

Machine learning from wristband sensor data for wearable, noninvasive seizure forecasting

Christian Meisel^{1,2,3}  | Rima El Atrache³  | Michele Jackson³ | Sarah Schubach³ | Claire Ufongene³ | Tobias Loddenkemper³

¹Department of Neurology, Charité–Universitätsmedizin Berlin, Berlin, Germany

²Berlin Institute of Health, Berlin, Germany

³Boston Children's Hospital, Boston, MA, USA

Correspondence

Christian Meisel, Department of Neurology, Charité – Universitätsmedizin Berlin, Berlin 10117, Germany.

Email: christian@meisel.de

Funding information

Epilepsy Research Foundation; Brain and Behavior Research Foundation

Abstract

Objective: Seizure forecasting may provide patients with timely warnings to adapt their daily activities and help clinicians deliver more objective, personalized treatments. Although recent work has convincingly demonstrated that seizure risk assessment is in principle possible, these early approaches relied largely on complex, often invasive setups including intracranial electrocorticography, implanted devices, and multichannel electroencephalography, and required patient-specific adaptation or learning to perform optimally, all of which limit translation to broad clinical application. To facilitate broader adaptation of seizure forecasting in clinical practice, non-invasive, easily applicable techniques that reliably assess seizure risk without much prior tuning are crucial. Wristbands that continuously record physiological parameters, including electrodermal activity, body temperature, blood volume pulse, and actigraphy, may afford monitoring of autonomous nervous system function and movement relevant for such a task, hence minimizing potential complications associated with invasive monitoring and avoiding stigma associated with bulky external monitoring devices on the head.

Methods: Here, we applied deep learning on multimodal wristband sensor data from 69 patients with epilepsy (total duration > 2311 hours, 452 seizures) to assess its capability to forecast seizures in a statistically significant way.

Results: Using a leave-one-subject-out cross-validation approach, we identified better-than-chance predictability in 43% of the patients. Time-matched seizure surrogate data analyses indicated forecasting not to be driven simply by time of day or vigilance state. Prediction performance peaked when all sensor modalities were used, and did not differ between generalized and focal seizure types, but generally increased with the size of the training dataset, indicating potential further improvement with larger datasets in the future.

Significance: Collectively, these results show that statistically significant seizure risk assessments are feasible from easy-to-use, noninvasive wearable devices without the need of patient-specific training or parameter optimization.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2020 The Authors. *Epilepsia* published by Wiley Periodicals LLC on behalf of International League Against Epilepsy

KEYWORDS

precision medicine, seizure forecasting, wearable devices

1 | INTRODUCTION

Reliable methods to assess seizure risk could alleviate a major burden for epilepsy patients by providing timely warning or relief when seizure risk is high or low. From a clinician perspective, robust seizure risk assessments are desirable because of their ability to improve treatment by optimizing dosing and timing of antiseizure medication regimen, utilizing objective, personalized standards, as well as by potentially enabling timely interventions to avert impending seizures.¹ Following initial attempts,² there has been a recent surge of studies demonstrating the possibility of accurate seizure forecasting.^{3–7} To this end, most studies have utilized either electrocorticography (ECoG) or scalp electroencephalography (EEG) as well as, to a lesser extent, electrocardiography (ECG), and have demonstrated that robust differentiation between preictal and interictal periods as well as early seizure detection is possible with a better-than-chance performance.^{8–13} Furthermore, seizure forecasting has traditionally performed best when algorithms were trained or optimized at the individual patient level, which often required some sort of training or adjustment phase^{8,14} prior to deployment.

To make seizure risk assessments available for broader clinical use, however, methods that build on noninvasive, easily recordable data streams and that can be readily used without the need of an adjustment phase or expert parameter setting are desirable.¹⁵ Peripheral signals recorded using wearable devices, such as wristbands, are particularly interesting in this respect, because these signals permit continuous, noninvasive recording of several physiological parameters, such as electrodermal activity, body temperature, blood volume pulse, and actigraphy.¹⁶ At the same time, the compact design may limit the risk of stigmatization, affords more easy application, and may altogether increase patient adherence relevant for long-term ambulatory use. Monitoring of such physiological parameters has already been demonstrated to assist in the detection of generalized tonic-clonic seizures.¹⁷ Similar autonomous system measures may also provide information on detection of preictal patterns or periods. There is an urgent medical need for early identification of seizures, as this may permit earlier intervention, and ultimately improve seizure management and epilepsy care.

Deep learning has been shown to exhibit strong classification performance from complex feature sets.¹⁸ It therefore constitutes a promising technique to differentiate pre- from interictal periods based on complex, multimodal wristband data. Whereas more traditional machine learning approaches rely on hand-designed feature sets, deep learning uses multiple layers of connections to perform classification tasks

Key Points

- Wearable devices may provide patients with timely seizure warning
- Here, we applied deep learning to long-term data from wearable devices worn by epilepsy patients
- Seizures from about one-half of the patients could be predicted with better-than-chance performance
- Forecasting performance increased with the amount of training data used, suggesting that more data may lead to further improvements in the future

without the need of feature designing, which may be an advantage in relatively underexplored, multimodal datasets, such as data from wrist-worn devices.

In this work, we used neural networks on a unique dataset comprised of multimodal wristband sensor data recorded from patients with epilepsy during multiday, in-hospital monitoring. Our aims were to (1) evaluate whether forecasting solely based on wristband data could deliver better-than-chance performance, (2) assess the role of seizure timing on forecasting performance, (3) determine whether seizure forecasting differs for generalized and focal seizure types, (4) analyze the contribution of different wrist-worn modalities to seizure forecasting, and (5) determine the algorithm performance's dependence on training data size.

2 | MATERIALS AND METHODS**2.1 | Data recording and preprocessing****2.1.1 | Patients**

We recruited patients with epilepsy admitted to the long-term video-EEG monitoring (LTM) unit at Boston Children's Hospital and placed a biosensor wristband (E4, Empatica¹⁶) on either the left or right wrist or ankle for long-term recording. For the purpose of this study, we considered all patients with wristband recordings from February 2015 until October 2018. Data from one wristband per patient only were considered. When a patient recording involved multiple wristbands (eg, from wrist and ankle), we selected the data from the biosensor wristband with the longest total recording time

for further analysis. Upon patient enrollment, wristband data were continuously collected as long as technically and logistically possible throughout the LTM period, taking into consideration the availability of wristband devices, the need to charge device batteries approximately every 24 hours, and clinical considerations including tests where the device may have had to be removed. To allow stable recording conditions, we removed data from the first and last hour of each recording from further analysis. Apart from this criterion, all recorded data (including interictal, preictal, ictal data) were subsequently considered for further processing and analysis in an attempt to be as close to real-life settings as possible, and to include as much data as possible. During recording, all sensor modalities (electrodermal activity [EDA], accelerometer data in three dimensions [ACC], blood volume pulse [BVP], and temperature [TEMP]) were simultaneously monitored. From all the patients monitored by wristband recording ($n = 317$), we only included the patients with at least one seizure during the wristband recording period, limiting our analysis to 69 patients (Table 1). We analyzed all epileptic seizure types occurring in a patient, which included primary and secondary generalized, and focal seizures (Table 1), as determined by board-certified epileptologists. Written informed consent was obtained from all participants, or their guardians, enrolled in the study. We received approval from the Boston Children's Hospital Institutional Review Board.

2.1.2 | Data selection and labeling

A prerequisite for seizure risk assessment is the reliable distinction between pre- and interictal periods. For this purpose, we analyzed continuous, nonoverlapping 30-second segments of wristband recordings composed of six sensor data streams (EDA, ACC, BVP, and TEMP; Figure 1A).

To train a forecasting algorithm, data needed to be labeled as either pre- or interictal. To define preictal periods during training of the algorithm, we focused only on lead seizures. We defined a lead seizure as a clinical seizure that occurs at least 2 hours after a preceding seizure. We considered a 30-second data segment as preictal if it occurred between 61 minutes and 1 minute prior to such a seizure, thus leaving a 1-minute buffer prior to seizure onset (Figure 1B, red). This preictal window definition was chosen to be commensurate with other seizure-forecasting research using EEG and ECoG³⁻⁶ and to account for potential small ambiguities in determining the exact seizure onset between EEG data and wristband data. Thirty-second data segments were classified as interictal if they were 2 hours or more away from any seizure (Figure 1B, green). Electrographic seizure onset was determined using video and EEG recordings. For training, we thus excluded intervals directly after the onset of a seizure or when many seizures occurred in rapid progression, to

not bias our analyses with seizure effects and with postictal period findings.¹⁹ Note again that for forecasting on test patients, all available sensor data were used.

2.2 | Preparation of training, validation, and test data

For the main results of our study, we applied a leave-one-subject-out cross-validation approach, where data from 68 patients were used for training, and testing was done on the full dataset of the one remaining out-of-sample patient. This allowed us to maximize the number of patients included in training and testing. Additional analyses were performed to train on a lower number of patients and use other patients for validation before testing on one out-of-sample test patient (see below; Figures 1D and 4).

For the preparation of the training datasets, we first identified all 30-second preictal segments and matched them by an equal number of randomly chosen interictal segments in each patient (Figure 1C). This resulted in a total of 39 814 30-second segments, which (minus the ones belonging to the test patient) were used for training. This matching was done to handle the imbalanced data during training (interictal data mostly outnumber preictal data by a large factor in each patient).^{20,21} Next, the matched data from 68 patients were used for training, whereas testing was generally performed on the remaining patient's full dataset.

We also performed our analysis on training data with <68 patients, where the remaining patients (apart from the test patient) were used for validation. Based on monitoring these validation data, we observed no indication of overfitting (Figure 1D; in this case: training on 67 patients, validation on one patient). Results for the main part of the article are thus reported for the leave-one-out cross-validation approach where data from 68 patients was used for training and testing was done on the one remaining patient. Additionally, we also report summary results for seizure-forecasting performance when training was done on fewer patients (Figure 4).

2.3 | Neural networks and training

We used long short-term memory (LSTM) networks, as they are specifically designed for learning underlying representations in time series data and have been shown to provide robust classification performance based on multidimensional time series data.¹⁸ The wearable device records data from different modalities at different sampling frequencies: EDA and TEMP are recorded at 4 Hz, BVP is recorded at 64 Hz, and ACC are recorded at 32 Hz. To use the wristband sensor data in LSTM networks, data were downsampled to 4 Hz for all sensors (downsampling after applying an antialiasing order

TABLE 1 Demographics of patients

Patient	Gender	Age, yr	Age of first seizure, yr	Seizure types	MRI findings	Etiology	Wristband location
1	Male	15	13	Focal onset	Not noted	Structural	Left wrist
2	Male	13	0	Focal onset	Normal	Unknown	Left ankle
3	Female	17	Unknown	Focal onset	Normal	Unknown	Left wrist
4	Female	2	Unknown	Focal onset	Volume loss, unspecified	Structural	Left ankle
5	Female	5	4	Unclassified	Malformation	Structural	Left wrist
6	Female	15	3	Generalized onset, unclassified	Gliosis, unspecified	Structural	Right wrist
7	Female	22	10	Focal onset	Infarction	Structural	Right ankle
8	Female	16	15	Focal onset	Normal	Unknown	Left ankle
9	Male	17	7	Focal onset	Normal	Unknown	Right wrist
10	Male	3	2	Focal onset	Dysplasia	Structural	Right ankle
11	Female	14	Unknown	Generalized onset	Cyst	Unknown	Left wrist
12	Male	11	1	Focal onset	Normal	Unknown	Left wrist
13	Male	10	1	Focal onset	Malformation	Structural	Right wrist
14	Female	13	11	Focal onset, unclassified	Volume loss, unspecified	Structural	Left wrist
15	Male	16	10	Focal onset	Normal	Unknown	Left wrist
16	Male	8	7	Generalized onset	Normal	Unknown	Left ankle
17	Female	2	Unknown	Focal onset, unclassified	Tuberous sclerosis/ hamartoma	Genetic	Left ankle
18	Male	9	1	Focal onset, generalized onset	Not noted	Unknown	Right ankle
19	Male	10	8	Focal onset	Volume loss, unspecified	Unknown	Right wrist
20	Male	9	5	Focal onset, generalized onset	Normal	Unknown	Left ankle
21	Male	5	0	Generalized onset	Normal	Unknown	Left wrist
22	Female	14	7	Focal onset	Infarction	Structural	Right wrist
23	Female	15	13	Focal onset, generalized onset	Volume loss, unspecified	Genetic	Left ankle
24	Female	3	0	Generalized onset, unclassified	Resection	Structural	Right ankle
25	Male	13	0	Focal onset	Tuberous sclerosis/ hamartoma	Genetic	Right ankle
26	Male	0	0	Focal onset	Malformation	Structural	Right ankle
27	Female	27	14	Generalized onset	Normal	Unknown	Right wrist
28	Male	17	15	Focal onset	Tumor	Structural	Right ankle
29	Male	13	0	Focal onset	Resection	Unknown	Left wrist
30	Female	2	1	Generalized onset	Not noted	Unknown	Right ankle
31	Male	8	1	Focal onset	Volume loss, unspecified	Structural	Right wrist
32	Female	15	0	Focal onset, unclassified	Hippocampal sclerosis	Structural	Right ankle
33	Male	7	4	Focal onset	Infarction	Structural	Left wrist
34	Female	17	1	Generalized onset	Not noted	Unknown	Right wrist
35	Female	3	Unknown	Focal onset	Volume loss, unspecified	Structural	Right ankle
36	Female	6	3	Generalized onset, unclassified	Gliosis, unspecified	Structural	Left wrist
37	Male	1	0	Generalized onset	Not noted	Unknown	Left ankle
38	Female	11	7	Generalized onset	Normal	Unknown	Right wrist
39	Male	12	6	Generalized onset	Volume loss, unspecified	Structural	Right wrist
40	Male	7	Unknown	Generalized onset	Volume loss, unspecified	Metabolic	Right ankle
41	Male	3	1	Focal onset	Resection	Structural	Right wrist

(Continues)

TABLE 1 (Continued)

Patient	Gender	Age, yr	Age of first seizure, yr	Seizure types	MRI findings	Etiology	Wristband location
42	Female	10	1	Focal onset, unclassified	Dysplasia	Structural	Right wrist
43	Female	13	0	Focal onset	Infarction	Structural	Left wrist
44	Male	2	0	Generalized onset, unclassified	Tuberous sclerosis/ hamartoma	Genetic	Left ankle
45	Male	4	2	Focal onset	Volume loss, unspecified	Structural	Left ankle
46	Female	5	0	Focal onset	Infarction	Structural	Right ankle
47	Male	8	5	Focal onset	Tumor	Structural	Right ankle
48	Female	13	9	Focal onset, unclassified	Normal	Unknown	Left ankle
49	Male	12	7	Focal onset	Normal	Unknown	Right wrist
50	Male	0	0	Generalized onset, unclassified	Not noted	Unknown	Right ankle
51	Male	9	Unknown	Generalized onset	Normal	Unknown	Right wrist
52	Female	13	3	Focal onset	Volume loss, unspecified	Structural	Right wrist
53	Female	11	1	Focal onset	Hippocampal sclerosis	Structural	Right ankle
54	Male	1	Unknown	Focal onset	Dysplasia	Structural	Left ankle
55	Male	14	0	Focal onset, unclassified	Resection	Structural	Left wrist
56	Female	11	10	Generalized onset	Tumor	Structural	Right wrist
57	Female	19	0	Generalized onset	Volume loss, unspecified	Unknown	Right ankle
58	Male	7	1	Focal onset	Normal	Unknown	Left ankle
59	Male	14	7	Generalized onset	Dysplasia	Structural	Left ankle
60	Male	2	0	Generalized onset	Dysplasia	Unknown	Left ankle
61	Male	12	1	Focal onset	Not noted	Structural	Right wrist
62	Male	9	8	Focal onset	Malformation	Unknown	Left ankle
63	Male	0	0	Unclassified	Infarction	Structural	Right ankle
64	Male	9	1	Focal onset	Normal	Unknown	Right ankle
65	Male	10	0	Focal onset	Cyst	Unknown	Left ankle
66	Male	5	Unknown	Focal onset	Malformation	Structural	Right ankle
67	Female	3	1	Generalized onset, unclassified	Normal	Genetic	Left ankle
68	Male	11	Unknown	Generalized onset	Normal	Unknown	Left wrist
69	Male	21	3	Focal onset, generalized onset	Volume loss, unspecified	Genetic	Right ankle

Abbreviation: MRI, magnetic resonance imaging.

8 Chebyshev type I filter; Python function `scipy.signal.decimate`), to provide the same vector length for each 30-second segment (ie, 120 sampling points). To limit LSTM networks from overfitting, network architecture was kept simple and shallow (Table S1, Figure 1D), and training was performed on matched data, that is, where both classes appeared equally often. As stated above, no indication of significant overfitting was observed in learning metrics (Figure 1D).

We also compared the results using the LSTM network to forecast performance from a one-dimensional convolutional neural (1DConv) network. 1DConv networks are often easier and faster to train than LSTM networks while also being known to exhibit good performance on time series data.¹⁸ Convolutional networks have recently also been applied to classification problems in epilepsy.^{14,22} Table S1 shows the

summary of the LSTM and 1DConv network hyperparameters used and a graphical schema depicting the network. All networks were trained for 200 epochs. We employed a learning rate of 0.001, that is, the parameter determining the step size at each iteration while moving toward minimizing the loss function, and Adam optimizer with a binary cross-entropy loss function. Analyses were performed with in-house-written code using Python (version 2.7) and Keras with Tensorflow backend.

2.4 | Performance metrics and statistical tests

Seizure-forecasting performance was assessed on the full time series data (after removal of potential dropouts, for example,

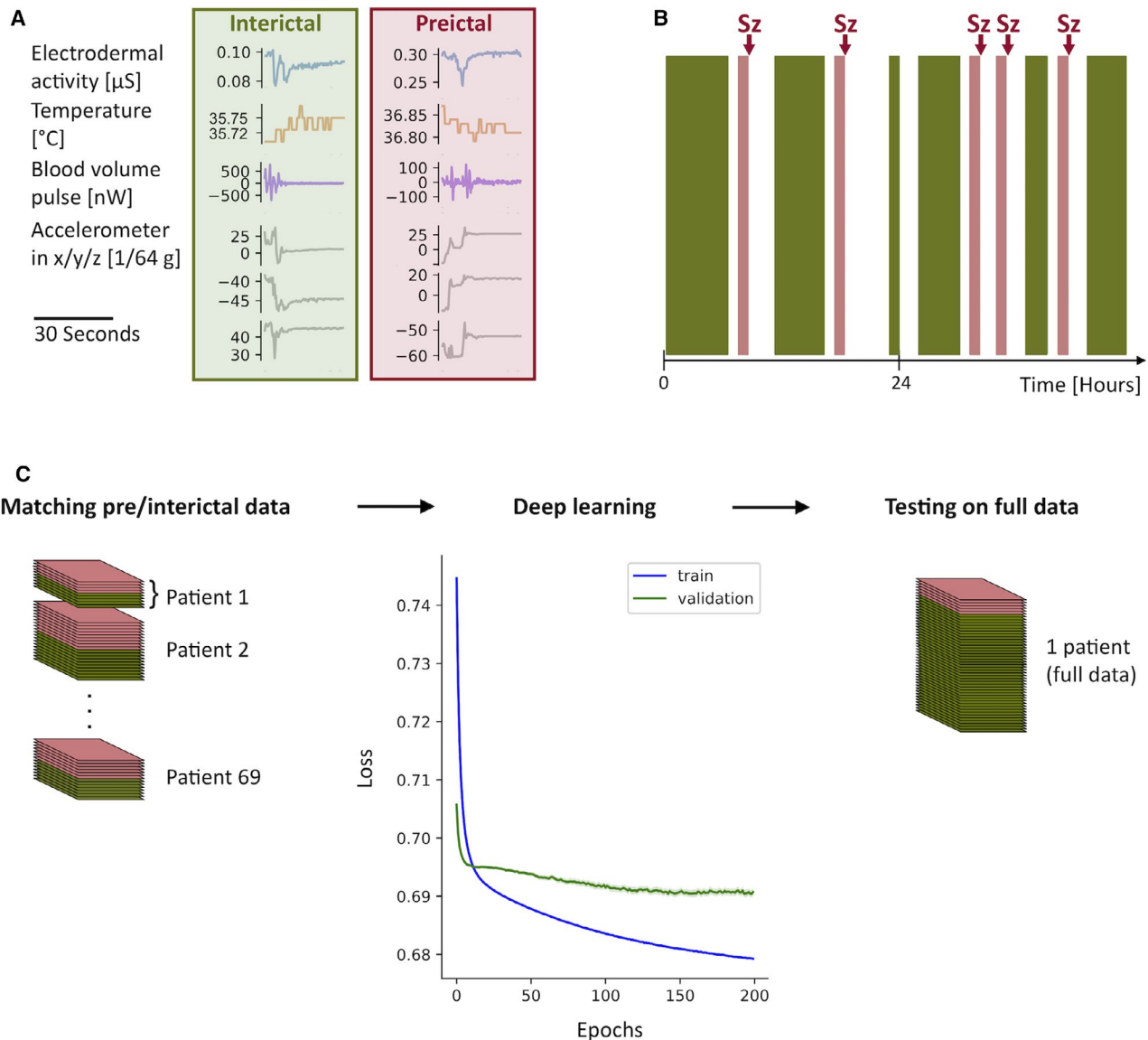


FIGURE 1 Outline of data processing for seizure forecasting. A, Continuous multimodal time series data were separated into consecutive 30-second segments. B, Data segments in training data were labeled either preictal (red) or interictal (green). Sz, seizure. C, For each patient, preictal segments were matched with the same number of interictal segments for training. Results reported in this article are generally based on a leave-one-subject-out cross-validation approach, where training was performed on 68 patients and testing on the one left-out patient. Loss curves in the middle correspond to control analyses, where 67 patients were used for training, one patient for validation, one for testing. Solid lines and shading indicate mean \pm SEM across all patients

when the wristband was replaced by a new one for charging purposes) from out-of-sample test patients. Figure S3 shows the full data from one exemplary test patient. To be clinically useful, the binary classification has to be translated into a suitable user interface.²³ We here used a sliding window approach in which the individual 30-second segment predictions were averaged over an *integration window*. If this averaged value crossed a *threshold*, an alarm would be initiated which would last for the duration of a *seizure occurrence period*. A new alarm could only be initiated once the seizure occurrence period had passed (Figure S3). This postprocessing thus requires the determination of three additional variables: integration window, threshold, and seizure occurrence period. In

long-term recordings, parameters like this can in principle be optimized at the individual patient level, for example, by optimizing these parameters during an initial adjustment phase.¹⁴ Here, in our leave-one-subject-out cross-validation approach, we used the training data to find the optimal parameters using a grid search (Figure S2; integration window values: 150, 300, 600, 1200 seconds; seizure occurrence period values: 150, 300, 600, 1200, 2400, 3600, 7200 seconds; threshold values: 0.5, 0.52, 0.54, 0.56, 0.58, 0.6). Therefore, for forecasting of a test patient using a network trained on 68 patients, the parameters yielding the on average highest improvement over chance (IoC) for predictions on the training data predictions from these 68 patients were chosen (Figure S2, blue square).

Note that this parameter selection process ensures that no information from the test patient is used to infer these three parameters. This grid search yielded highly similar results across patients (because training datasets overlap by 67 patients) and provided an integration window of 600 seconds, seizure occurrence period of 3600 seconds, and threshold of 0.54 as the best parameter set in all 10 cases where it was applied (Figure S2). Although other parameters also yielded good forecasting performance results, these parameters were consequently chosen for the reported results in this article. As stated above, in future long-term trials, it is conceivable that these parameters can be further optimized for the individual patient, for example, by determining these parameters during an initial training period¹⁴ or to tune sensitivity versus time in warning (TiW) to suit the individual patient's needs.²⁴

As clinically useful metrics to evaluate forecasting algorithm performance,^{2,25} we used the metrics applied in Cook et al⁸: sensitivity, TiW, and improvement over chance (IoC)¹⁴:

- *Sensitivity*: Any seizures occurring during the alarm were considered to be true positive (TP), and seizures occurring outside alarms to be false negative (FN). Sensitivity (S) is then defined as: $S = TP / (TP + FN)$.
- *TiW*: The fraction of time spent in warning.
- *IoC*: $IoC = S - TiW$.

All metrics are reported in percent.

We report mean prediction scores for these variables over 10 independent runs for each patient where each run corresponds to an independently trained deep neural network. As an additional control, we furthermore show IoC results for time-shuffled predictions, that is, when the 30-second segment predictions are randomly distributed over time prior to calculation of the algorithm performance metrics sensitivity, TiW, and IoC (Figure 4, gray line).

A two-sided Wilcoxon signed-rank test was used to assess the significance of IoC values in each patient. We controlled for multiple comparisons with the Benjamini and Hochberg false discovery rate using a threshold of 0.05.²⁶ A Kruskal-Wallis *H* test followed by a two-sided Wilcoxon signed-rank test was used to assess significance of IoC values depending on individual sensor modalities. A two-sided Mann-Whitney *U* test was applied for comparison between groups with respect to age, seizure type, and location where the wristband was worn. We deemed $P \leq .05$ to be significant.

2.4.1 | Assessment of seizure timing by generation of surrogate data

To better understand the dependence of forecasting performance on time of day and vigilance state, we employed the forecasting algorithm also to time-matched seizure

surrogate data²⁷ in the patients who had sufficiently long data for such an analysis. The underlying idea is that, if forecasting performance was primarily based on detecting a certain time of day (or better: the respective vigilance state, which should be similar at the same time of day), then shifting seizure onset times to the same time of a random other day should not lead to a strong decline in forecasting performance. Conversely, if performance declined significantly under time-matched seizure surrogate data, then it would be unlikely that the forecasting algorithm was mainly detecting time of day (or the respective vigilance state). In our data, we identified $n = 28$ patients with sufficiently long data recordings to randomly shift all seizure onset times to another day while maintaining the exact same time of day. Using the metrics outlined above, forecasting performance was then compared between data containing the true seizure onset times and data containing the time-matched seizure surrogate data (Figure S5).

2.4.2 | Assessment of performance contribution of individual sensor modalities

To assess the influence of individual data modalities, we performed additional experiments where, during training, the sensor data of individual modalities (EDA, TEMP, BVP, or ACC) were all set to zero. Any potential information with respect to the preictal period contained in this particular modality could thus not be learned during training. We then assessed performance with and without this modality in terms of IoC (Figure 3).

2.5 | Data and code availability

The datasets and code generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

3 | RESULTS

3.1 | Demographics

Our dataset was comprised of multiday recordings from 69 epilepsy patients (mean age = 9.8 ± 5.9 years [mean \pm SD], 28 females, total duration = 2311.4 hours, 452 seizures; Table 1).

3.2 | Overall prediction performance

We evaluated performance of the proposed seizure-forecasting system in terms of sensitivity, TiW, and IoC.^{2,5,8,25} We applied a leave-one-subject-out cross-validation

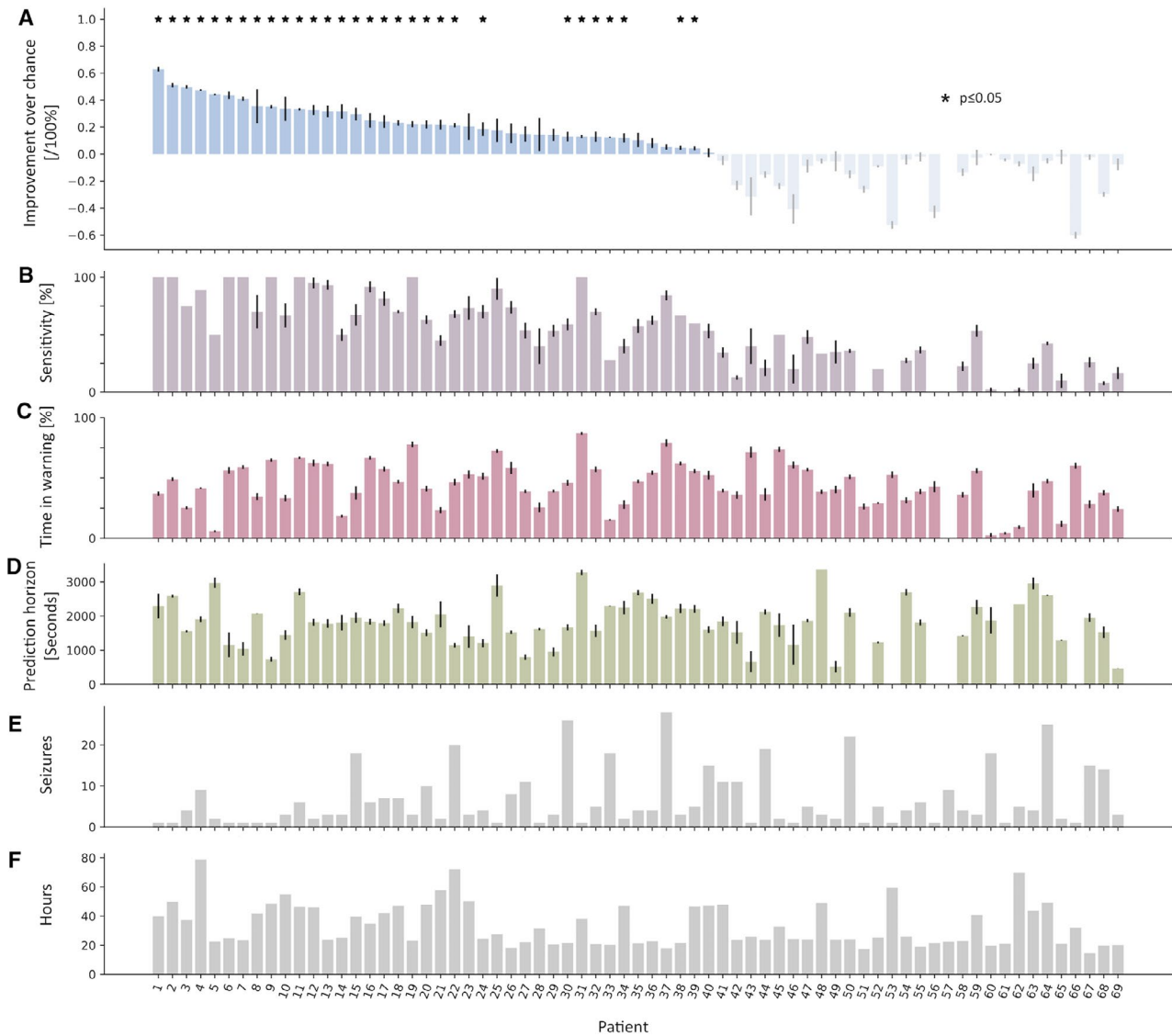
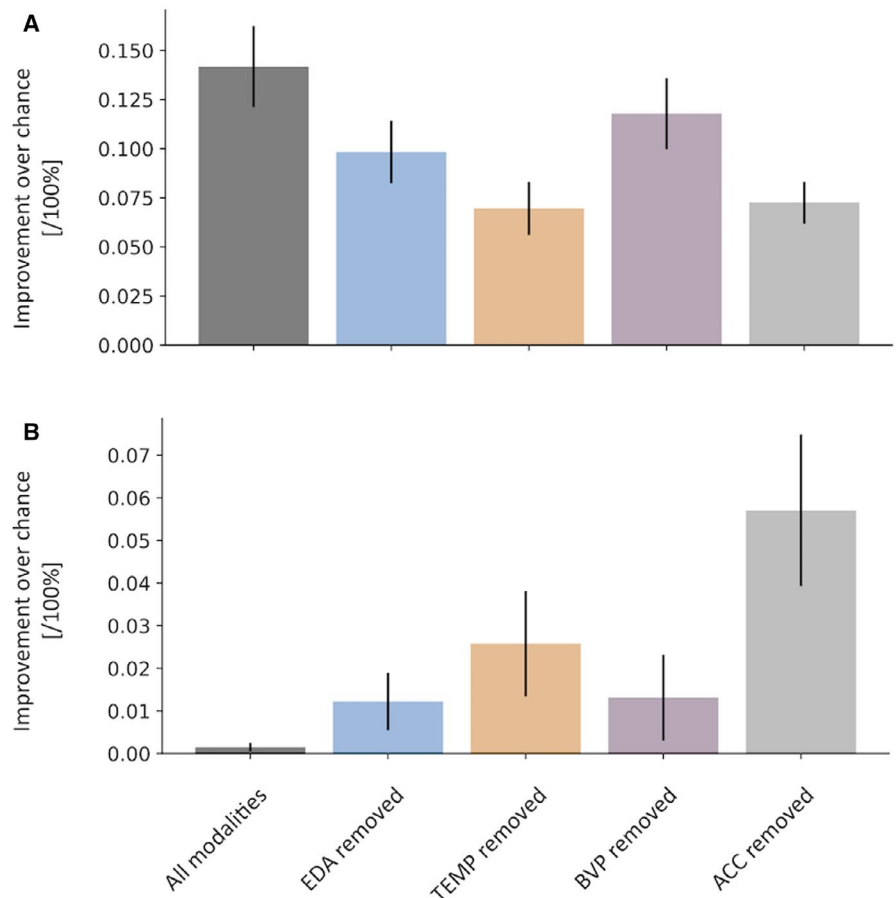


FIGURE 2 Forecasting performance results in a pseudoprospective approach for all patients. Bars represent mean values over 10 runs; error bars indicate SEM. A, Improvement over chance. B, Sensitivity. C, Time in warning. D, Prediction horizon. E, Number of seizures. F, Number of hours in recording. *IoC values significantly larger than zero after correcting for multiple comparison

approach where matched pre-/interictal data from 68 patients were used for training (Figure 1C), and testing was done on the full dataset of the one remaining out-of-sample patient. Control analyses indicated no signs of significant overfitting during training (Figure 1D). Performance was calculated for predictions from 10 independently trained LSTM networks (Table S1) for each patient. Figure 2 shows the average performance for all 69 patients. Significant IoC larger than zero is an indication that seizure forecasting is useful in a clinical setting.⁵ In our data, seizure forecasting was significantly better than chance for 43.5% of patients (30 of 69 patients). For these patients, we obtained a mean IoC of $28.5 \pm 2.6\%$, a mean sensitivity of $75.6 \pm 3.8\%$, and a mean TiW (ie, the percentage of time spent in warning) of $47.2 \pm 3.4\%$ (mean \pm SEM).

Across all patients, including those with nonsignificant, nonpositive IoC, mean IoC was $14.1 \pm 1.9\%$, mean sensitivity was $51.2 \pm 3.8\%$, and mean TiW was $43.7 \pm 2.3\%$. Note that mean IoC in patients with nonpositive mean IoC was set to zero prior to averaging across patients. To investigate whether an increase in model complexity might lead to an improved performance, we also compared our results to a more complex LSTM network (100 LSTM units instead of 10 units; Table S1). This did not result in an overall better performance (mean IoC across patients of $13.7 \pm 1.8\%$). Mean age in the group of patients with better-than-chance forecastability was higher than in the remainder of patients (mean age = 10.7 ± 5.3 years vs 9.1 ± 6.2 years), although this difference did not reach statistical significance. ($P = .16$, Wilcoxon signed-rank test).

FIGURE 3 Influence of individual sensor modalities on overall performance. Each bar indicates the improvement over chance (IoC) averaged across patients. Results are based on five independent network runs for each patient and modality. A, IoC averaged across all patients. Results indicate that, on average, best performance is achieved when all modalities are included for training and each modality contributes to seizure forecasting. B, Results averaged across the 20 patients with worst overall forecasting performance. These results indicate that accelerometer data in three dimensions (ACC) may particularly contribute to the poor performance in these patients. BVP, blood volume pulse; EDA, electrodermal activity; TEMP, temperature



3.3 | Evaluation of seizure timing as potential prediction confounder

Seizures are more likely to occur at certain times of the day or vigilance states, and we also noted a bimodal peak in the histogram of seizure occurrences in our dataset (Figure S4). To rule out detection driven simply by time of day and respective vigilance state, we also employed the forecasting algorithm to time-matched seizure surrogate data²⁷ in the patients who had sufficiently long data records ($n = 28$ patients; Figure S5). Performance was found strongly and significantly decreased for time-matched seizure surrogates in comparison to actual seizure times (Figure S5A; $P = .011$, Wilcoxon signed-rank test). These results consequently provide strong indication that forecasting is not primarily based on detection of a time of day or the respective vigilance state.

3.4 | Prediction horizon

For practical application as a warning system for patients, the expected time between alarm onset and seizure onset, the prediction horizon (Figure S3), is of particular interest. A sufficiently long alarm may afford patients to take precautionary

steps or avoid certain activities. The mean prediction horizon across all patients observed in our data was 1844 ± 80 seconds. For the 30 patients with significantly better-than-chance forecastability, the mean prediction horizon was 1896 ± 101 seconds (minimum = 738 seconds, maximum = 3273 seconds), a period that may be long enough to afford reasonable warning in advance and that excludes the possibility that some early seizure-related activity might have contributed to the detection.

3.5 | Prediction performance in relationship to seizure type and sensor location

Next, we investigated whether prediction performance was dependent on seizure type, in particular whether performance depended on seizures of focal or generalized onset. We thus compared prediction performance between patients with only focal onset seizures ($n = 35$ patients) and patients with only generalized onset of seizures ($n = 16$ patients; Table 1). Group comparison revealed no significant difference in IoC values ($P = .32$). Similarly, no significant dependence on where the device was worn (wrist: $n = 31$ patients, ankle: $n = 38$ patients) in terms of IoC performance values was revealed by group comparison ($P = .14$).

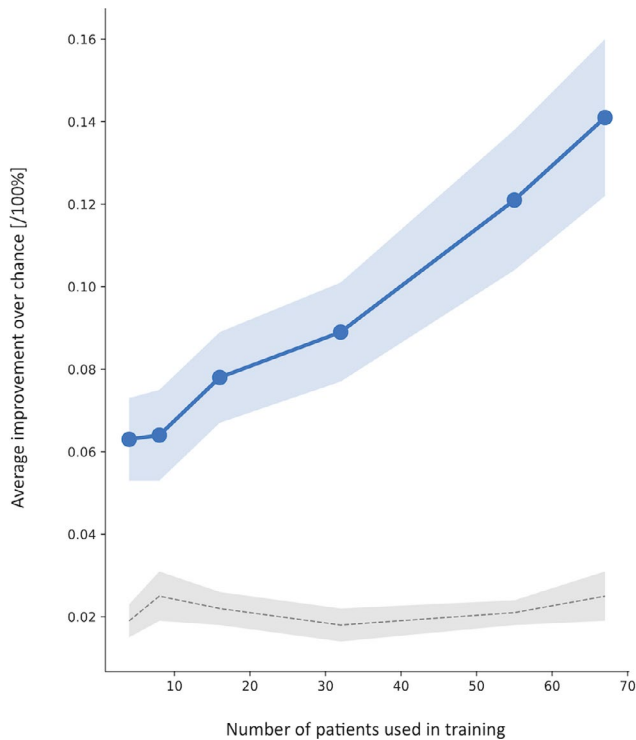


FIGURE 4 Seizure-forecasting performance improves with larger training datasets. Blue indicates average improvement over chance for increasing sizes of training data. Performance was assessed at the individual patient level (test data), where training of neural network was performed with training data comprised of 4, 8, 16, 32, 55, or 68 patients. Plot indicates averages across all patients (mean \pm SEM). Gray indicates average improvement over chance for time-shuffled predictions (patients with mean negative improvement over chance [IoC] are set to IoC = 0 prior to averaging across patients). Note that the gray markers indicate what true chance prediction at the group level would look like, that is, when predictions are uniformly distributed across recording time, and accumulations of predictions indicative of proictal dynamics are destroyed

3.6 | Contribution of individual modalities to seizure prediction

As a purely data-driven approach that gives machine learning methods free reign to identify the most useful data features across modalities, our approach considered all data modalities (EDA, TEMP, BVP, ACC). To better understand the information that is provided by the multimodal data in terms of each modality's contribution for successful seizure forecasting, we performed additional analyses by removing the information of each modality individually. Specifically, we repeated neural network training where all data from each modality individually were set to zero, thereby providing no information about the preictal period (Figure 3). Across all patients, these analyses indicated that performance was on average best when all modalities

were included (Figure 3A). Thus, each modality contributed to seizure forecasting. Relative influences for seizure-forecasting performance varied between modalities, but variations did not reach statistical significance ($P = .099$, Kruskal-Wallis H test). These analyses also confirmed that ACC contributed strongly to algorithm performance also under downsampling. Finally, to obtain some insights into which modalities may have contributed most to the poor performance of the worst-performing patients, we assessed relative influence of individual modalities in the 20 worst-performing patients only (Figure 3B). There, results indicated ACC in particular to have had a large detrimental influence on IoC values.

3.7 | Performance based on training set size

Machine learning, particularly deep learning, benefits from large datasets that afford learning of the underlying data representations while also containing enough variability to permit generalization to unseen data. In a use-case like ours, it is conceivable that performance requires a certain amount of data and, more generally, benefits from training on larger datasets. To determine the relationship between seizure-forecasting performance and size of the training dataset, and to obtain a better understanding of how our approach might benefit from more data in the future, we evaluated performance under different amounts of training data. For this purpose, instead of training on all 68 patients in a leave-one-out approach, as described above, we systematically reduced the amount of training data by considering only a smaller number of patients ($n = 4, 8, 16, 32, \text{ or } 55$ patients) for training. Specifically, performance for each test patient was calculated for 10 independently trained networks, where training data were composed of only n number of randomly chosen patients in each run. Figure 4 shows the dependence of forecasting performance in terms of average IoC across all patients (mean IoC in patients with nonpositive mean IoC again set to zero prior to averaging). Performance increased monotonically as more patients were used for training (blue line), whereas control analyses based on randomly time-shuffled predictions remained low and did not improve with increasing training data size (gray line). Note that the gray markers indicate what true chance prediction at the group level would look like, that is, when predictions are uniformly distributed across recording time, and accumulations of predictions indicative of proictal dynamics are destroyed (Figure 4). The performance increase with training size demonstrates the benefits of creating larger datasets for training. As no saturation effect under larger training size is apparent, these results suggest that larger datasets than the one used in this study might, on average,

improve seizure-forecasting performance even more in the future.

4 | DISCUSSION

We showed that forecasting of seizures is feasible with wrist-worn data. Forecasting was independent of time of day and independent of focal or generalized seizure type, suggesting that such an approach might be useful for a broad range of epilepsy patients. All wrist-worn data streams contributed to forecasting. Preliminary analyses suggest that more data will improve prediction.

Of the patients included in our analysis, about one-half displayed a significant IoC. Although future research will need to quantify the clinical value of these forecasts for patients, our results indicate that seizure forecasting is feasible with relatively noisy, multimodal signals recorded from devices far away from the brain. As such, the results can be regarded as a benchmark for future methodological refinements and clinical usability assessments. The better-than-chance classification performance in about one-half of the patients was obtained despite the comparably brief duration of data and despite the variability in seizure types, age, and wristband location. Our approach used the data in raw format ("as is") in an attempt to maximize the transferability of our approach to real-world, noisy conditions, and utilized a pseudoprospective assessment scheme, that is, using only data from the past for future predictions, which is essential for use in real-world conditions to evaluate future seizure risk.

Others have shown effects of seizure timing and seizure location on seizure occurrence.²⁸ Adding these data among other clinical predictors may improve performance in larger datasets in the future.²⁹ Furthermore, it is possible that the autonomic variability and normative ranges, which vary significantly across pediatric age groups (eg, lower resting heart rates with increasing age), may have complicated seizure forecasting. Training on more homogenous or adult patient populations may thus potentially improve prediction rates.

Our results demonstrate that there exists a preictal signature in autonomous nervous system³⁰ and actigraphy data, which, despite possibly not being detectable by visual inspection, may be picked up by deep learning. Importantly, our work demonstrates that this signature may be learned across patients and is therefore not patient-specific. This can be considered an advance from most traditional seizure prediction work, which mostly succeeded when using algorithms trained specifically for the individual patient.^{5,6,8} An algorithm that can be trained across patients has the advantage that it can be readily employed to a new patient, without any training to learn patient-specific factors or expert knowledge to set parameters. For the first-in-man study, for example, that reported statistically significant seizure forecasting for

10 patients, forecasting algorithm parameters were set for each individually after an initial training phase.⁸ If the IoC values of these patients are averaged, a mean IoC of 38.1% is obtained.

The average IoC for the 30 patients with significantly better-than-random predictability in our study was slightly less (mean IoC = 28.5%). Clearly, a direct comparison between these studies is limited, and care has to be taken when comparing the results obtained from 10 of 11 patients⁸ to results from 30 of 69 patients here. However, considering that our results reported here did not require invasive ECoG, patient-specific training and parameter optimization, or significant patient selection (we included patients with both focal and generalized seizure onset; Table 1), and were obtained from an algorithm trained on out-of-sample data, which may allow constant improvement with larger and larger datasets (Figure 4), our results are overall encouraging and provide a step forward.

Seizure forecasting builds on the notion that a preictal period, during which a seizure is more likely to occur soon, can be reliably distinguished from interictal periods. To this end, most studies have focused on data recorded either from ECoG and EEG or from ECG. ECG has thus been a long-standing example that peri- and preictal changes not only can be detected within the central nervous system but are also reflected in a variety of cardiac effects.^{31–33} Cardiac activity is controlled by parasympathetic and sympathetic branches of the autonomic nervous system, with the former producing an inhibitory response and the latter producing an excitatory response on heart rate.²⁴ Preictal changes in brain activity that occur in or propagate to autonomic control centers may affect this autonomic balance and, consequently, affect cardiac activity during the lead-up to a seizure. A recent study that compared the information content in ECoG, EEG, and ECG in terms of identifying preictal periods found that single-channel ECG contains a comparable amount of information to multichannel EEG,³⁵ which highlights the relevance of peripheral sensors for seizure forecasting. Autonomous nervous system changes are captured by the wristband sensors used in this study in several ways. Electrodermal activity is known to be sensitive to sympathetic innervation. Blood volume pulse curves contain information about heart rate, which is controlled by the parasympathetic and sympathetic interplay. Similarly, body temperature is known to be maintained by the autonomic nervous system. The approach proposed in this study builds on monitoring these autonomous nervous system functions along with actigraphy, which indirectly also monitors resting periods and sleep, and therefore pioneers the seizure forecasting based on such multimodal sensor data, going beyond more traditional ECoG/EEG and ECG approaches.⁷

In our approach, we used the same LSTM model for all patients, albeit models were trained for each patient

separately. Although it is possible that model hyperparameters individualized for each patient might bring about better performance, we chose to have the same model architecture across patients, which could potentially be implemented "out-of-the-box" in future prospective settings. Ultimately, the usefulness of a seizure-forecasting system may depend also on a patient's preference in terms of sensitivity and time spent in warning. Our approach has three parameters—integration window, seizure occurrence period, and threshold (Figure S3)—that could be tuned for these purposes in a straightforward manner. With longer data, one may tune these parameters for the individual patient's needs, for example, during an initial adjustment phase, for optimal future performance and patient preferences.^{14,24} Due to the relative shortness of our data, which only covered up to a few days per patient, we did not attempt this and used the parameters identified in a grid search from the training data (Figure S2). Our results show that statistically significant seizure forecasting is possible, and it is likely that a parameter tuning at the individual patient level on longer data may improve performance even more. Furthermore, although we tested different models with different complexity, it is certainly possible that other, more refined machine learning methods will provide even better results and should be explored in the future. Crowd-sourcing competitions may be a great way to allow such exploration of many different approaches to find the best-performing algorithm.^{36,37} It is furthermore possible that seizures may manifest very differently for some subset of patients, in which case increasing the amount of training data might not help, but different approaches (eg, other sources of data, higher sampling rate, personalized or semisupervised models) may help to develop useful methods also for these patients.

Results need to be interpreted in the setting of data acquisition. One limitation of our study is the relative short duration of recordings, with only a few days of continuous data per patient. Training data benefits from long periods of data acquisition, where algorithms can better learn to generalize, and which give a more realistic account of seizure-forecasting capabilities. However, the current dataset is unique in the sense that it contains multimodal sensor data over several days from a relatively large number of epilepsy patients along with ground-truth video-EEG, which is essential for proper training. The better-than-chance predictability in about one-half of the patients in this study is therefore encouraging for future, longer trials using these sensors. This encouragement is further supported by the improvement of seizure-forecasting performance as a function of dataset size (Figure 4). Based on these observations, to improve seizure-forecasting performance even more, it may be useful to pool datasets from different institutions and laboratories together. The creation of additional datasets, similar to the one used here, will also afford validation of findings.

Another limitation is the absence of benchmarks to compare our approach to. Although we attempted a comparison of our results to the first-in-man seizure-forecasting study⁸ and used chance predictors as well as well-established statistical frameworks, the uniqueness and novelty of the current dataset limits more comprehensive comparison to other approaches. Lastly, we did our best to balance patient recruitment and data acquisition, but cannot rule out selection and information bias, and confounding by clinical data, which are inherent to this and similar studies. By including timing and seizure type analyses, we did our best to evaluate the effect of confounders and covariates. There is growing awareness of the benefits of creating data warehouse ecosystems that allow rigorous and continuous reevaluation and benchmarking by making data and algorithms available to many researchers, and evaluating larger patient datasets.^{36,37} We expect that these open-science efforts will increase the reproducibility and help benchmark and improve algorithms, such as the ones proposed in the current study, in the future.

We here assessed the utility of physiological sensor data recorded from a wearable device to estimate seizure risk in a clinically useful way. Our work is motivated by the potential benefits for patients and clinicians from a robust seizure gauge.¹ Forecasting seizures would provide patients with timely warning to adapt daily activities and allow clinicians to titrate therapies and develop novel interventions that potentially could prevent impending seizures.^{1,38} Peripheral sensor data that can be recorded easily and noninvasively with a wristband would be desirable for such a purpose, because approaches based on ECoG⁸ or a large number of scalp EEG channels¹⁵ limit broad clinical application.

Seizure forecasting is likely to bring about notable benefits for many epilepsy patients and may improve seizure management from the perspective of clinicians. To make seizure forecasting available for broad use, noninvasive, easily applicable techniques are crucial. We here demonstrated the capability of multimodal wristband sensor data from easy-to-use, noninvasive devices in combination with deep learning to provide statistically significant seizure forecasting. These results provide a step toward patient empowerment and more objective epilepsy diagnostics with feasibility for broad application.

ACKNOWLEDGMENTS

This work was supported in part by the Epilepsy Research Fund. C.M. acknowledges support by a National Alliance for Research on Schizophrenia & Depression Young Investigator Grant from the Brain & Behavior Research Foundation. Open access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST

C.M. is part of patent applications to detect and predict clinical outcomes and to manage, diagnose, and treat

neurological conditions. T.L. serves on the Council of the American Clinical Neurophysiology Society, on the American Board of Clinical Neurophysiology, as founder and consortium principal investigator of the Pediatric Status Epilepticus Research Group, as an Associate Editor for *Wyllie's Treatment of Epilepsy* 6th and 7th editions, and as a member of the NORSE Institute, PACS1 Foundation, and CCEMRC. He served as Associate Editor of *Seizure* and on the Laboratory Accreditation Board for Long Term (Epilepsy and Intensive Care Unit) Monitoring in the past. He is part of patent applications to detect and predict clinical outcomes and to manage, diagnose, and treat neurological conditions, epilepsy, and seizures. T.L. is coinventor of the TriVox Health technology, and T.L. and Boston Children's Hospital might receive financial benefits from this technology in the form of compensation in the future. He has received research support from the Epilepsy Research Fund, the NIH, the Epilepsy Foundation of America, the Epilepsy Therapy Project, and the Pediatric Epilepsy Research Foundation, and has received research grants from Lundbeck, Eisai, Upsher-Smith, Mallinckrodt, Sunovion, Sage, Empatica, and Pfizer, and past device donations from various companies, including Empatica, SmartWatch, and Neuro-electrics. In the past, he served as a consultant for Zogenix, Upsher-Smith, Amzell, Engage, Elsevier, UCB, Grand Rounds, Advance Medical, and Sunovion. He performs video-electroencephalographic long-term and intensive care unit monitoring, electroencephalograms, and other electrophysiological studies at Boston Children's Hospital and affiliated hospitals and bills for these procedures, and he evaluates pediatric neurology patients and bills for clinical care. He has received speaker honorariums from national societies, including the AAN, AES, and ACNS, and for grand rounds at various academic centers. His wife, Dr Karen Stannard, is a pediatric neurologist; she performs video-electroencephalographic long-term and intensive care unit monitoring, electroencephalograms, and other electrophysiological studies and bills for these procedures, and she evaluates pediatric neurology patients and bills for clinical care. None of the other authors has any conflict of interest to disclose.

ETHICAL PUBLICATION STATEMENT

We confirm that we have read the Journal's position on issues involved in ethical publication and affirm that this report is consistent with those guidelines.

ORCID

Christian Meisel  <https://orcid.org/0000-0003-2984-5480>
Rima El Atrache  <https://orcid.org/0000-0002-5600-1120>

REFERENCES

- Dumanis SB, French JA, Bernard C, Worrell G, Fureman BE. Seizure forecasting from idea to reality. Outcomes of the My Seizure Gauge Epilepsy Innovation Institute Workshop. *eNeuro*. 2017;4(6):ENEURO.0349-17.2017.
- Mormann F, Andrzejak RG, Elger CE, Lehnertz K. Seizure prediction: the long and winding road. *Brain*. 2007;130:314–33.
- Freestone DR, Karoly P, Cook MJ. A forward-looking review of seizure prediction. *Curr Opin Neurol*. 2017;30:167–73.
- Melbourne University AES/MathWorks/NIH Seizure Prediction. Available at: <https://www.kaggle.com/c/melbourne-university-seizure-prediction>. Accessed October 2, 2016.
- Mormann F, Andrzejak RG. Seizure prediction: making mileage on the long and winding road. *Brain*. 2016;139:1625–7.
- Kuhlmann L, Lehnertz K, Richardson M, Schelter B, Zaveri HP. Seizure prediction—ready for a new era. *Nat Rev Neurol*. 2018;14:618–30.
- Meisel C, Loddenkemper T. Seizure prediction and intervention. *Neuropharmacology*. 2020;172: 107898.
- Cook MJ, O'Brien TJ, Berkovic SF, Murphy M, Morokoff A, Fabinyi G, et al. Prediction of seizure likelihood with a long-term, implanted seizure advisory system in patients with drug-resistant epilepsy: a first-in-man study. *Lancet Neurol*. 2013;12:563–71.
- Lehnertz K, Dickten H, Porz S, Helmstaedter C, Elger CE. Predictability of uncontrollable multifocal seizures—towards new treatment options. *Sci Rep*. 2016;6:24584.
- Fujiwara K, Miyajima M, Yamakawa T, Abe E, Suzuki Y, Sawada Y, et al. Epileptic seizure prediction based on multivariate statistical process control of heart rate variability features. *IEEE Trans Biomed Eng*. 2016;63:1321–32.
- Kerem DH, Geva AB. Forecasting epilepsy from the heart rate signal. *Med Biol Eng Comput*. 2005;43:230–9.
- Acharya UR, Oh SL, Hagiwara Y, Tan JH, Adeli H. Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. *Comput Biol Med*. 2018;100:270–8.
- Hügler M, Heller S, Watter M, Blum M, Manzouri F, Dimpelmann M, et al. Early Seizure Detection with an Energy-Efficient Convolutional Neural Network on an Implantable Microcontroller. Paper presented at: IJCNN 2018: International Joint Conference on Neural Networks; July 8–13, 2018; Rio de Janeiro, Brazil.
- Kiral-Kornek I, Roy S, Nurse E, Mashford B, Karoly P, Carroll T, et al. Epileptic seizure prediction using big data and deep learning: toward a mobile system. *EBioMedicine*. 2018;27:103–11.
- Schulze-Bonhage A, Sales F, Wagner K, Teotonio R, Carius A, Schelle A, et al. Views of patients with epilepsy on seizure prediction devices. *Epilepsy Behav*. 2010;18:388–96.
- Poh MZ, Loddenkemper T, Swenson N, Goyal S, Madsen JR, Picard RW. Continuous monitoring of electrodermal activity during epileptic seizures using a wearable sensor. *Conf Proc IEEE Eng Med Biol Soc*. 2010;2010:4415–8.
- Poh M-Z, Loddenkemper T, Reinsberger C, Swenson NC, Goyal S, Sabtala MC, et al. Convulsive seizure detection using a wrist-worn electrodermal activity and accelerometry biosensor. *Epilepsia*. 2012;53:e93–7.
- Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw*. 2015;61:85–117.
- So NK, Blume WT. The postictal EEG. *Epilepsy Behav*. 2010;19:121–6.
- Chawla N. Data mining for imbalanced datasets: an overview. In: Maimon O, Rokach L, eds. *Data Mining and Knowledge Discovery Handbook*. Boston, MA: Springer; 2005:853–67.
- Jeni LA, Cohn JF, De La Torre F. Facing imbalanced data recommendations for the use of performance metrics. *Int Conf Affect Comput Intell Interact Workshops*. 2013;2013:245–51.

22. Truong ND, Nguyen AD, Kuhlmann L, Bonyadi MR, Yang J, Ippolito S, et al. Convolutional neural networks for seizure prediction using intracranial and scalp electroencephalogram. *Neural Netw.* 2018;105:104–11.
23. Snyder DE, Echaz J, Grimes DB, Litt B. The statistics of a practical seizure warning system. *J Neural Eng.* 2008;5:392–401.
24. Elger CE, Mormann F. Seizure prediction and documentation—two important problems. *Lancet Neurol.* 2013;12:531–2.
25. Winterhalder M, Maiwald T, Voss HU, Aschenbrenner-Scheibe R, Timmer J, Schulze-Bonhage A. The seizure prediction characteristic: a general framework to assess and compare seizure prediction methods. *Epilepsy Behav.* 2003;4:318–25.
26. Glickman ME, Rao SR, Schultz MR. False discovery rate control is a recommended alternative to Bonferroni type adjustments in health studies. *J Clin Epidemiol.* 2014;67:850–7.
27. Andrezjak RG, Mormann F, Kreuz T, Rieke C, Kraskow A, Elger CE, et al. Testing the null hypothesis of the nonexistence of a pre-seizure state. *Phys Rev E Stat Nonlin Soft Matter Phys.* 2003;67:10901.
28. Stirling RE, Cook MJ, Grayden DB, Karoly PJ. Seizure forecasting and cyclic control of seizures [published online ahead of print July 26, 2020]. *Epilepsia.* doi: <https://doi.org/10.1111/epi.16541>
29. Beniczky S, Karoly P, Nurse E, Ryvlin P, Cook M. Machine learning and wearable devices of the future [published online ahead of print July 26, 2020]. *Epilepsia.* doi: <https://doi.org/10.1111/epi.16555>.
30. Vieluf S, Reinsberger C, El Atrache R, Jackson M, Schubach S, Ufongene C, et al. Autonomic nervous system changes detected with peripheral sensors in the setting of epileptic seizures. *Sci Rep.* 2020;10(1):11560.
31. van Buren JM. Some autonomic concomitants of ictal automatism: a study of temporal lobe attacks. *Brain.* 1958;81:505–28.
32. Marshall DW, Westmoreland BF, Sharbrough FW. Ictal tachycardia during temporal lobe seizures. *Mayo Clin Proc.* 1983;58:443–6.
33. Smith P, Howell SJ, Owen L, Blumhardt LD. Profiles of instant heart rate during partial seizures. *Electroencephalogr Clin Neurophysiol.* 1989;72:207–17.
34. Robinson BF, Epstein SE, Beiser GD, Braunwald E. Control of heart rate by the autonomic nervous system. Studies in man on the interrelation between baroreceptor mechanisms and exercise. *Circ Res.* 1966;19:400–11.
35. Meisel C, Bailey K. Identifying signal-dependent information about the preictal state: a comparison across ECoG, EEG and EKG using deep learning. *EBioMedicine.* 2019;45:422–31.
36. Brinkmann BH, Wagenaar J, Abbot D, Adkins P, Bosshard SC, Chen M, et al. Crowdsourcing reproducible seizure forecasting in human and canine epilepsy. *Brain.* 2016;139:1713–22.
37. Kuhlmann L, Karoly P, Freestone DR, Brinkmann BH, Temko A, Barachant A, et al. Epilepsyecosystem.org: crowd-sourcing reproducible seizure prediction with long-term human intracranial EEG. *Brain.* 2018;141:2619–30.
38. Baud MO, Rao VR. Gauging seizure risk. *Neurology.* 2018;91:967–73.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Meisel C, El Atrache R, Jackson M, Schubach S, Ufongene C, Loddenkemper T. Machine learning from wristband sensor data for wearable, noninvasive seizure forecasting. *Epilepsia.* 2020;61:2653–2666. <https://doi.org/10.1111/epi.16719>