyperventilation by mouth (HVM), rest 4 (Rest 4), open/close eyes event and rest until the end of the recording. The lengths of all segments were described in [7] in further detail. Here we focused on the effect of hyperventilation. We used only segments Rest 3, HVN, HVM, and Rest 4. However, while

HVN and HVM were recorded in a consistent duration of 120 seconds, this did not apply at Rest 3 (187 ± 35 sec, ranging from 82 to 304 sec) and Rest 4 (172 ± 42 sec, ranging from 56 to 301 sec). Therefore, for Rest 3, we used the last 120 seconds of the ECG signal preceding HVN. In a total of 4 patients, the 120-sec area interacted with the end of the photo-stimulation. Similarly, we used the first 120 seconds of the ECG signal after HVM for the Rest 4; this area interacted with subsequent open/close eyes events in 10 patients.

The HVN and HVM segments did not follow seamlessly but with a delay of an unequal length of 11 ± 17 sec. Although data points inside this delay were used as a part of 30-sec long floating windows from connected HVN and HVM segments, this segment was not evaluated separately.

III. METHOD

We focused on dynamic behavior related to hyperventilation. Therefore we examined areas before, inside, and after the hyperventilation.

A. Baseline correction

We selected the baseline area as a range <-120,-60> seconds before the start of the HVN. Mean RR and pNN50 values were computed from this baseline area and were stored for each patient as baseline values.

B. pNN50 and mean RR computation

We examined two features related to RR intervals – pNN50 and mean RR. The feature pNN50 describes the percentage of successive RR intervals with more than 50 ms increase in their length. The second feature - mean RR - describes the arithmetic average of interbeat intervals in a given window. Both these features were evaluated in a 30-second-long floating window with a 1-second step. Each segment was examined in 120 steps. Overlapping windows near the segment ends were not cropped and used data from the overlapping area.

We iterated through each segment in a 30-sec floating window, found associated QRS complexes, and computed RR intervals as distances between them. Any elimination or correction of outlying RR intervals was not used. The baseline mean RR value was subtracted from the nominal value.

The same window was used for the pNN50 feature, computed using the function *get_time_domain_features()* from the "hrv-analysis" Python package. The baseline value of pNN50 was also subtracted from the nominal value.

Computed pNN50 and mean RR values were used to obtain respective Grand averages from all patients.

C. Statistical signification

The pNN50 and mean RR values for responders and non-responders for each step (i.e., each second) were collected, and statistical significance was evaluated using the Mann-Whitney U test [8] from the *Scipy.stats* package [9]. Since evaluating consecutive time locations in a signal could be considered multiple testing, we used False-Discovery-Rate (FDR) - Benjamini/Hochberg correction to suppress possible false significance.

We used the *statsmodels* package [10] for the FDR. These corrected results we reported as adjusted p-values. We also evaluated the Area Under the Precision-Recall Curve (AUC) using the *sklearn.metrics* package [11]. All computations were prepared in Python 3.6 and 3.8.

IV. RESULTS

A. Baseline values

We calculated baseline values for pNN50 and mean RR in all 73 patients. Differences between responders and non-responders were not significant at the baseline (Tab.I).

TABLE I. BASELINE VALUES OF EXAMINED FEATURES

Feature	*Responders*	*Non-responders*	*p-value*
pNN50	18.50±21.34	16.80±22.24	0.30
Mean RR (s)	0.86±0.15	0.84±0.18	0.18

B. pNN50 and mean RR results

Tab. II shows the total count of significantly different steps in every 120-second long segment for pNN50 and mean RR. Resultant p-values were adjusted using FDR, and Tab. 2 also shows how many points remained significant after the FDR. Fig.1 shows the detailed behavior of both these features during the experiment.

The pNN50 remained significant only in the HVN segment; however, this applies to nearly 87% of its duration shown as a hatched area in Fig. 1 (the first row, second column). The pNN50 tended to significant differentiation also in the Rest 3 and the Rest 4, but none of these remained significant after the FDR. The best p-value was 0.0006 and 0.01 without and with the FDR correction, respectively. Peak AUC reached 0.72.

Contrary to the pNN50, the mean RR was (FDR) significant in the Rest 4. However, the significant differentiation came with an approximately 40 seconds delay after the Rest 4 start, which is a noticeable point in dynamic behavior (Fig. 1, the second row, hatched area at the right end). The lowest non-adjusted and FDR-adjusted p-value was 0.001 and 0.01, respectively. Peak AUC reached 0.71.

TABLE II. COUNT OF SIGNIFICANTLY DIFFERENT STEPS IN EXAMINED SEGMENTS BEFORE AND AFTER THE FDR ADJUSTMENT

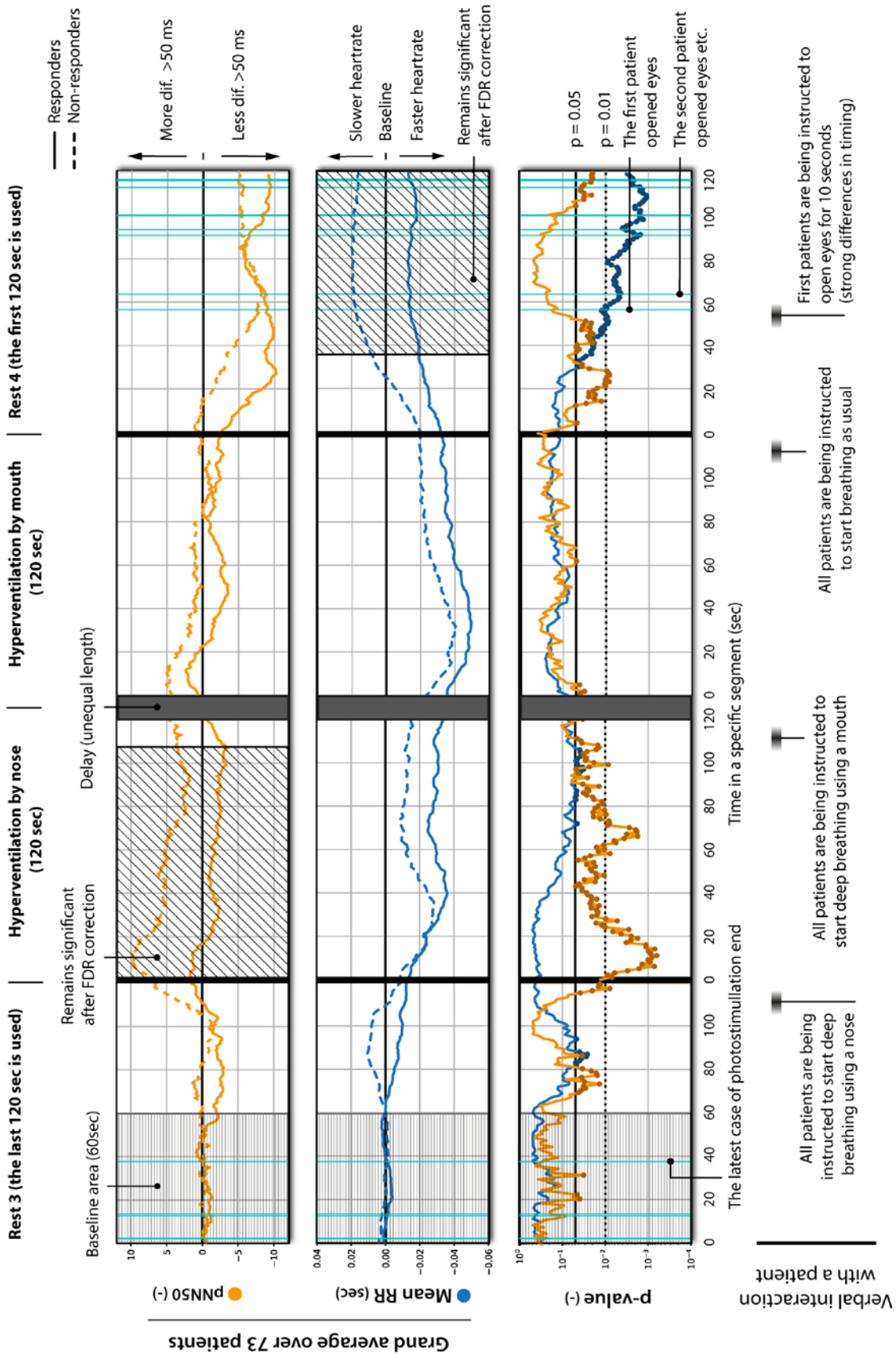| Segment name	pNN50		Mean RR	
	Significant steps		*Significant steps*	
	N	*Adjusted N*	*N*	*Adjusted N*
Rest 3	20	0	8	0
HVN	109	104 (95%)	6	0
HVM	5	0	0	0
Rest 4	50	0	87	84 (97%)

Figure 1. Dynamics of the pNN50 (top row), mean RR (middle row) and their significance (the bottom row). Non-responders (dashed line) show a significantly higher pNN50 in the hyperventilation by nose in comparison to responders (diagonally hatched area). In the resting phase 4, non-responders show significantly longer RR intervals (i.e., slower heartrate) then responders (diagonally hatched area).

V. Discussion

Results showed that both pNN50 and mean RR have a significant ability to distinguish between VNS responders and non-responders. Nevertheless, this behavior applies only to specific time regions.

A. Behavior of the pNN50

The pNN50 was reported before to distinguish between VNS responders and non-responders [6] in 24-hour recordings. The study reported that non-responders have lower pNN50 than responders. Our experiment with much shorter time segments found only one (FDR) significant segment – the HVN. During the HVN, non-responders have significantly higher pNN50 than responders. The pNN50 measured in the baseline area (part of the Rest 3) cannot separate responders and non-responders (Tab. I). VNS responders have higher baseline pNN50 but non-significantly. In the HVN, the pNN50 loses significance approximately 10 seconds before its end. This could be caused by a verbal interaction with a patient (instruction to breathe deeply using a mouth) or by the patient's early action. We also noticed the decreasing trend of pNN50 in both responders and non-responders during the HVN, meaning that pNN50 should not be computed from the whole HVN as a single number. It might be hypothesized that this decreasing trend is caused by decreasing patient effort to hyperventilate. In the HVM, the pNN50 is not significant, and all significant steps in Rest 4 disappeared after the FDR.

B. Behavior of the mean RR

The behavior of mean RR at the baseline was not significant. However, it becomes significant (even with the FDR) in the Rest 4, showing that non-responders have longer mean RR (i.e., slower heart rate) than responders after hyperventilation. This follows our previous findings that the biggest difference in heart rate is between Rest 3 and Rest 4 [7]. The original study used full segment lengths, a slightly smaller cohort, and performed baseline normalization using a division, not a subtraction. The new observation is that Rest 4 becomes significant in mean RR after a specific time, approximately 40 seconds. Since Rest 4 length differs across patients, it is shorter than 120 seconds in 10 cases of 73. Therefore, an ECG signal from consecutive segments (open/close eyes or later) was used. However, because we have already found significant mean RR differences in clean Rest 4 areas [7], we expect that this minor contamination is not essential for mean RR significance. Therefore, presented results showed that even simple mean RR is significant when proper excitation and baseline are used and that it should be evaluated after the proposed time delay.

VI. Conclusion

In this study, we examined the dynamic behavior of the pNN50 and the mean RR in subjects to VNS before implantation. The behavior was examined concerning the hyperventilation by nose and mouth, including preceding and successive rests.

We showed that presented features significantly differ in responders and non-responders during hyperventilation (pNN50) and following rest (mean RR). More importantly, we showed how examined features evolve in time, which was not described before. Observed dynamic behavior means that presented measurement segments should not be evaluated as homogenous blocks. This dynamic behavior also suggests that precise and consistent timing during measurements, even in rests, might be crucial for future predictive models.

References

[1] D. J. Englot, E. F. Chang, and K. I. Auguste, "Vagus nerve stimulation for epilepsy: a meta-analysis of efficacy and predictors of response," *J. Neurosurg.*, vol. 115, no. 6, pp. 1248–1255, Dec. 2011.

[2] C. Bodin *et al.*, "Responders to vagus nerve stimulation (VNS) in refractory epilepsy have reduced interictal cortical synchronicity on scalp EEG," *Epilepsy Res.*, 2015.

[3] M. Brázdil *et al.*, "EEG Reactivity Predicts Individual Efficacy of Vagal Nerve Stimulation in Intractable Epileptics," *Front. Neurol.*, vol. 10, May 2019.

[4] S. Hödl *et al.*, "Neurophysiological investigations of drug resistant epilepsy patients treated with vagus nerve stimulation to differentiate responders from non-responders," *Eur. J. Neurol.*, 2020.

[5] H. Y. Liu *et al.*, "Preoperative Heart Rate Variability as Predictors of Vagus Nerve Stimulation Outcome in Patients with Drug-resistant Epilepsy," *Sci. Rep.*, 2018.

[6] H. Liu, Z. Yang, L. Huang, W. Qu, H. Hao, and L. Li, "Heart-rate variability indices as predictors of the response to vagus nerve stimulation in patients with drug-resistant epilepsy," *Epilepsia*, 2017.

[7] F. Plesinger *et al.*, "Response to Vagal Stimulation by Heart-rate Features in Drug-resistant Epileptic Patients *," in *2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2020, vol. 2020-July, pp. 46–49.

[8] H. B. Mann and D. R. Whitney, "On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other," *Ann. Math. Stat.*, vol. 18, no. 1, pp. 50–60, Mar. 1947.

[9] E. Jones, T. Oliphant, P. Peterson, and Others, "SciPy.org," *SciPy: Open source scientific tools for Python2*, 2001.

[10] S. Seabold and J. Perktold, "Statsmodels: Econometric and Statistical Modeling with Python," in *Proceedings of the 9th Python in Science Conference*, 2010.

[11] F. Pedregosa *et al.*, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, 2011.