Forecasting seizure risk in adults with focal epilepsy: a development and validation study

Proix, Timothée; Truccolo, Wilson; Leguia, Marc G; Tcheng, Thomas K; King-Stephens, David; Rao, Vikram R; Baud, Maxime O

Forecasting seizure risk in adults with focal epilepsy: a development and validation study

Timothée Proix[1,2], PhD, Wilson Truccolo[1], PhD, Marc G. Leguia[3], PhD, Thomas K. Tcheng[4], PhD, David King-Stephens[5], MD, Vikram R. Rao[6*], MD & Maxime O. Baud[3,7*†], MD

[1] *Department of Neuroscience, Carney Institute for Brain Science, Brown University, Providence, Rhode Island, United States*

[2] *Department of Basic Neurosciences, Faculty of Medicine, University of Geneva, Geneva, Switzerland*

[3] *Sleep-Wake-Epilepsy Center and Center for Experimental Neurology, Department of Neurology, Inselspital Bern, University Hospital, University of Bern, Switzerland*

[4] *NeuroPace, Inc., Mountain View, California, United States*

[5] *California Pacific Medical Center, San Francisco, United States*

[6] *Department of Neurology and Weill Institute for Neurosciences, University of California, San Francisco, United States*

[7] *Wyss Center for Bio- and Neuro-technology, Geneva, Switzerland.*

*Senior authors contributed equally

†*Corresponding author: maxime.baud.neuro@gmail.com; Department of Neurology, Bern University Hospital, Freiburgstrasse 8, CH-3010 Bern*

Research in context

Evidence before this study:

We searched the literature on seizure prediction in MEDLINE from January 1, 1946, to June 1, 2020, in Embase from January 1, 1974, to June 1, 2020, and in Google Scholar (first 200 relevant references) using comprehensive electronic search strategies combining terms “epilepsy”, “seizures”, “prediction”, “forecasting”, “cycles”, “patterns”, “circadian”, and “multidien”, with no language restrictions. Identified studies used different outcome measures, but most involved analyses of electroencephalography (EEG) to predict seizures minutes in advance, with variable success. A single prospective trial of an implanted device for chronic EEG demonstrated above-chance accuracy of warnings for imminent seizures in 9 out of 15 enrolled subjects. Two studies in independent cohorts of subjects chronically implanted with intracranial electrodes showed that rates of interictal epileptiform activity oscillate in circadian and multiday (multidien) cycles that help determine seizure likelihood. Circadian cycles and seizure diaries were used in three studies to forecast seizures over short horizons, but we found no results on forecasting seizures several days in advance.

Added value of this study:

In a large cohort of people with drug-resistant focal epilepsy who had chronic EEG recorded by an approved clinical device, we demonstrate that circadian and multidien cycles can be leveraged to forecast seizures up to three days in advance in some subjects and 24 hours in advance in the majority of subjects. These results highlight the feasibility of seizure forecasting over horizons longer than previously possible.

Implications of all the available evidence:

Seizures are not entirely random events. Using cyclical patterns of brain activity to forecast seizures hours to days in advance may enable novel seizure warning systems and prevention strategies. Convergence of findings from multiple independent datasets suggests the generalizability of this approach in people with epilepsy, though this will require direct testing in prospective clinical trials.

Abstract

Background

For people with epilepsy, much suffering stems from the apparent unpredictability of seizures. Recently, converging evidence from studies using chronic electroencephalography (cEEG) revealed that brain activity in epilepsy demonstrates robust cycles, operating over hours (circadian) and days (multidien), which help determine fluctuating seizure risk. We hypothesized that cycles of brain activity can be leveraged to estimate future seizure probability, and we tested the feasibility of forecasting seizures days in advance.

Methods

This feasibility study involved retrospective analysis of cEEG (≥ 6 months; recorded between January 2004 and May 2018) collected with an FDA-approved implanted device in 175 adults with drug-resistant focal epilepsy followed at 35 centers across the USA. In distinct development and validation cohorts, subjects had ≥ 20 electrographic and disabling clinical (self-reported) seizures, respectively. In all subjects, the device stored interictal epileptiform activity (IEA) that revealed cycles of abnormal brain activity. Point process statistical models trained on initial portions of each subject's data generated forecasts of seizure probability that were tested on subsequent unseen data and evaluated against surrogate time-series. The primary outcome was the percentage of subjects with forecasts showing improvement over chance (IoC).

Findings

Models incorporating information about IEA cycles generated daily seizure forecasts with IoC in 15/18 (83%) subjects and 104/157 (66%) subjects in the development and validation cohorts, respectively. In many subjects, the forecasting horizon could be extended up to three days. Hourly forecasts, possible only in the development cohort, showed IoC in 18/18 (100%) subjects.

Interpretation

Seizure probability can be reliably forecasted days in advance using data from an approved device. For adults with focal epilepsy, personalized risk-stratification over days is unprecedented and may enable novel seizure prevention strategies. This study paves the way for prospective clinical trials that will establish how people with epilepsy may benefit from long-horizon seizure forecasting.

Funding

None.

Introduction

Epilepsy is defined by the seemingly random occurrence of spontaneous seizures. Although seizures are typically brief events that cumulatively amount to a small fraction of time, their unpredictability necessitates standing treatments and causes significant disability.[1] People with epilepsy are plagued by constant uncertainty, and the looming threat of seizures has implications for personal safety, independence, and psychological well-being. Reliable methods to anticipate seizures would mark a paradigm shift in clinical epilepsy, mitigating this uncertainty and enabling time-varying, risk-based seizure prevention strategies.

Despite decades of progress in the field of seizure prediction, such methods remain elusive.[2] The landmark NeuroVista trial[3] demonstrated feasibility of a cEEG-based advisory system that warned of seizures minutes in advance. Subsequent analyses[4-10] of data from this trial yielded numerous transformative insights that propelled the field for years. However, limitations of these pioneering efforts include the relatively small size of the trial—ten subjects participated in the seizure advisory phase—and the fact that the implanted device used is no longer available.

In the decade since the NeuroVista trial, cEEG from another device (RNS® System), one that is FDA-approved and increasingly used in clinical care for epilepsy,[11] revealed pervasive daily (circadian[12,13]) and multi-day (multidien[13]) cycles of interictal epileptiform activity (IEA) that are biomarkers of seizure risk.[13] With these long-timescale biomarkers,[14-16] interest has recently shifted to probabilistic approaches to seizure forecasting,[5,17,18] akin to weather forecasting, which leverage prior knowledge about cyclical patterns of seizure risk to estimate seizure probability over future time horizons. Since most prior work in the field has sought to identify seizure precursors in the minutes preceding seizure onset,[2,8] an unresolved question concerns whether periods of heightened seizure risk (pro-ictal states[18]) can be anticipated over longer horizons. We hypothesized that seizure probability is determined by alignment of cyclical influences at multiple timescales as well as the temporal distribution of recent seizures.[4,9,10,18] Models that incorporate these factors to generate seizure risk forecasts will ultimately require validation in large, prospective clinical trials. Imminent feasibility of such trials hinges on evaluating how generalizable and valuable the approach may be on existing cEEG data, using a probabilistic framework that accounts for the fact that seizures may not occur every time risk is accurately forecasted to be high.

Here, we address these questions in a feasibility study aimed at developing and validating statistical models to forecast the individual risk for electrographic and self-reported seizures based on temporal features extracted from up to ten years of cEEG data. The primary study outcome was the percentage of subjects for whom forecasting models demonstrated Improvement over Chance (IoC) at different forecasting horizons. Secondary outcomes involved quantifying model performance using statistical methods suitable for probabilistic forecasts.

Methods

Study design and participants

This feasibility study involved development of seizure forecasting models in a 'development cohort' of 18 subjects who were implanted with the RNS® System (NeuroPace, Inc., Mountain View, CA, USA) for clinical indications and followed at two centers (University of California, San Francisco, and California Pacific Medical Center, USA). Forecasting models were subsequently validated by including cEEG data and self-reported seizures obtained from a 'validation cohort' of 157 participants in the nine-year long-term treatment trial (LTT) of the RNS System[11,19] that took place between January 2004 and May 2018 across 34 centers in the USA (appendix, pp 10–11; ClinicalTrials.gov identifiers: NCT00079781, NCT00264810, and NCT00572195). All involved centers obtained authorization from their institutional review board to recruit adults with medically-refractory focal epilepsy in the original trials and for subsequent data analysis. Existing cEEG data and seizure logs were screened for eligibility: > 6 months of continuous hourly IEA count data without large gaps and ≥ 20 electrographic or self-reported seizures but < 50% days with seizures, as the utility of forecasting in individuals with very frequent seizures is likely low.[3] All 175 included subjects provided written informed consent for analysis of their data.

Procedure

The RNS System utilizes customizable algorithms to detect pathological brain activity, as previously detailed.[13] For each subject, IEA time-series from two RNS System detectors (appendix, pp 10–16) were selected for periods of continuous data with stable detection settings > 6 months. For all subjects, the first few months of cEEG (median [range] 222 d [28–362 d]) after device implantation were discarded to account for time needed by clinicians to optimize detection parameters.

Self-reported seizures are the current gold-standard for clinical trials in epilepsy, but electrographic seizures evident on cEEG are more objective and obviate subjects' reporting biases.[20] Therefore, we examined two types of seizures drawn from distinct, non-overlapping cohorts of subjects: (1) Timestamps of electrographic seizures from cEEG in the development cohort (N=18), identified through detections of prolonged epileptiform activity exceeding a clinically pre-specified duration, typically 15−40s. For each subject and each period of stable detection settings, a Board-certified epileptologist (V.R.R.) verified visually that ≥90% of these detections corresponded to electrographic seizures in stored electrocorticograms, as described in detail previously[13]; (2) Diaries of self-reported seizures in the validation cohort (N=157), recorded by participants in the LTT as number of seizures ('simple motor', 'simple other', 'complex partial', and 'generalized tonic-clonic') per calendar day. According to the 2017 International League Against Epilepsy classification, we considered 'complex partial' and 'generalized tonic-clonic' as the disabling 'seizures with impaired awareness' studied here and

excluded subjects without disabling seizures. As subjects did not report the time of day for their seizures, these data could only be used for daily and not for hourly forecasts.

Statistical analysis

Forecasting models: To forecast seizure probabilities—continuous values between 0 (no risk of seizure occurrence) and 1 (seizure occurrence is certain)—we used past IEA, occurrence times of past seizures, and cyclical variables (hereafter, collectively referred to as 'temporal features') as inputs for point process generalized linear models (PP-GLMs). Models were estimated on training data and evaluated on chronologically subsequent test data. PP-GLMs are established tools in neuroscience research[21,22] that provide a flexible statistical framework to evaluate the association between sequences of event (seizure) times represented as binary (or count) time-series and temporal features upon which event probability may depend. Hourly IEA time-series were available for electrographic and self-reported seizure cohorts, allowing determination of circadian and multidien cycles of epileptic brain activity in all subjects (N=175).[13] We trained PP-GLMs with a log-link function and a conditionally Poisson distribution[22] to output the probability of a seizure as a function of these cycles and other temporal features (appendix, pp 9–10) from subject-specific datasets comprising the shorter of 480 d or 60% of the subject's total data. To prevent inflation of our performance metrics (see below) through the well-known phenomenon of seizure clustering, we defined 'seizure-days' or 'seizure-hours' as binary events regardless of the seizure count and used these as training labels. The large amount of previously unseen testing data (Individually: minimum of 40% of data, >800 d in most subjects and up to 8 y; In total: 73% of data with 211,005 d) ensured that the models were not overfit for a small number of seizures and enabled assessment of forecasting performance in a probabilistic framework.

Outcomes: Subject-specific forecast performance was quantified on held-out test datasets (i.e. index test) containing unseen seizures (i.e. reference standard: days or hours with self-reported or electrographic seizures) using two complementary metrics that are fully described in the appendix (pp 1–9): (i) for various seizure warning threshold probabilities, the area under the curve (AUC) of sensitivity (proportion of all seizures captured during warning) vs. corrected proportion of time in warning;[23] (ii) Brier skill score (BSS), adapted from meteorology,[5,17] which evaluates performance in relation to a naïve predictor (here, a randomly shuffled forecast).

Based on these metrics, the primary outcome was Improvement over Chance (IoC), a binary outcome defined at the individual level through comparison of the original AUC to chance-level AUC, calculated from forecasts issued on surrogate data (see Statistical significance).[2,5,8,23] In addition, secondary outcomes involved quantifying individual forecast performance. We calculated the median AUC across subjects with IoC (and the entire cohort) to evaluate discrimination, the goal of a deterministic forecast (do forecasts differ when their corresponding observations differ? see appendix, pp 1–9). AUC depends heavily on forecast horizon and pro-ictal state duration, and AUC is less than 1 even for a reliable forecast (appendix, p 8). This motivated the additional use of the BSS, which assesses model resolution (are different forecasts associated with different outcomes?) and calibration (how close are

forecasted probabilities to observed probabilities?), the goals of a probabilistic forecast (appendix, pp 1–9). Reliability diagrams[5] were used to compare observed and forecasted seizure probabilities.

Post-hoc analyses: To characterize individualized forecasts in terms of time-varying risk, we defined pro-ictal states as periods of time with forecasted probability above the individual expected seizure probability (appendix, pp 1–9). Based on these adjusted values, we report the average duration of pro-ictal states and the relative risk for seizures in pro-ictal as compared to low-risk states. We evaluated the overlap between forecasted probabilities and observed seizures as a function of circadian and multidien cycle phases and AUC as a function of the strength of seizure cycles, quantified as the phase-locking value (appendix, p 24).[13]

Sensitivity analyses: Robustness of our results was assessed by systematically varying the amount of training data and the retraining interval (appendix, pp 21–23).

Statistical significance: To determine individual chance-level AUCs, 200 surrogates were generated for each temporal feature: (1) for the recent seizure, circadian, and weekly distribution models, by randomly shuffling the seizure time-series under the null hypothesis that the seizure process is memoryless (i.e. events are independent of one another); (2) for the IEA-based features, by randomizing phases of underlying cycles, under the null hypothesis that seizure timing does not depend on trends in IEA.[24,25] Significance was assessed with a false discovery rate (FDR) at $\alpha = 0{\cdot}05$ across all subjects to correct for multiple testing. As a supplementary statistical analysis, significance of AUC was assessed by comparing the number of seizures correctly identified by the model and by chance for a given fraction of time under warning (appendix, p 28).[23,26] Analyses were performed with MATLAB R2019a, R 3·4·4, and Python 3·7·4.

Role of funding source

This study received no targeted funding. All authors had full access to data and had responsibility for the decision to submit for publication.

Data sharing

Deidentified individual data in the form of IEA counts and electrographic seizures from the 18 subjects in the development cohort, as well as code created and used for this paper, will be freely available at **#DOI will be made available upon acceptance#** as of July 1, 2021 for at least 20 years. A short explanation of the data is also provided. Original study protocols, statistical analysis plan, and informed consent forms are not available for this retrospective study. Data for the validation (self-reported seizure) cohort is property of NeuroPace, Inc. and is not available.

Results

We collected retrospective cEEG data from two cohorts of subjects implanted with the RNS System (Fig. 1a, mean duration per subject 1484 d, range 227–3502 d). Between January 1, 2018 and October 1, 2019, we screened 72 and 256 subjects, and we included 18 and 157 subjects in distinct cohorts for forecasting model development and validation, respectively (flow diagram). The development cohort comprised 10 subjects whose cEEG data we previously published[13] but here extended to include two years of subsequent recordings, plus 8 new subjects. The validation cohort comprised a subset of participants in the nine-year RNS System Long-term Treatment Trial (LTT)[19], from which only limited cEEG data has been published[12]. cEEG data (Fig. 1b) from this cohort was used to validate forecasting models, which were then tested against the published dataset of self-reported seizures from the LTT[19]. Baseline characteristics for the two cohorts were similar, with median age 38 [IQR 32–51] and 35 [IQR 25–43], and 44% (8/18) and 47% (74/157) females, respectively, with a preponderance of multifocal and mesio-temporal epilepsies (Table 1).

In both cohorts, forecasting models were individually estimated on the first portion of each subject's data, the 'training datasets', and tested on non-overlapping individual 'testing datasets' containing a total of 767 electrographic (median 19% [IQR 13–29] days with seizures) and 27,658 self-reported seizures (median 9% [IQR 5–16] days with seizures) that were previously unseen (see Methods). To forecast seizure probability with horizons of hours to days, models incorporated past IEA, occurrence times of past seizures, and cyclical variables as inputs (hereafter, 'temporal features;' Fig. 1c, Table 2). Individual subjects had excellent correspondence between forecasts and seizures (Fig. 1d–h).

As a primary outcome, and for each temporal feature, we determined which subjects might benefit from forecasting with our models by calculating improvement over chance (IoC: AUC relative to chance-level, Table 2). Daily forecasts incorporating information only about recent seizures, weekly seizure distribution, or recent IEA produced IoC less often than models using information from multidien IEA cycles, for which IoC was observed in 15/18 (83%) and 104/157 (66%) subjects for electrographic and self-reported seizures, respectively (Fig. 2a; Table 2). With multidien IEA cycles alone, the forecast horizon could be extended up to three days while maintaining IoC in 2/18 (11%) and 61/157 (39%) subjects for electrographic and self-reported seizures, respectively (Fig. 2b).

As secondary outcomes, we quantified forecast performance for subjects with IoC using two complementary metrics, each addressing a distinct question (appendix, pp 1–9): (i) Area under the curve (AUC, sensitivity vs. corrected time in warning)—How valuable is a forecast given the amount of time spent in warning?, and (ii) Brier skill score[5,17]—How well does the forecast perform relative to a reference strategy (BSS = 1 for perfect forecast; BSS = 0 for no improvement over a random predictor)? Median AUC was 0·74 [IQR 0·70-0.79] and 0·70 [IQR 0·65-0·75], and median BSS was 0·23 [IQR 0·18-0·30] and 0·13 [IQR 0·05-0·20] in the development (electrographic seizures) and validation (self-reported seizures) cohorts, respectively (Fig. 2a, c; all median values in Table 2; appendix, pp 17–18). A reliability diagram[5]

showed that resolution (Fig. 2d, highest bin average is below 1) and calibration were good, but not perfect, with forecasted probabilities above 25% being overconfident (i.e. below the diagonal line of perfect calibration, Fig. 2d).

As a post-hoc analysis, we characterized the durations of forecasted pro-ictal states, i.e. the tendency for daily forecasts to remain high over consecutive days (Fig. 1f). To allow for comparison across subjects, we averaged peak-aligned forecasts centered within subjects around expected seizure probability (appendix, pp 1–9). This enabled visualization of pro-ictal states as contiguous periods of heightened seizure probability lasting 3–9 d and aligning well with the distributions of observed electrographic and self-reported seizures (Fig. 3). Average relative risk (RR) for self-reported and electrographic seizures occurring during forecasted pro-ictal versus low-risk states was 9·4 [95% CI 4·5–14·9] and 3·7 [95% CI 2·8–4·7]) across subjects with IoC. Model performance also correlated with phase-locking values between seizures and multidien IEA cycles[13] (Pearson r=0·6547±2·7×10^{-3}, Wald test, p<0·0001; appendix, p 24), suggesting that the most forecastable individuals can be identified in advance.

To further characterize performance of daily forecasts, we carried out sensitivity analyses to inclusion criteria (appendix, pp 19–20) and to training conditions. In both cohorts, longer training duration and iterative retraining (appendix, p 23), improved model performance and the calibration of output forecast probability (Fig. 2d; appendix, pp 21–22).

Forecasting days-long pro-ictal states over long horizons may not be ideal for all patients,[27,28] so we asked whether our approach allows refinement of forecasts to shorter horizons. Equivalent outcomes were obtained for hourly forecasting, which was only possible for electrographic seizures, as subjects in the validation cohort reported seizure days but not hours. Multivariate models incorporating instantaneous phases of circadian and multidien cycles and the recent circadian distribution of seizures[5] yielded the best-performing hourly forecasts of electrographic seizures (Fig. 4a, c; appendix, p 25), and IoC was observed in 18/18 subjects (100%; Table 2; appendix, pp 25–26). The forecasting horizon could be extended up to 14 h while maintaining IoC in 8/18 subjects (44%; Fig. 4b). Across subjects, highest forecasted seizure probabilities occurred when critical phases of multidien and circadian cycles aligned (Fig. 4d).

Discussion

Here, we forecasted electrographic seizures and self-reported seizures—a gold standard metric for clinical trials in epilepsy—up to three days in advance. To our knowledge, this represents an unprecedented horizon for personalized seizure risk-stratification. Daily forecasts were above chance in the majority of the 175 adults with focal epilepsy involved in this feasibility study (15/18 and 104/157 in development and validation cohorts, respectively). In all subjects for whom it was possible (18/18), forecasts of electrographic seizures achieved finer temporal resolution on the scale of hours. Included subjects were treated with an implanted neurostimulation device and may not be representative of all people with epilepsy, though diverse focal epilepsies were represented in our cohorts, and seizure cycles are independent of brain stimulation.[13,15]

To date, there has been only one prospective trial (NeuroVista[3]) of a seizure advisory system, which provided short-term (minutes) warnings of imminent seizures demonstrating above-chance accuracy in 9 out of 15 (60%) enrolled subjects (10 of these subjects completed a 4-month testing period). Subsequent analyses on the same dataset showed that even the most difficult cases were predictable to some extent through crowd-sourced computational efforts.[6-8] In comparison, our feasibility study involved ten times more subjects, testing and validating a single computational approach for periods up to 10 years, and forecast horizons several orders of magnitude longer (hours to days).

To evaluate forecasting model performance rigorously, we comprehensively report measurements of risk, discrimination, resolution, and calibration (explained in appendix, pp 1–9). During forecasted pro-ictal states, the average RR of occurrence of electrographic and self-reported seizures was 9·4 and 3·7, respectively, placing cycles of epileptic brain activity among the strongest predictors of seizures discovered to date. While RR is a well-established metric in medicine, it is limited to the evaluation of probabilistic forecasts at a single threshold value, whereas the BSS circumvents this limitation, offering a refined interpretation of forecast performance as a continuum (appendix, pp 1–9). A recent study in nine subjects employed probabilistic methods similar to ours within a circadian framework and yielded BSS ranging 0·02–0·2 at a forecast horizon of one minute.[5] In comparison, our study provided well-calibrated forecasts, as illustrated in a reliability diagram (Fig. 2b), and median BSS of 0·23 [IQR 0·18–0·30] (electrographic seizures) and 0·13 [IQR 0·05–0·20] (self-reported seizures). Another key distinction of our work is that daily forecasts of higher seizure probabilities were aggregated over days-long pro-ictal states (Fig. 3), providing smooth forecasts rather than flickering alerts based on real-time detection of evanescent seizure precursors. This may improve the interpretability of forecasts for people with epilepsy.[3]

This study has limitations. Implanted devices are associated with surgical risks and may not be suitable for all people with epilepsy who desire seizure forecasts, motivating development of minimally-invasive methods to monitor seizure risk biomarkers.[1,4] Cycles of IEA may be more tractable than biomarkers requiring high sampling rate intracranial EEG,[29] opening the possibility that certain novel methods, like sub-scalp EEG,[30] could be viable for forecasting.

Our models did not incorporate common seizure triggers, such as medication non-compliance, which could account for some apparent 'false negatives.' Self-reported seizure data was drawn from a large, prospective, nine-year clinical trial[11,19]—arguably the most well-curated clinical seizure dataset of this chronicity—but inaccuracy of seizure self-reports[3,20] and small gaps in the data could have led to under-estimation of model performance. Finally, to dissect the potential contribution of different temporal features, this feasibility study focused on explicit statistical models that are computationally efficient, modest in their training requirements, and incorporated cycles of IEA using an accurate but non-causal estimation of the instantaneous phases. Thus, conclusions should be regarded as hypothesis-generating rather than clinical evidence.[8]

In summary, our results corroborate an emerging view that seizures are not entirely random events.[8] Given the large sample size, these results validate and powerfully extend our previous findings based solely on electrographic seizures,[13] and they suggest the generalizability of using multiscale cyclical biomarkers in epileptic brain activity to forecast clinically-relevant seizures over long horizons. Moreover, our study indicates that seizure forecasting is feasible with existing neurotechnology in widespread clinical use (~3,000 patients currently implanted in the U.S.) and need not await novel industrial developments. Future prospective clinical trials should assess directly the ways in which people with epilepsy benefit from replacing constant uncertainty about seizures with "measured uncertainty" (forecasted risk) at different horizons, which has not been established by this or prior studies. To that end, we propose a nested approach to personalized seizure forecasting: (1) patient-specific multidien cycles reveal pro-ictal states days in advance; (2) circadian IEA cycles and peak seizure times reveal hours of high risk;[5] and (3) real-time detections of seizure precursors[2] provide imminent seizure warnings conditioned on prior probability from (1) and (2). Future work will also involve miniaturization of devices, integration of cEEG with multimodal physiological data,[1] optimization of forecasting models, and elucidation of mechanisms underlying cycles in epilepsy.

Legends

	Development cohort	Validation cohort
N	18	157
Age in years (median [IQR])	38 [32-51]	35 [25-43]
Percent females	44% (8/18)	47% (74/157)
Percent males	56% (10/18)	53% (83/157)
Seizure studied	Electrographic seizures	Self-reported disabling seizures
Bilateral focus	50% (9/18)	46% (73/157)
Left-sided focus	33% (6/18)	39% (62/157)
Right-sided focus	17% (3/18)	14% (22/157)
Mesiotemporal lobe epilepsy	83% (15/18)	64% (101/157)
Frontal lobe epilepsy	0% (0/18)	9% (14/157)
Multilobar epilepsy	6% (1/18)	12% (19/157)
Other neocortical epilepsy	11% (2/18)	15% (23/157)
Percentage of days with seizures in training datasets (median [IQR])	25% [17-29]	15% [10-25]
Percentage of days with seizures in testing datasets (median [IQR])	19% [13-29]	9% [5-16]

Table 1. Demographics and seizure characteristics of all subjects in the development and validation cohorts.

		Forecasts								
Horizon		**Daily**						**Hourly**		
Cohort		**Development (N=18)**			**Validation (N=157)**			**Development (N=18)**		
Reference standard		**Electrographic seizures**			**Self-reported disabling seizures**			**Electrographic Seizures**		
Study outcome		**1°**	**2°**		**1°**	**2°**		**1°**	**2°**	
Metric		**IoC**	**AUC**	**BSS**	**IoC**	**AUC**	**BSS**	**IoC**	**AUC**	**BSS**
Temporal features	Recent seizures	2/18 (11%)	0.62 (0.60)	0.06 (0.03)	43/157 (27%)	0.58 (0.57)	0.012 (0.009)	6/18 (33%)	0.57 (0.52)	0.002 (0.00)
	Recent IEA	0/18 (0%)	NA (0.61)	NA (0.02)	51/157 (32%)	0.62 (0.58)	0.04 (0.01)	5/18 (28%)	0.64 (0.60)	0.008 (0.006)
	Circadian IEA phases	NA	NA	NA	NA	NA	NA	8/18 (44%)	0.65 (0.62)	0.01 (0.01)
	Circadian seizure distribution	NA	NA	NA	NA	NA	NA	11/18 (61%)	0.62 (0.59)	0.008 (0.002)
	Weekly seizure distribution	0/18 (0%)	NA (0.54)	NA (0.00)	0/157 (0%)	NA (0.56)	NA (0.004)	NA	NA	NA
	Multidien phases	15/18 (83%)	0.74 (0.73)	0.23 (0.17)	103/157 (66%)	0.70 (0.66)	0.13 (0.07)	15/18 (83%)	0.70 (0.70)	0.024 (0.018)
	Multivariate	NA	NA	NA	NA	NA	NA	18/18 (100%)	0.75 (0.75)	0.036 (0.035)

Table 2. Primary and secondary study outcomes. Percentage of subjects with Improvement over Chance (IoC) for different temporal features, where IoC was obtained by comparing the area under the curve (AUC) of the original data with the AUC of 200 surrogate time-series with alpha < 0.05 adjusted for false-discovery rate correction. Data are median AUC and median Brier skill score (BSS) among subjects with IoC (entire cohort). 1°: primary, 2°: secondary.

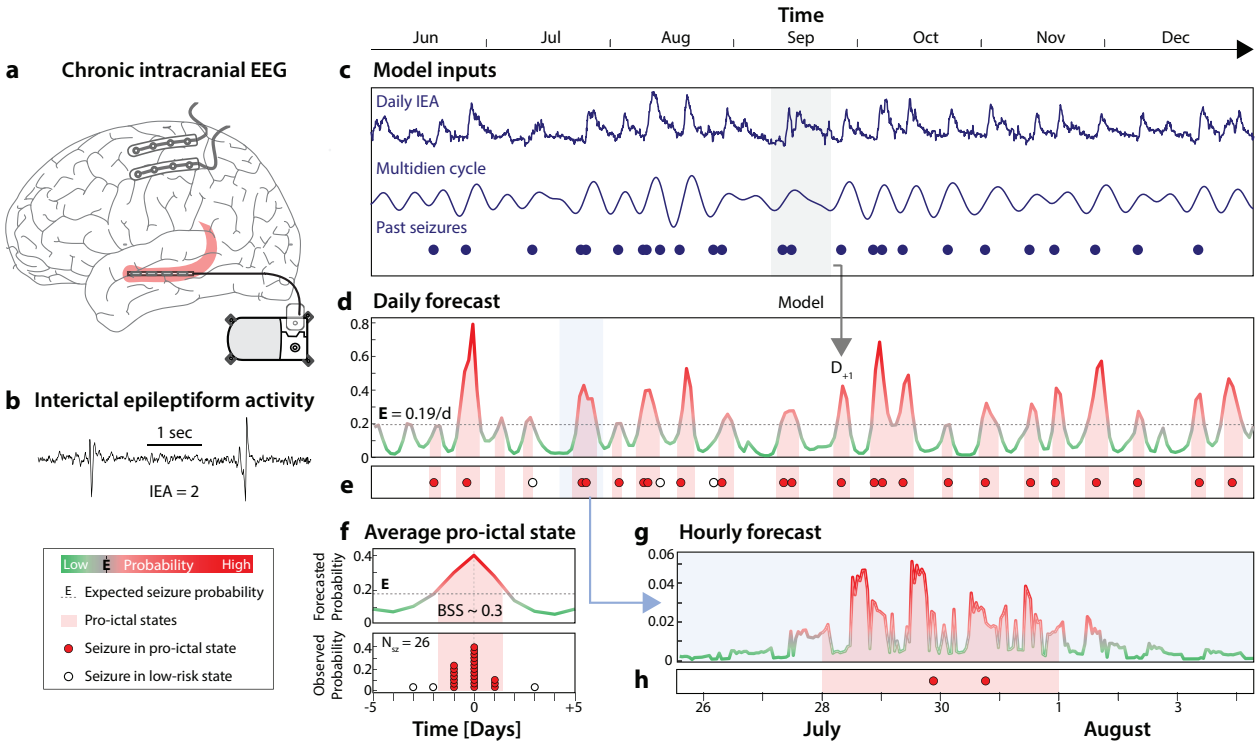

Figure 1. Individual seizure risk forecasting in one subject. (**a**) Responsive Neurostimulation (RNS®) System, comprising a cranially-implanted neurostimulator connected to two four-contact intracranial depth leads (shown, for example, in hippocampus, red) and/or cortical strip leads (shown unconnected) that provide chronic electroencephalography (cEEG). (**b**) From these recordings, the RNS System provides hourly counts of detections of interictal epileptiform activity (IEA) and electrographic seizures (not shown). (**c-e**) Entire test dataset from one subject (S7) showing input temporal features, output daily forecasts, and observed seizures. (**c**) Time-series of IEA averaged over one calendar day ('daily IEA'), underlying multidien cycle, and electrographic seizures that serve as some of the input temporal features for the forecasting model. (**d**) Daily forecast of seizure probability (gradient-colored lines) at 24-hour horizon (D_{+1}) generated by a model (grey arrow) trained on ten months of data (not shown) and run on seven months of held-out test data (shown here) using input variables from **c**. Higher forecasted probabilities (red) form days-long pro-ictal states (red shadow) during which daily probability of seizures is continuously above the expected probability, defined as the long-term average daily seizure frequency calculated over months of training data ('E', here 0·19 seizures per day). (**e**) Seizures observed during and outside of pro-ictal states over these seven months. (**f**) Average pro-ictal state illustrated by peak-aligned average probability forecasts (top) and corresponding temporal distribution of seizures (bottom, shown as stacked individual events and percentage of total count on y-axis). (**g**) Hourly forecasts of seizure probability based on hourly IEA and its circadian cycle (not shown) refining pro-ictal states into hours of relatively higher and lower seizure risk. BSS: Brier skill score. (**h**) Seizures observed over this period of nine days.

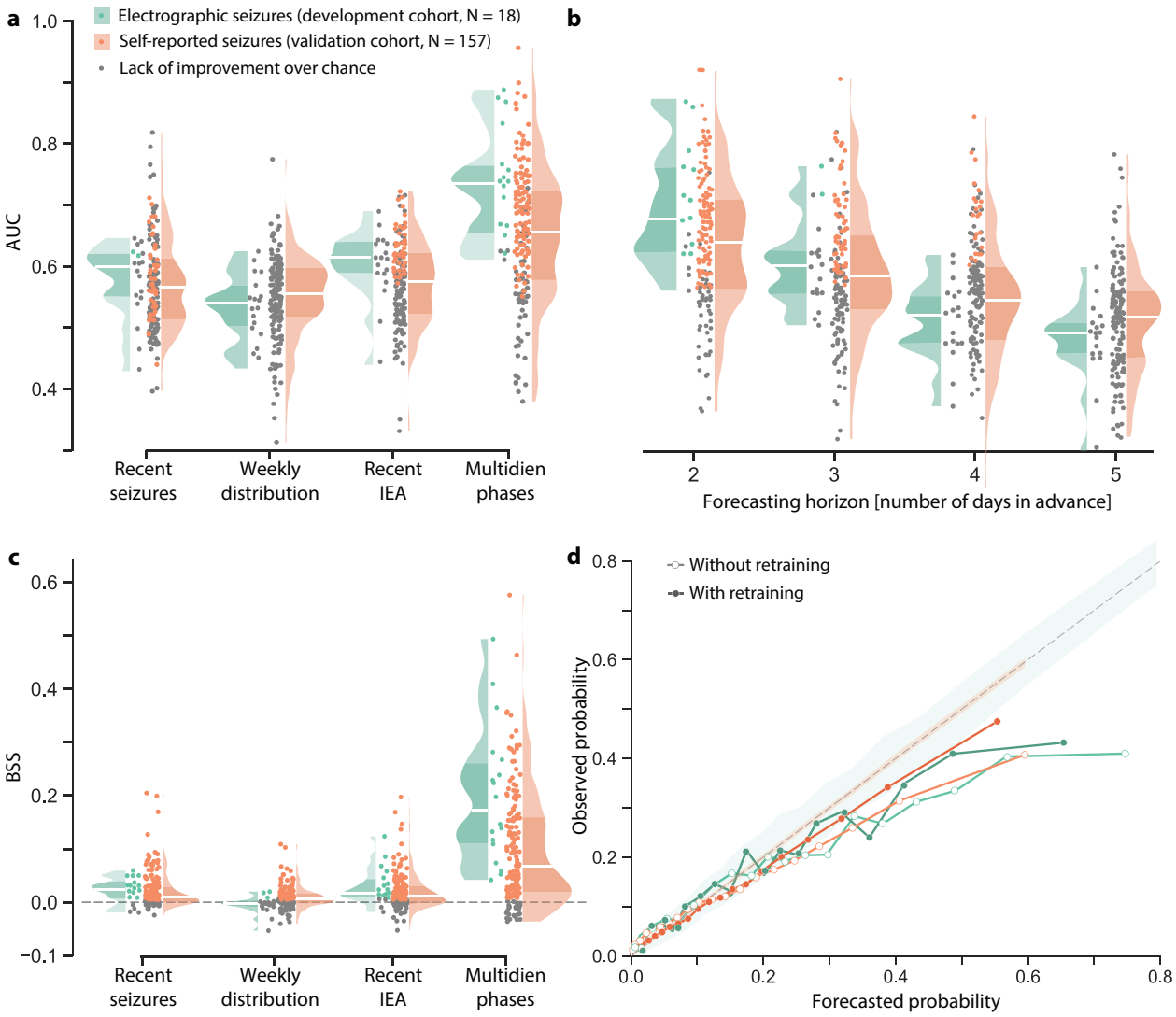

Figure 2. Performance of daily forecasts of electrographic and self-reported seizures. (**a**) Distributions of univariate daily forecast performance (at 24-h horizon) quantified as the area under the curve (AUC) across subjects. When models incorporated multidien phase information, AUC showed improvement over chance (IoC) for 83% and 66% of subjects (color dots, $p<0.05$) in the development cohort (with recorded electrographic seizures) and the validation cohort (with self-reported seizures), respectively. Shaded areas in these and subsequent violin plots show kernel density estimates to highlight the shape of the distribution of the entire cohort; darker shading is the interquartile range and horizontal white line is the median. (**b**) AUC as a function of forecasting horizon longer than 24-hour using multidien phase as the input variable, to be compared to forecast at 24-hour horizon in **a**. (**c**) As in (a), daily forecasts based on multidien phases of IEA yielded both higher AUC and Brier skill score (BSS) than other models. The BSS represents improvement (skill, color dots) of mean squared forecast error (Brier score) relative to a reference randomly shuffled forecast; BSS range is $-\infty$ to 1, with 0 being no skill relative to reference forecast and 1 being a perfect forecast. (**d**) Reliability diagram showing observed

seizure probability vs. binned forecasted probabilities of electrographic seizures (green, N=18 subjects) and self-reported seizures (orange, N=157 subjects). Empirical curves for a set of forecasts generated by models before (empty dots) and after (filled dots) re-training after every seizure are compared to the dashed diagonal line of perfect calibration (shading indicates 95% confidence intervals (CI)).

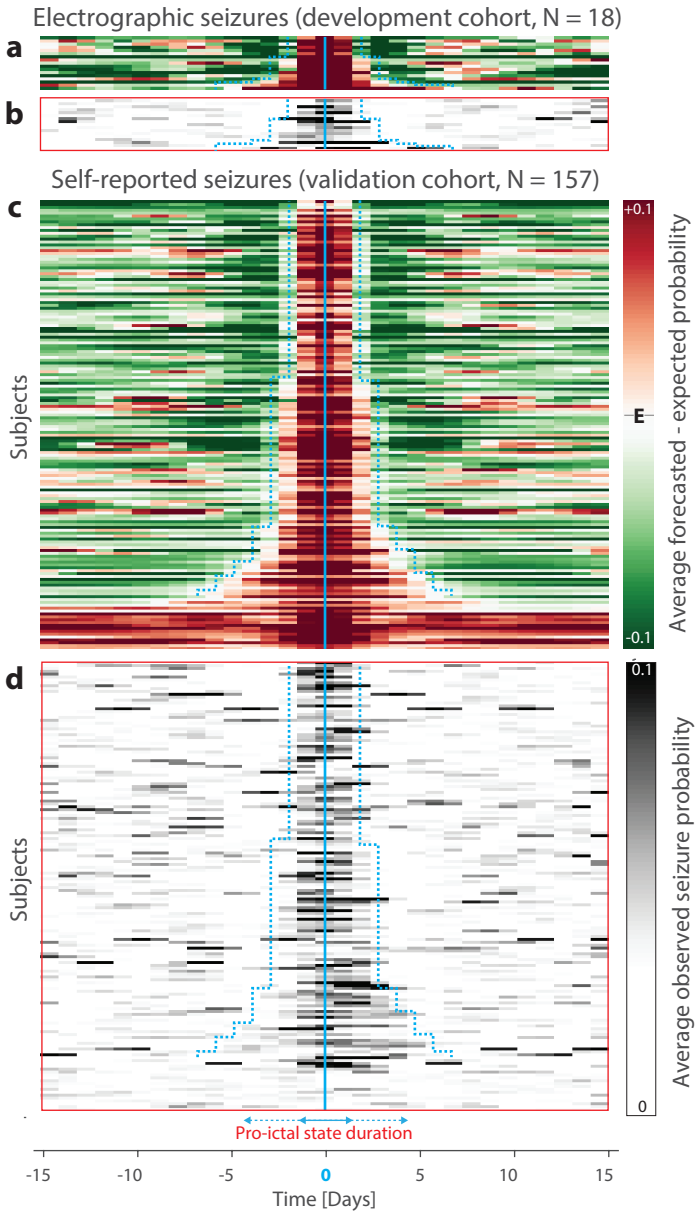

Figure 3. Pro-ictal states. (**a**) Peak-aligned normalized average forecast probabilities for all subjects in the electrographic seizures (development) cohort (N=18, rows, ranked by width of pro-ictal state) reveal days-long periods of seizure probability higher than the expected seizure probability (E). **(b)** Distributions of observed seizure probabilities averaged in the same way align well with periods of high risk. (**c**) and (**d**) show data analogous to (**a**) and (**b**) from the self-reported seizures (validation) cohort (N=157). Cyan boundaries depict estimated durations of pro-ictal states, which range from three to five days in the majority of subjects and more than seven days in a minority of subjects. Most subjects whose forecasts did not show IoC reside at the bottom, outside of the cyan boundaries.

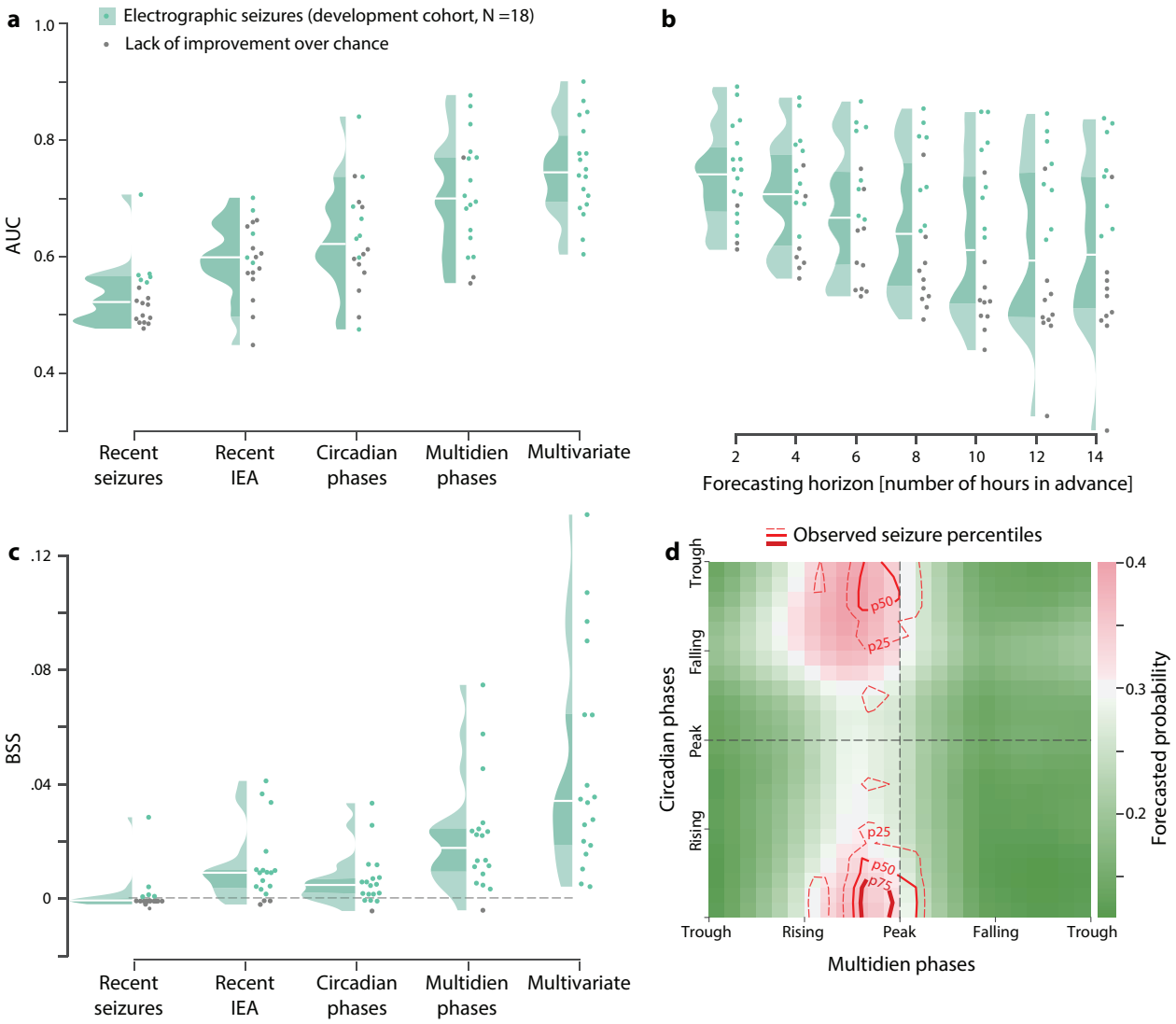

Figure 4. Hourly forecasts of electrographic seizures. Hourly forecasts were not possible for self-reported seizures because time resolution of these data was one day. (**a**) Distributions of univariate and multivariate hourly forecast performance (at 1-h horizon) quantified as the AUC across subjects. Multivariate models incorporated information from circadian and multidien phases of IEA, as well as the circadian distribution of seizures, yielding AUC with IoC in 18 out of 18 (100%) subjects (color dots, $p<0.05$). (**b**) AUC as a function of forecasting horizon hours in advance of seizures. (**c**) As in (a), multivariate models yielded both higher AUC and BSS than univariate models. (**d**) Phase-space map across 18 subjects showing alignment of critical phases of circadian and multidien cycles with observed seizures (contours represent percentiles), coinciding with times of highest forecasted seizure probability.

Online content

Any methods, additional references, reporting summaries, source data, statements of data availability, and associated accession codes are available in the online version of the paper.

Acknowledgements

W.T. is supported by the National Institute of Neurological Disorders and Stroke (NINDS), grant R01NS079533, and the Pablo J. Salame ’88 Goldman Sachs endowed Associate Professorship of Computational Neuroscience at Brown University. V.R.R. is supported by the Ernest Gallo Foundation Distinguished Professorship in Neurology at the University of California, San Francisco. M.O.B. is supported by an Ambizione Grant from the Swiss National Science Foundation and by the Velux Stiftung. Part of this research was conducted using computational resources and services at the Center for Computation and Visualization, Brown University. We thank Eva Pool and Ben Meuleman for advice with the statistical analysis.

Author contributions

M.O.B., T.P., V.R.R., and W.T. designed the study. V.R.R. and D.K.S collected the data of the development cohort. T.K.T. provided the data of the validation cohort. V.R.R., T.K.T., and M.O.B. accessed, selected, and verified the data. T.P., W.T., M.G.L., and M.O.B. performed the analysis. T.P., M.O.B., and V.R.R. wrote the manuscript, which all authors edited.

Competing interests

M.O.B reports personal fees from Wyss Center for neurotechnology as part-time employee, grants from Wyss Center for neurotechnology, outside the submitted work; In addition, Dr. Baud has a pending patent under the Patent Cooperation Treaty (#62665486). V.R.R reports personal fees from NeuroPace, Inc., outside the submitted work. T.K.T. is an employee of NeuroPace, Inc. and receives salary and stock options as compensation. T.P., W.T., M.G.L., and D.K.-S. have nothing to disclose.

References

1. Dumanis SB, French JA, Bernard C, Worrell GA, Fureman BE. Seizure Forecasting from Idea to Reality. Outcomes of the My Seizure Gauge Epilepsy Innovation Institute Workshop. *eNeuro* 2017; **4**(6).
2. Mormann F, Andrzejak RG, Elger CE, Lehnertz K. Seizure prediction: the long and winding road. *Brain* 2007; **130**(Pt 2): 314-33.
3. Cook MJ, O'Brien TJ, Berkovic SF, 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**(6): 563-71.
4. Karoly PJ, Cook MJ, Maturana M, et al. Forecasting cycles of seizure likelihood. *Epilepsia* 2020; **61**(4): 776-86.
5. Karoly PJ, Ung H, Grayden DB, et al. The circadian profile of epilepsy improves seizure forecasting. *Brain* 2017; **140**(8): 2169-82.
6. Kuhlmann L, Karoly P, Freestone DR, et al. Epilepsyecosystem.org: crowd-sourcing reproducible seizure prediction with long-term human intracranial EEG. *Brain* 2018; **141**(9): 2619-30.
7. Brinkmann BH, Wagenaar J, Abbot D, et al. Crowdsourcing reproducible seizure forecasting in human and canine epilepsy. *Brain* 2016; **139**(Pt 6): 1713-22.
8. Kuhlmann L, Lehnertz K, Richardson MP, Schelter B, Zaveri HP. Seizure prediction - ready for a new era. *Nat Rev Neurol* 2018; **14**(10): 618-30.
9. Karoly PJ, Goldenholz DM, Freestone DR, et al. Circadian and circaseptan rhythms in human epilepsy: a retrospective cohort study. *Lancet Neurol* 2018; **17**(11): 977-85.
10. Cook MJ, Varsavsky A, Himes D, et al. The dynamics of the epileptic brain reveal long-memory processes. *Front Neurol* 2014; **5**: 217.
11. Skarpaas TL, Jarosiewicz B, Morrell MJ. Brain-responsive neurostimulation for epilepsy (RNS((R)) System). *Epilepsy Res* 2019; **153**: 68-70.
12. Spencer DC, Sun FT, Brown SN, et al. Circadian and ultradian patterns of epileptiform discharges differ by seizure-onset location during long-term ambulatory intracranial monitoring. *Epilepsia* 2016; **57**(9): 1495-502.
13. Baud MO, Kleen JK, Mirro EA, et al. Multi-day rhythms modulate seizure risk in epilepsy. *Nat Commun* 2018; **9**(1): 88.
14. Baud MO, Ghestem A, Benoliel JJ, Becker C, Bernard C. Endogenous multidien rhythm of epilepsy in rats. *Exp Neurol* 2019; **315**: 82-7.
15. Gregg NM, Nasseri M, Kremen V, et al. Circadian and multiday seizure periodicities, and seizure clusters in canine epilepsy. *Brain Commun* 2020; **2**(1): fcaa008.
16. Maturana MI, Meisel C, Dell K, et al. Critical slowing down as a biomarker for seizure susceptibility. *Nat Commun* 2020; **11**(1): 2172.
17. Jachan M, Feldwisch-Drentrup H, F. P, et al. Probabilistic forecasts of epileptic seizures and evaluation by the Brier score. *4th European conference of the International Federation for Medical and Biological Engineering* 2009: 1701-5.

18. Baud MO, Proix T, Rao VR, Schindler K. Chance and risk in epilepsy. *Curr Opin Neurol* 2020; **33**(2): 163-72.
19. Nair DR, Laxer KD, Weber PB, et al. Nine-year prospective efficacy and safety of brain-responsive neurostimulation for focal epilepsy. *Neurology* 2020; **95**(9): e1244-e56.
20. Elger CE, Hoppe C. Diagnostic challenges in epilepsy: seizure under-reporting and seizure detection. *Lancet Neurol* 2018; **17**(3): 279-88.
21. Truccolo W, Hochberg LR, Donoghue JP. Collective dynamics in human and monkey sensorimotor cortex: predicting single neuron spikes. *Nat Neurosci* 2010; **13**(1): 105-11.
22. Truccolo W, Eden UT, Fellows MR, Donoghue JP, Brown EN. A point process framework for relating neural spiking activity to spiking history, neural ensemble, and extrinsic covariate effects. *J Neurophysiol* 2005; **93**(2): 1074-89.
23. Snyder DE, Echauz J, Grimes DB, Litt B. The statistics of a practical seizure warning system. *J Neural Eng* 2008; **5**(4): 392-401.
24. Andrzejak RG, Kraskov A, Stogbauer H, Mormann F, Kreuz T. Bivariate surrogate techniques: necessity, strengths, and caveats. *Phys Rev E Stat Nonlin Soft Matter Phys* 2003; **68**(6 Pt 2): 066202.
25. Mormann F, Kreuz T, Rieke C, et al. On the predictability of epileptic seizures. *Clin Neurophysiol* 2005; **116**(3): 569-87.
26. Mormann F. Seizure prediction. *Scholarpedia* 2008; **3**(10): 5770.
27. Schulze-Bonhage A, Sales F, Wagner K, et al. Views of patients with epilepsy on seizure prediction devices. *Epilepsy Behav* 2010; **18**(4): 388-96.
28. Janse SA, Dumanis SB, Huwig T, Hyman S, Fureman BE, Bridges JFP. Patient and caregiver preferences for the potential benefits and risks of a seizure forecasting device: A best-worst scaling. *Epilepsy Behav* 2019; **96**: 183-91.
29. Alvarado-Rojas C, Valderrama M, Fouad-Ahmed A, et al. Slow modulations of high-frequency activity (40-140-Hz) discriminate preictal changes in human focal epilepsy. *Sci Rep* 2014; **4**: 4545.
30. Duun-Henriksen J, Baud M, Richardson MP, et al. A new era in electroencephalographic monitoring? Subscalp devices for ultra-long-term recordings. *Epilepsia* 2020: online ahead of print; doi: 10.1111/epi.16630.