



HHS Public Access

Author manuscript

Annu Int Conf IEEE Eng Med Biol Soc. Author manuscript; available in PMC 2021 January 22.

Published in final edited form as:

Annu Int Conf IEEE Eng Med Biol Soc. 2019 July ; 2019: 3502–3506. doi:10.1109/EMBC.2019.8857552.

Ictal autonomic activity recorded via wearable-sensors plus machine learning can discriminate epileptic and psychogenic nonepileptic seizures

A Zsom¹, WC LaFrance Jr^{2,3}, AS Blum³, P Li², LA Wahed², MA Shaikh², G Sharma², R Ranieri², L Zhang³, S Tsekhan³, T Hamid³, J Levin³, W Truccolo^{4,5}

¹Center for Computation and Visualization, Brown University, 3 Davol Square, Providence RI, USA;

²Division of Neuropsychiatry and Behavioral Neurology, Rhode Island Hospital, Brown University, Providence RI, USA;

³Department of Neurology, Rhode Island Hospital, Brown University, Providence RI, USA;

⁴Department of Neuroscience & Carney Institute for Brain Science, Brown University, Providence RI, USA;

⁵U.S. Department of Veterans Affairs, Center for Neurorestoration and Neurotechnology, Providence RI, USA;

Abstract

Differentiating epileptic seizures (ES) and psychogenic nonepileptic seizures (PNES) is commonly based on electroencephalogram and concurrent video recordings (vEEG). Here, we demonstrate that these two types of seizures can be discriminated based on signals related to autonomic nervous system activity recorded via wearable sensors. We used Empatica E4 Wristband sensors worn on both arms in vEEG confirmed seizures, and machine learning methods to train classifiers, specifically, extreme gradient boosting (XGBoost). Classification performance achieved a predictive accuracy of $78 \pm 1.5\%$ on previously unseen data for whether a seizure was epileptic or psychogenic, which is 6 standard deviations above the baseline of 68% accuracy. Our dataset contained altogether 35 seizures from 18 patients out of which 8 patients had 13 convulsive seizures. Prediction of seizure type was based on simple features derived from the segments of autonomic activity measurements (electrodermal activity, body temperature, blood volume pulse, and heart rate) and forearm acceleration. Features related to heart rate and electrodermal activity were ranked as the top predictors in XGBoost classifiers. We found that patients with PNES had a higher ictal heart rate and electrodermal activity than patients with ES. In contrast to existing published studies of mainly convulsive seizures, our classifier focuses on autonomic signals to differentiate convulsive or nonconvulsive semiology ES from PNES. Our results show that autonomic activity recorded via wearable sensors provides promising signals for detection and discrimination of psychogenic and epileptic seizures, but more work is necessary to improve the predictive power of the model.

I. Introduction

Psychogenic nonepileptic seizures (PNES) consist of paroxysmal events with or without alterations in consciousness that resemble epileptic seizures (ES) without the common associated signatures in neocortical, mesial and subcortical epileptiform electrophysiological activity [1], [2]. For this reason, PNES are considered functional neurological or conversion disorders. Misdiagnosis of seizures is a common problem. About one quarter of the patients referred to comprehensive epilepsy centers, for apparent drug-resistant epilepsy, are found to have nonepileptic seizures [3], [4], [5]. Furthermore, approximately 10% of individuals with PNES also have ES [6].

Commonly, laboratory differentiation between ES and PNES is based on examination of concurrent video and electroencephalogram (vEEG) recordings, which is the gold standard for seizure diagnosis [7]. However, vEEG can be costly, time consuming, not always available, and inconclusive in some cases. More recently, there have been attempts to examine the contribution of autonomic nervous system (ANS) activity to both understanding PNES, as well as using it for diagnosis, detection and discrimination [8]. Changes in autonomic system physiology are mediated by the spread of paroxysmal activity to centers involved in autonomic function such as amygdala, hypothalamus, hippocampus, insula, and spinal cord [9]. Autonomic dysfunction is known to be associated with both ES and PNES. For example, autonomic dysfunction can lead to sudden death in patients with epilepsy [10], [11], [12].

Despite this common association, autonomic dysfunction may manifest itself in different ways in ES and PNES. In this study, we examine ANS signals recorded by Empatica E4 Wristband © sensors and use machine learning methods to attempt to discriminate between ES and PNES recorded in the epilepsy monitoring unit. Previous studies examining ANS signals recorded via wearable sensors have reported mostly kinematic biomarkers (e.g. wrist acceleration or surface EMG) limiting analyses to only convulsive seizures (e.g. [13], [14], [15]) or have focused only on epileptic seizure identification [16]. Likewise, studies of heart rate variability (HRV) have distinguished only convulsive ES and PNES [17]. In contrast, our dataset included a mix of convulsive and non-convulsive seizures. We assessed the discrimination power of several ANS related signals, including electrodermal activity (EDA), heart rate (HR), blood volume pressure (BVP) and body temperature. We hypothesized that EDA and HR would distinguish between ES and PNES.

II. Methods

A. Data acquisition and preprocessing

Of 30 participants with inpatient video EEG long-term monitoring, 18 had at least one seizure during the monitoring period of the data collection. The 18 patients had altogether 35 seizures: 10 patients had 23 ES, and 8 patients had 12 PNES. Most of the seizures were measured by two devices - one device worn on each wrist. 13 out of the 35 seizures are convulsive. Seizure type (ES or PNES) and semiology (convulsive or non-convulsive) was

determined by a board-certified neurologist/neurophysiologist (AB). An example of on participant's recordings of a 24 hours epoch is illustrated in Fig. 1.

Data preprocessing included dividing the variable-length seizure periods into fixed-length segments and removing segments if the Empatica device was not in contact with the patients skin. The 35 seizures were divided into 1-minute segments with no temporal overlap. Data from periods when the wrist sensor was not in contact with the patients skin were excluded from our model. Excluded periods were identified when the standard deviation of the BVP was less than 20 or larger than 600 in a segment. If the standard deviation of the BVP was outside the range, we observed the HR to be above 100–120 beats per minute, which is inconsistent with sedentary activity for patients lying in bed, and the HR changed erratically (see the grey region on Fig. 1). The BVP limits were determined by visually inspecting the data.

The final dataset contains 86 ES and 190 PNES segments, each of which are 1 minute long from 16 patients. A larger number of PNES segments were present with the smaller number of PNES because of the longer duration of the PNES. Some PNES lasted more than 10 minutes, while ES were maximum 4–5 minutes long, which is consistent with their typical ictal types.

B. The machine learning pipeline

We generated a set of simple features to illustrate that it is possible to differentiate ES from PNES based on autonomic activity signals. Each 1-minute segment was described by nine summary features: the mean, standard deviation, skewness, minimum, maximum, 25th, 50th, 75th percentiles, and the linear trend of the measurement (slope). As we have five measurements (net acceleration, blood volume pulse, electrodermal activity, heart rate, and body temperature - see Fig. 1), each segment was described by 45 features (nine features calculated for the five measured quantities).

Extra care was taken to ensure that the trained classification model generalizes to data from new patients. We performed three-fold patient-level cross validation (CV). The dataset was divided into training, test and holdout sets based on anonymous patient IDs. The classification model learns on the training set, the test set is used to select the model parameters that maximize the accuracy of the test set, and the holdout set is used to report the accuracy of the final trained model on previously unseen data. We used two CV methods. CV1 collects all segments of one patient in the test set, all segments of another patient are in the holdout set, and the remaining segments from the other patients are in the training set. CV2 collects segments from one patient with psychogenic and one with epileptic seizures in the test set, and segments of another patient with psychogenic and another with epileptic seizures are in the holdout set. The remaining segments of the other patients are in the training set.

XGBoost was used to train the classification model [18]. XGBoost (eXtreme Gradient boosting) is a tree-based classification/regression algorithm. We chose a tree-based method because they generally perform well on small datasets. XGBoost is the most advanced tree-based method to our knowledge and it generates models that are easy to interpret. Several

parameters can be optimized to improve the accuracy of model (learning rate, subsampling parameters, regularization parameters, etc). We tuned the depth of the trees using grid search and explored tree depth values between 2 and 10. The number of trees is another parameter which is tuned by early stopping. In early stopping, the number of trees is increased until the accuracy of the resulting model on the cross-validation set does not improve for some time (a parameter described by `early_stopping_rounds` in XGBoost). Other parameters were unexplored: `learning_rate` = 0.03, `subsample` = 0.66, `colsample_by_tree` = 0.9.

We looped through all possible ways the dataset can be split into the three folds and repeated the process 10 times with different random seeds to calculate uncertainty estimates. There were 240 different combinations to split our dataset into three folds with CV1, and 3024 combinations with CV2. We looped through each possible combination and accumulated the predicted and true labels of the holdout set. The accumulated information was used to calculate the accuracy score and the confusion matrix. The uncertainty estimate was calculated by repeating this whole process 10 times. Even-though the exact same dataset was used in each iteration, tree-based methods have an inherent randomness and this uncertainty is what we assessed with our pipeline. We developed the machine learning pipeline in python 3.6 using packages like scikit-learn, numpy, pandas, and xgboost.

III. Results

A. Classification accuracy

The predictive power of a classifier was assessed based on the confusion matrix and derived metrics. The confusion matrix (C) is defined such that $C_{i,j}$ is equal to the observations that are known to be in group i but are predicted to be in group j . Thus the $i=j$ elements are the correctly predicted observations (true positives and true negatives, TP and TN respectively) and the $i \neq j$ elements are the misclassified observations. It is not obvious in our case what should be called a false positive (FP or false alarm) and a false negative (FN or miss). The reason is that we do not distinguish between a positive and a negative condition (i.e., no seizure or seizure) rather differentiate between two seizure types. As FP and FN cannot be unambiguously defined, we report the accuracy metric of the classifier ($(TP + TN) / (n_{\text{samples}})$) and the confusion matrix.

The classification accuracy was 78% which is at least 6 standard deviations above the baseline accuracy (see Table I). The baseline accuracy (the accuracy if all segments in the holdout set are predicted to be PNES) is 68% with CV1 and 74% with CV2. The standard deviation of the accuracy is 1.5% with CV1 and 0.5% with CV2. The difference is due to the larger number of combinations and more accumulated data in the CV2 pipeline where we select data from two patients in the test set and data from two patients in the holdout set.

B. Predictive features

Feature importance measures helped us identify that the most predictive features are derived from the HR and EDA measurements. Feature importance measures how useful/valuable a feature is in constructing the decision trees. We calculated the feature importances in both

models and the HR and EDA emerged as the two most predictive features of differentiating ES from PNES.

We plot the histograms of the two most predictive features during interictal and ictal periods (see Figs. 2 and 3) to understand their clinical relevance, which indicate how the autonomic functions differ during ES and PNES. As the interictal periods are much longer than the ictal periods, the interictal distributions are much better constrained ($\approx 10^4$ interictal data points vs 276 ictal data points.) We find that the minimum HR measured within the 1-minute segments is the most predictive feature. The interictal HR of ES and PNES patients are 68 ± 11 BPM and 72 ± 10 BPM, respectively. The ictal HR of ES and PNES patients are 73 ± 14 BPM and 94 ± 12 BPM, respectively. We performed two two-sample two-sided K-S tests on the HR data. First, we compared the interictal vs ictal HRs of ES patients and found that the null hypothesis that the distributions of the two samples are the same cannot be rejected at a 1% level ($p = 0.018$). However, when we compared the interictal vs ictal HRs of PNES patients ($p = 3 \times 10^{-82}$), we can confidently reject the null hypothesis. Thus the heart rates of ES patients during the ictal and interictal periods are similar, while PNES patients have a higher ictal heart rate. The second most important feature is the minimum EDA. The interictal EDA of ES and PNES patients are 0.71 ± 1.5 and 1.3 ± 2.4 , respectively. The ictal EDA of ES and PNES patients are 0.33 ± 0.6 and 3.5 ± 3.6 , respectively. The K-S test of the interictal vs ictal EDA of ES patients was not significant ($p = 0.09$, null hypothesis cannot be rejected at a 5% level), while the interictal vs ictal EDA of PNES patients is significant ($p = 2 \times 10^{-32}$, the null hypothesis is confidently rejected). These findings indicate that the EDA of ES patients are similar during the interictal and ictal periods, but PNES patients have a high EDA during seizures.

Limitations of this pilot study are the sample size, and collection from a single site. Novel aspects of the study include comparing multiple ANS signals to differentiate convulsive and nonconvulsive ES from PNES. The preliminary findings reflect a robust signal for differentiating ES from PNES, and can be strengthened with a multi-site large sample trial.

IV. Summary and outlook

Our results indicate that ANS signals recorded via wrist-worn wearable sensors provide promising biomarkers for differentiating vEEG confirmed ES and PNES, both convulsive and nonconvulsive. Among the ANS signals, statistical features related to HR and the EDA were ranked as the top discriminating features. In particular, the ictal HR during PNES appears to be higher than in ictal ES. Other studies have previously examined HR differences in ES and PNES using different sensors and statistical methods [19], [17], [20]. In contrast to our findings, [19] has found higher HR changes in ES ictal events than in PNES, using HR measures relative to baseline references, while we have used absolute rates. Other studies have found prominent peri-ictal increases in EDA in association with ES [21] but comparative studies in PNES versus ES have not been previously reported. We find that EDA responses are higher in PNES than in ES in our cohort. The results of this pilot ictal physiology biomarker study are promising and warrant further data collection and more sophisticated feature-engineering in future studies.

ACKNOWLEDGMENT

We thank the participants in this study, Anita Curran and Preston Linson-Gentry for study setup, and collaborators at Empatica, in particular Chiara Carboni, for helping with data transfer and data annotation, and Francesco Onorati for data consultation. We specifically thank Roz Picard (MIT) who helped create the Empatica device. The wristbands were provided by Empatica.

* This research was supported by a Brown Institute for Brain Science/Norman Prince Neuroscience Institute Frontiers pilot grant (to WCL, AB, WT); the National Institute of Neurological Disorders and Stroke (NINDS), grant R01NS079533 (to WT); the U.S. Department of Veterans Affairs, Merit Review Award I01RX000668 (to WT); and the Pablo J. Salame 88 Goldman Sachs endowed Assistant Professorship of Computational Neuroscience (to WT). The contents do not represent the views of the U.S. Department of Veterans Affairs or the United States Government.

References

- [1]. Devinsky O, Gazzola D, and LaFrance WC Jr, "Differentiating between nonepileptic and epileptic seizures," *Nature Reviews Neurology*, vol. 7, no. 4, p. 210, 2011. [PubMed: 21386814]
- [2]. LaFrance WC Jr, Reuber M, and Goldstein LH, "Management of psychogenic nonepileptic seizures," *Epilepsia*, vol. 54, no. s1, pp. 53–67, 2013. [PubMed: 23458467]
- [3]. Smith D, Defalla B, and Chadwick D, "The misdiagnosis of epilepsy and the management of refractory epilepsy in a specialist clinic," *Qjm*, vol. 92, no. 1, pp. 15–23, 1999. [PubMed: 10209668]
- [4]. Reuber M, Fernandez G, Bauer J, Helmstaedter C, and Elger CE, "Diagnostic delay in psychogenic nonepileptic seizures," *Neurology*, vol. 58, no. 3, pp. 493–495, 2002. [PubMed: 11839862]
- [5]. LaFrance WC Jr, Baird GL, Barry JJ, Blum AS, Webb AF, Keitner GI, Machan JT, Miller I, and Szaflarski JP, "Multicenter pilot treatment trial for psychogenic nonepileptic seizures: a randomized clinical trial," *JAMA psychiatry*, vol. 71, no. 9, pp. 997–1005, 2014. [PubMed: 24989152]
- [6]. Benbadis SR, Agrawal V, and Tatum WO, "How many patients with psychogenic nonepileptic seizures also have epilepsy?" *Neurology*, vol. 57, no. 5, pp. 915–917, 2001. [PubMed: 11552032]
- [7]. LaFrance WC Jr., Baker GA, Duncan R, Goldstein LH, and Reuber M, "Minimum requirements for the diagnosis of psychogenic nonepileptic seizures: A staged approach," *Epilepsia*, vol. 54, no. 11, pp. 2005–2018, 2013. [PubMed: 24111933]
- [8]. Perez DL, Dworetzky BA, Dickerson BC, Leung L, Cohn R, Baslet G, and Silbersweig DA, "An integrative neurocircuit perspective on psychogenic nonepileptic seizures and functional movement disorders: neural functional unawareness," *Clinical EEG and neuroscience*, vol. 46, no. 1, pp. 4–15, 2015. [PubMed: 25432161]
- [9]. Osorio I, "Automated seizure detection using ekg," *International journal of neural systems*, vol. 24, no. 02, p. 1450001, 2014. [PubMed: 24475899]
- [10]. Nashef L, Walker F, Allen P, Sander J, Shorvon S, and Fish D, "Apnoea and bradycardia during epileptic seizures: relation to sudden death in epilepsy." *Journal of Neurology, Neurosurgery & Psychiatry*, vol. 60, no. 3, pp. 297–300, 1996.
- [11]. Blum AS, Ives JR, Goldberger AL, Al-Aweel IC, Krishnamurthy KB, Drislane FW, and Schomer DL, "Oxygen desaturations triggered by partial seizures: Implications for cardiopulmonary instability in epilepsy," *Epilepsia*, vol. 41, no. 5, pp. 536–541, 2000. [PubMed: 10802758]
- [12]. Leutmezer F, Scherthaner C, Lurger S, Pötzelberger K, and Baumgartner C, "Electrocardiographic changes at the onset of epileptic seizures," *Epilepsia*, vol. 44, no. 3, pp. 348–354, 2003. [PubMed: 12614390]
- [13]. Poh M-Z, Loddenkemper T, Reinsberger C, Swenson NC, Goyal S, Sabtala MC, Madsen JR, and Picard RW, "Convulsive seizure detection using a wrist-worn electrodermal activity and accelerometry biosensor," *Epilepsia*, vol. 53, no. 5, pp. e93–e97, 2012. [PubMed: 22432935]
- [14]. Halford JJ, Sperling MR, Nair DR, Dlugos DJ, Tatum WO, Harvey J, French JA, Pollard JR, Faight E, Noe KH, Henry TR, Jetter GM, Lie OV, Morgan LC, Girouard MR, Cardenas DP, Whitmire LE, and Cavazos JE, "Detection of generalized tonic-clonic seizures using surface

- electromyographic monitoring,” *Epilepsia*, vol. 58, no. 11, pp. 1861–1869, 2017. [PubMed: 28980702]
- [15]. Kusmakar S, Karmakar C, Yan B, Muthuganapathy R, Kwan P, O’Brien TJ, and Palaniswami MS, “Novel features for capturing temporal variations of rhythmic limb movement to distinguish convulsive epileptic and psychogenic nonepileptic seizures,” *Epilepsia*, 2018.
- [16]. Onorati F, Regalia G, Caborni C, Migliorini M, Bender D, Poh M-Z, Frazier C, Kovitch Thropp E, Mynatt ED, Bidwell J et al., “Multicenter clinical assessment of improved wearable multimodal convulsive seizure detectors,” *Epilepsia*, vol. 58, no. 11, pp. 1870–1879, 2017. [PubMed: 28980315]
- [17]. Reinsberger C, Perez DL, Murphy MM, and Dworetzky BA, “Pre-and postictal, not ictal, heart rate distinguishes complex partial and psychogenic nonepileptic seizures,” *Epilepsy & Behavior*, vol. 23, no. 1, pp. 68–70, 2012. [PubMed: 22100065]
- [18]. Chen T and Guestrin C, “Xgboost: A scalable tree boosting system,” in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*. ACM, 2016, pp. 785–794.
- [19]. Opherck C and Hirsch LJ, “Ictal heart rate differentiates epileptic from non-epileptic seizures,” *Neurology*, vol. 58, no. 4, pp. 636–638, 2002. [PubMed: 11865145]
- [20]. Jeppesen J, Beniczky S, Johansen P, Sidenius P, and Fuglsang-Frederiksen A, “Comparing maximum autonomic activity of psychogenic non-epileptic seizures and epileptic seizures using heart rate variability,” *Seizure-European Journal of Epilepsy*, vol. 37, pp. 13–19, 2016. [PubMed: 26921481]
- [21]. Picard RW, Migliorini M, Caborni C, Onorati F, Regalia G, Friedman D, and Devinsky O, “Wrist sensor reveals sympathetic hyperactivity and hypoventilation before probable sudep,” *Neurology*, vol. 89, no. 6, pp. 633–635, 2017. [PubMed: 28701502]

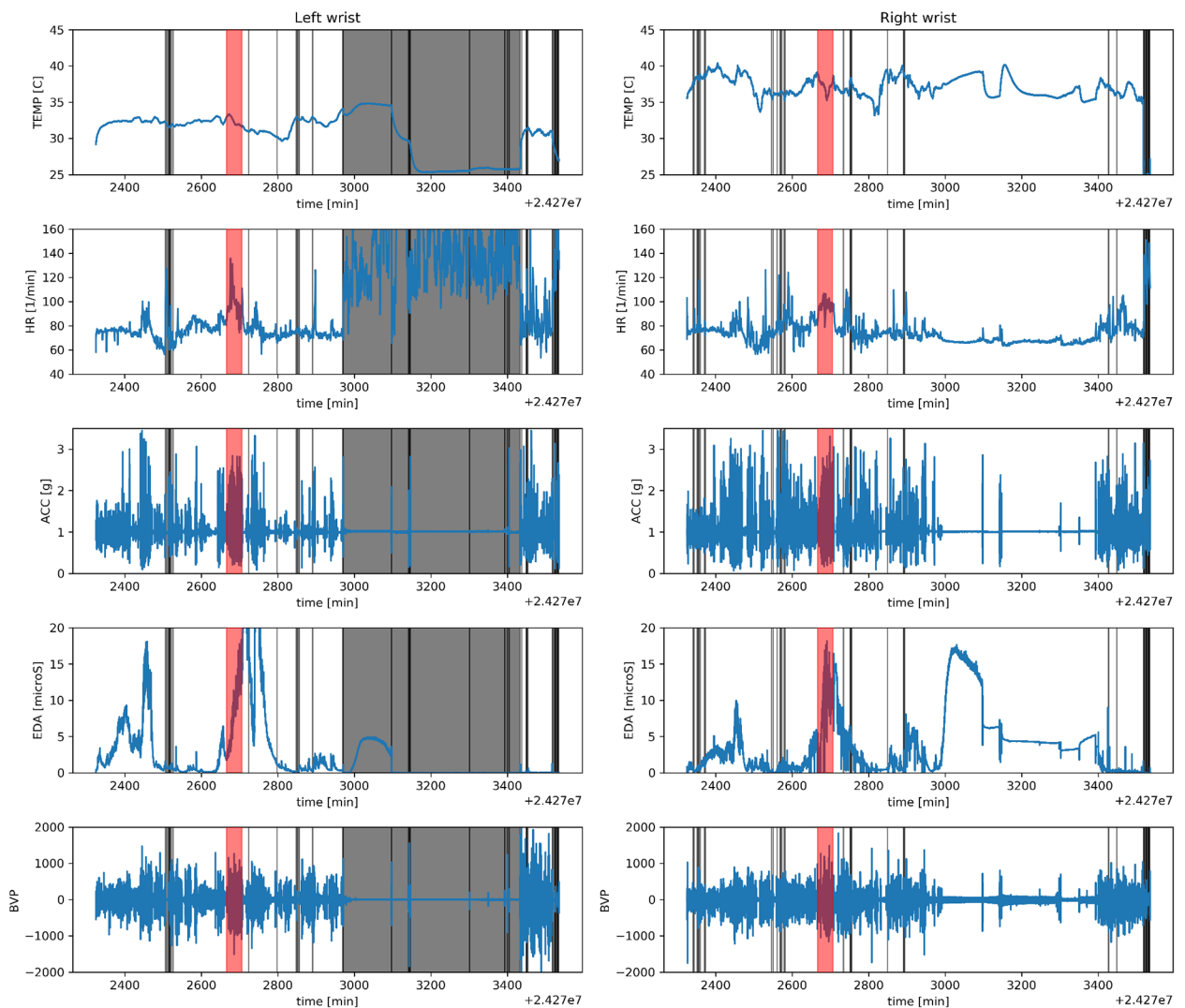
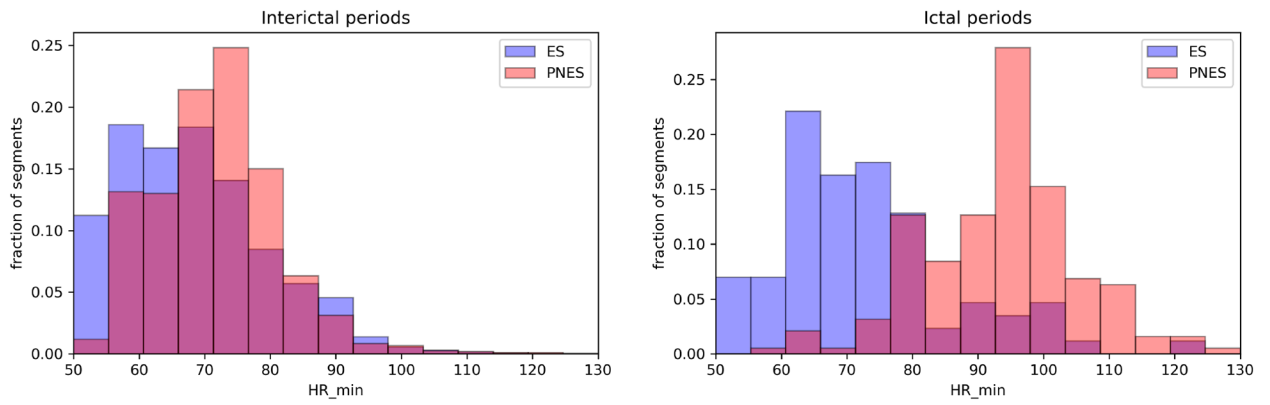


Fig. 1. Example of the Empatica data of one patients session covering roughly 24 hours. The skin temperature (TEMP), heart rate (HR), the net acceleration (ACC), electrodermal activity (EDA) and blood volume pulse (BVP) were measured by the wrist-worn sensor. This patient wore two sensors (one on each wrist) and the left/right side shows the measurements of the left/right wrist device, respectively. The shaded red regions show the time of a seizure, the ictal period. The shaded grey regions show time intervals when the sensor was not in proper contact with the skin of the patient based on the BVP measurements and were not included in the machine learning analysis. Only the ictal period is used in our analysis.

**Fig. 2.**

The most predictive feature in our classification model is the minimum heart rate. The left figure shows the ES and PNES HR histograms during the interictal period. The figure on the right shows the same histograms during the ictal periods. The ES histogram is shown in blue, the PNES histogram is red, the purple area is overlap between the two distributions. The y axis is the fraction of segments present in a bin (1 equals 100%). The heart rate of PNES patients is higher during the ictal period compared to the interictal period and the ictal HR of ES patients.

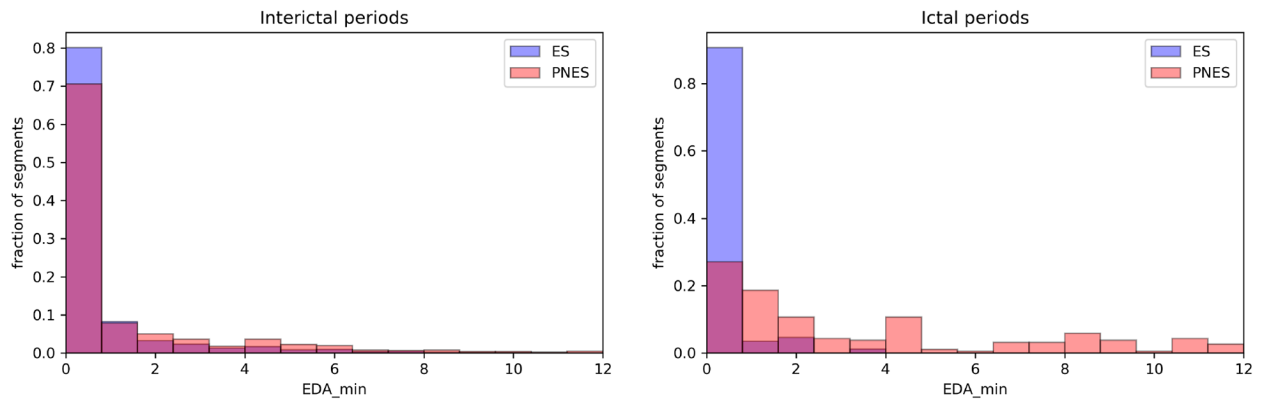


Fig. 3.

The second most predictive feature is the minimum electrodermal activity measured during the segments. Similarly to Fig. 2, the left(right) side shows the interictal(ictal) histograms. The EDA of PNES patients is higher during the ictal period compared to the interictal period and the ictal EDA of ES patients.

TABLE I

Confusion matrices of the holdout set. The values are in percentage points. The complete dataset was analysed with the CV1 and CV2 patient-level cross validation approaches described in Sect. II-B. All possible combinations were explored for the uncertainty estimates.

		Predicted label	
		ES	PNES
CV1, 240 combinations, $78 \pm 1.5\%$ accuracy with a balance of 68%.			
True label	ES	$23.1 \pm 0.4\%$	$8.0 \pm 0.4\%$
	PNES	$14.4 \pm 1.3\%$	$54.5 \pm 1.3\%$
CV2, 3024 combinations, $78.6 \pm 0.5\%$ accuracy with a balance of 74%.			
		Predicted label	
		ES	PNES
True label	ES	$19.3 \pm 0.1 \%$	$6.7 \pm 0.1 \%$
	PNES	$14.6 \pm 0.05 \%$	$59.3 \pm 0.05 \%$