Circadian and Multi-day Rhythms in Generalized Tonic-Clonic Seizure: A Probabilistic Approach

by

Boyu Zhang

B.S., University of Rochester (2021)

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning, in partial fulfillment of the requirements for the degree of

Master of Science

at the

Massachusetts Institute of Technology

September 2023

Author .....
Program in Media Arts and Sciences, School of Architecture and Planning,
June 7, 2023

Certified by .....
Rosalind W. Picard
Professor of Media Arts and Sciences

Accepted by .....
Joseph A. Paradiso
Academic Head, Program in Media Arts and Sciences

Circadian and Multi-day Rhythms in Generalized Tonic-Clonic Seizure: A Probabilistic Approach

by

Boyu Zhang

Submitted to the Program in Media Arts and Sciences, School of Architecture and Planning,
on June 7, 2023, in partial fulfillment of the
requirements for the degree of
Master of Science

Abstract

Epilepsy is a chronic neurological disorder characterized by recurrent seizures that affect more than 50 million people worldwide, representing approximately 0.6% of the global population. This condition poses significant public health challenges, with a heightened risk of premature mortality. Underdiagnosis and undertreatment remain pervasive, particularly in low- and middle-income countries.

Studies have discovered that seizure occurrences are phase-locking to subject-specific circadian and multi-day rhythms in human physiological signals. Also, various types of epilepsy have distinctive timing patterns with respect to sleep-wake cycles. However, it remains inconclusive how sleep parameters, non-invasive ambulatory physiological signals, and seizure occurrences are quantitatively related.

We first conduct an observational study on the association between sleep parameters, including duration, efficiency, fragmentation, and regularity, and generalized tonic-clonic seizure (GTCS) occurrences on the next day. We then conduct retrospective analyses of GTCS events phase-locking to rhythms in wrist electrodermal activity (EDA), validating previous claims. Ambulatory sleep-wake cycles and EDA recorded by smart wristbands from more than 1,000 patients diagnosed with GTCS are analyzed. GTCS events are detected by an FDA-cleared algorithm on the wristband.

Thesis Advisor:
Rosalind W. Picard
Professor of Media Arts and Sciences

Circadian and Multi-day Rhythms in Generalized Tonic-Clonic Seizure: A Probabilistic Approach

by

Boyu Zhang

This thesis has been reviewed and approved by the following committee members:

Rosalind W. Picard .....

Professor of Media Arts and Sciences

Massachusetts Institute of Technology

Mona Nasseri .....

Assistant Professor of Electrical Engineering

University of North Florida

Maxime Baud .....

Department of Neurology

Bern University Hospital

Contents

List of Figures

List of Tables

Chapter 1

Introduction

According to WHO, epilepsy is the fourth most common neurological disease globally.[1] CDC statistics show that 1 in 26 individuals in the U.S. will be diagnosed with epilepsy over a lifetime.[2] While there are 36 epilepsy drugs available on the market, 1 in 3 adult patients and 20 to 25% of child patients have drug-resistant epilepsy.[3]

When having a seizure episode, patients with epilepsy (PWE) may suffer from temporary loss of consciousness, sensation, and motor control. The notorious nature of GTCS has been known for a long time, especially the correlation between high-frequency GTCS and sudden unexpected death in epilepsy (SUDEP) [24]. Based on patient surveys, the seemingly random timing of seizures is one of the worst aspects of epilepsy [22]. Unexpected episodes are disruptive, significantly hindering their daily routines. Sometimes, there may be severe secondary damage, such as having an attack while driving or sporting.

Over the years, research has shown that seizure occurrence is not entirely random. Notably, studies have reported circadian and multi-day rhythms in seizure [31, 28, 39] and how they are modulated by brain excitability [3, 45]. Subjective self-reporting seizure diaries from PWE and objective chronic electroencephalogram (EEG) monitoring are consistent with the prevalence of such rhythms [29]. Lately, researchers have also explored multi-day synchrony between seizures and other human physiological systems [30, 20]. Causal evidence is

[1] https://www.who.int/news-room/fact-sheets/detail/epilepsy
[2] https://www.cdc.gov/epilepsy/data/index.html
[3] https://www.epilepsy.com/treatment/medicines/drug-resistant-epilepsy

lacking for the entrainment of multi-day epilepsy rhythms. Nevertheless, machine learning techniques combined with biosignal rhythms have shown promising results in forecasting seizure events [3, 8, 30].

In this thesis, we study sleep-wake cycles, circadian and multi-day rhythms in physiological signals, the risk of generalized tonic-clonic seizure (GTCS), and the potential modulation between them. We envision these rhythms captured by wearable devices to empower personalized and unobtrusive GTCS forecasting technologies. Specifically, we focus on sleep-wake behaviors, electrodermal activity (EDA), and GTCS events detected by wrist-worn smart wristbands in ambulatory settings. Data from over 1,000 to 2,000 patients diagnosed with GTCS are analyzed, depending on the availability in specific chapters. The structure of the thesis is outlined below:

- **Chapter 1**: Introduction
- **Chapter 2**: Sleep Duration, Quality, Regularity, and Generalized Tonic-Clonic Seizure Risk
  - **Previous Literature**: An overview of existing literature on circadian patterns in epilepsy and cues for such patterns.
  - **Data Preprocessing**: Preparation and cleaning of sleep-wake data and GTCS alerts provided by smart wristbands.
  - **Dataset Statistics**: Descriptive statistics of data in Chapter 2.
  - **Sleep Parameters and Next-day GTCS Risk**: Statistical analyses of sleep duration, efficiency, fragmentation, regularity, and their correlation with next-day GTCS risk.
- **Chapter 3**: Cycles in Electrodermal Activity and Generalized Tonic-Clonic Seizures
  - **Previous Literature**: An overview of existing literature on circadian and multi-day physiological rhythms and the relationship with epilepsy.
  - **Data Preprocessing**: Preparation and cleaning of long-term physiological signals and detected GTCS events.

- **Dataset Statistics**: Descriptive statistics of data in Chapter 3.
- **Cycle Extraction with Wavelet Analysis**: Detailed wavelet methods to extract cycles from physiological signals.
- **GTCS Alerts Phase-locking to Physiological Cycles**: Population and individual-level GTCS phase-locking to multi-day physiological rhythms.
- **Rhythmic Patterns in GTCS Alerts**: The discovery of cycles solely from objective GTCS alerts.
- **Rethinking Phase-locking with Simulated GTCS**: A simple and robust benchmark to evaluate GTCS phase-locking.

- **Chapter 4**: Conclusions and Future Work

Chapter 2

Sleep Duration, Quality, Regularity, and Generalized Tonic-Clonic Seizure Risk

In this chapter, we conduct an observational study using existing longitudinal data to investigate the association between sleep parameters, including duration, efficiency, fragmentation, and regularity, and GTCS occurrences on the next day using wearable monitoring. We study objective GTCS events and sleep-wake cycles detected by wrist-worn wearable devices (Embrace 2, Empatica, Boston, M.A., USA) in ambulatory settings. We hypothesize that inferior sleep episodes are correlated with elevated GTCS risk in the next 24 hours. Inferior sleep episodes can be separately defined by low sleep duration, low sleep efficiency, high sleep fragmentation, and low sleep regularity. We explore these four sleep parameters one by one in this chapter.

2.1 Previous Literature

Epilepsy patients are advised to maintain adequate sleep to reduce seizure risk. However, few studies have clarified the relationship between sleep parameters and seizure risk. A bidirectional relationship [37] has been proposed previously. On the one hand, nocturnal seizures may induce wakefulness or alter sleep structures, such as delaying REM sleep.

Diurnal seizures may trigger fatigue and elevate instantaneous sleepiness. On the other hand, along with sleep-wake schedules, alterations of the electrical excitability and hormonal levels in the brain may influence subsequent seizures.

2.1.1 Circadian Patterns in Epilepsy

Circadian patterns in epilepsy are described using two time standards: with respect to the clock time and with respect to sleep-wake cycles. It is challenging to disentangle the two-time references. When timing w.r.t. clock time, various seizure chronotypes have been identified, e.g., afternoon, dawn, or midnight peaks, depending on the type of epilepsy [39, 25, 28] and patient age [19, 26]. Anderson et al. showed that day–night patterns also exist in interictal epileptiform activity (IEA) [1].

When timing seizures w.r.t. sleep, seizures occur more commonly around sleep-wake transitions [29]. Compared to generalized epilepsy, patients with focal epilepsy are more likely to have nocturnal attacks [59]. Within focal epilepsy, frontal lobe seizures occur predominantly during sleep, while the occurrences of temporal lobe seizures are more individualized [59, 25, 32]. Nonetheless, seizure rhythms timed by sleep may not be stationary. Rao et al. reported that some patients experienced a reversal in peak seizure timing: sleep-related patterns may flip to diurnal patterns, and *vice versa* [45]. Some suggested that age-related factors, e.g., changes in hormone and sleep-wake schedule, may account for such reversal as they observed distinctive epilepsy chronotypes between adults and pediatric patients [19, 26].

Moreover, EEG of PWE when asleep demonstrated that nocturnal seizure frequency and IEA counts are both higher during non-rapid eye movement (NREM) sleep than during REM sleep [36, 17]. Pavlova et al. kept five patients with generalized epilepsy under constant dim light for three days, during which sleep is distributed evenly across their circadian phases assessed by plasma melatonin [40]. None of the patients had seizures during the course of the study. The researchers observed significantly higher IEA counts during NREM sleep, and IEA counts varied with circadian phases.

2.1.2 Cues for Circadian Patterns in Epilepsy

Given the circadian rhythm in sleep-related hormones such as dim-light melatonin onset (DLMO) and cortisol fluctuations, researchers have hypothesized endogenous drivers of

seizure cycles [29, 37]. However, there is a lack of evidence that hormones are causal triggers of seizures. Multiple studies reported that the correlation between baseline melatonin concentration and having a seizure is inconclusive [29], but patient-specific increases in melatonin after seizures appeared across patients [26, 34, 4]. Also, it has been found that patients with stress-sensitive epilepsy have cortisol fluctuations synced with IEA [57, 12].

Besides sleep-related hormones, researchers are also interested in sleep-wake behavioral cues for epilepsy. There are two frequently discussed aspects: sleep deprivation and sleep regularity. Sleep deprivation is often discussed as a potential trigger of seizure, but a robust relationship with IEA is only reported among generalized epilepsy patients [47]. Rossi et al. conducted a systematic review and identified five publications investigating sleep deprivation and focal epilepsy risks, and only one study involved randomized control experiments [48]. In contrast, Dell et al. found that an average increase in sleep duration by 1.66 hours corresponded to a 27% reduction in drug-resistant focal seizure likelihood in the next 48 hours[11]. Some researchers have summarized how the "circadian timing system, secondary circadian cycles of hormone secretion, sleep and wakefulness, and recurrent environmental factors" may influence neural excitability and seizure occurrences in temporal Lobe epilepsy with observational data from humans and mice [44, 13, 9].

Recently, Stirling et al. claimed that the time of sleep and wake onset differed significantly on days with and without self-reported seizures [53]. To draw individual-level conclusions on sleep duration, onset, and offset, Stirling et al. grouped the sleep episodes of each patient based on whether the next day contained seizures or not. Wilcoxon rank-sum tests were used to compare the two groups [53]. One caveat in this approach is that, to pass the statistical tests, sleep preceding seizure days and sleep preceding non-seizure days should be significantly different. However, low-quality sleep is not the only driver of seizures, and high-quality sleep cannot completely eliminate seizures. There is a *baseline* GTCS risk for each patient regardless of sleep (see Section 3.6.2), and GTCS events may be triggered by other factors such as stress, alcohol, or missed medication.

2.2 Data Preprocessing

In this section, we describe a preprocessing pipeline for the raw data from Embrace 2 wristbands. In particular, for GTCS events detected by Embrace 2, we discuss how to (i) select alerts produced after the FDA-clearance and (ii) assess user compliance and filter for reliable true-positive alerts. For sleep episodes detected by Embrace 2, we discuss how to (i) exclude invalid sleep episodes, such as naps, and (ii) generate four sleep parameters of interest: sleep duration, fragmentation, efficiency, and regularity. We identify a subset of data eligible for analysis in this chapter after preprocessing.

2.2.1 False Positive Generalized Tonic-Clonic Seizure Alerts

An FDA-cleared convulsive seizure detection algorithm on the wristband produces objective alerts at GTCS event onset and records the timestamp and duration [38]. In an offline setting with data from epilepsy monitoring units (EMU), this proprietary algorithm has a sensitivity of 0.98, rarely missing true GTCS events. However, on average, 0.94 false-positive alerts are produced every 24 hours. False-positive alerts are often triggered by excessive wrist movements, e.g., hand clapping. In ambulatory settings, patients may learn what triggers false alerts in their daily routines over time and try to avoid them, lowering their false alarm rate.

Every time an alert is triggered, the user will have a 10-second window to respond to the wristband indicating false-positive detection. If no response is received, the wristband will automatically notify medical and emergency contacts previously registered in the system, assuming a true-positive detection. Also, the user and caregiver may later label an alert as false-positive through a patient portal without a time limit. Therefore, each GTCS alert has a true-positive (TP) or false-positive (FP) label stored in the server along with the timestamp and duration. Table 2.1 below shows some basic statistics of all alerts in the raw data.

2.2.2 User Compliance Screening

Consequently, not all TP alerts are reliable, and many users have never reported any FP alert. It is almost impossible to check all alerts and concurrent wrist physiological signals manually. Thus, to screen users for compliance and study eligibility, we first plot the distri-

Embrace 2 User Group	Number
Number of Embrace 2 users with alerts on valid dates (e.g., excluding `1970-01-01` and `0000-00-00`)	28,893
Number of Embrace 2 users with a 0% false positive (FP) rate	18,524
Number of Embrace 2 users with a 100% FP rate	767
Number of Embrace 2 users with alerts after February 1st, 2018 (FDA clearance)	24,343

Table 2.1: Basic statistics of all alerts in the raw data.

bution of TP and FP alerts. Other quantities include the usage length (time elapsed between the first and last alert in days) and total alert count. Figure 2-1 shows the distribution of (a) FP count, (b) FP rate (FP count divided by total alert count), (c) usage length, and (d) total alert count in the raw data. Abnormal spikes at 0% and 100% FP rates indicate potential ineligible users such as internal testing accounts. For the rest of the chapter, a quantity of 0, such as 0 FP alert, is incremented by 1 for visualization on the logarithm scale. Figure 2-2 shows the empirical cumulative distribution function (ECDF) of the same quantities, indicating more than 60% of the users had a 0% FP rate. The high percentage of users with 0% FP rate is unrealistic, and it emphasizes the importance of scrutinizing and selecting reliable alerts for any subsequent analysis.

Combining two quantities may provide additional insights. Figure 2-3 shows the density contours of (a) alert count versus FP count and (b) usage length versus FP count, respectively. As expected, a smaller set of compliant users (along the diagonal) recorded more FP alerts as time progressed, and this trend holds as more total alerts are reported. Dark eclipse clusters on the bottom represent users who never reported FP alerts.

Figure 2-4 shows the 3-dimensional scatter plot where each dot represents a user. The color of the dot represents the FP rate. The dark layer on the bottom consists of users who never have reported FP alerts. Users with a 100% FP rate also clustered around low total usage length and low total alert count, potentially reflecting short-term internal testing, prototyping, or a commercial demo.

Thus, we introduce three quantities to screen for compliant users, i.e., users who regularly reported FP alerts over a long period. It will be irrational to threshold the exact value of FP since users may trigger FP at different frequencies depending on ambulatory conditions.

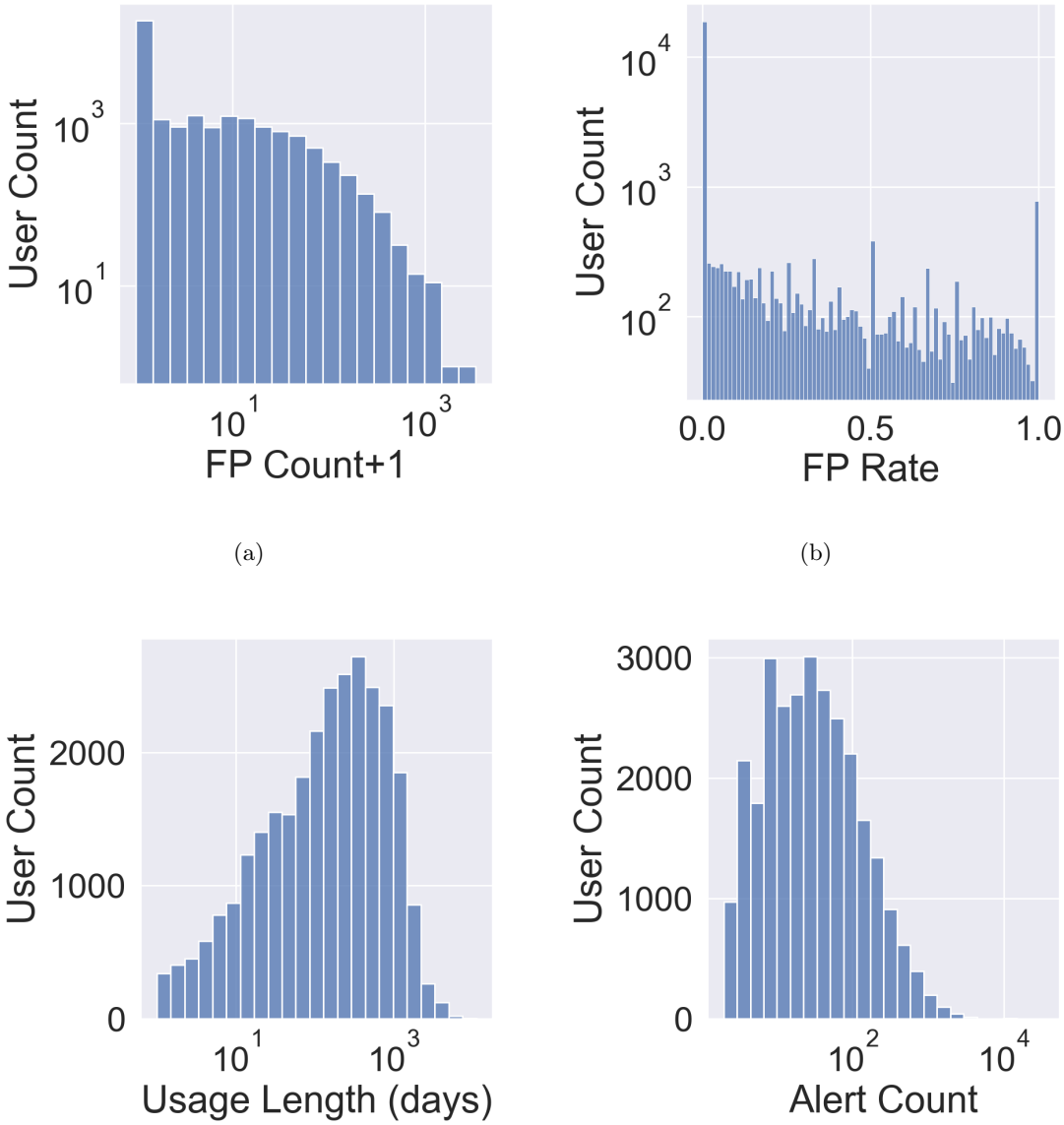


Figure 2-1: The distribution of (a) FP count, (b) FP rate, (c) usage length, and (d) total alert count in the raw data. A quantity of 0, such as 0 FP alert, is incremented by 1 for visualization on the logarithm scale.

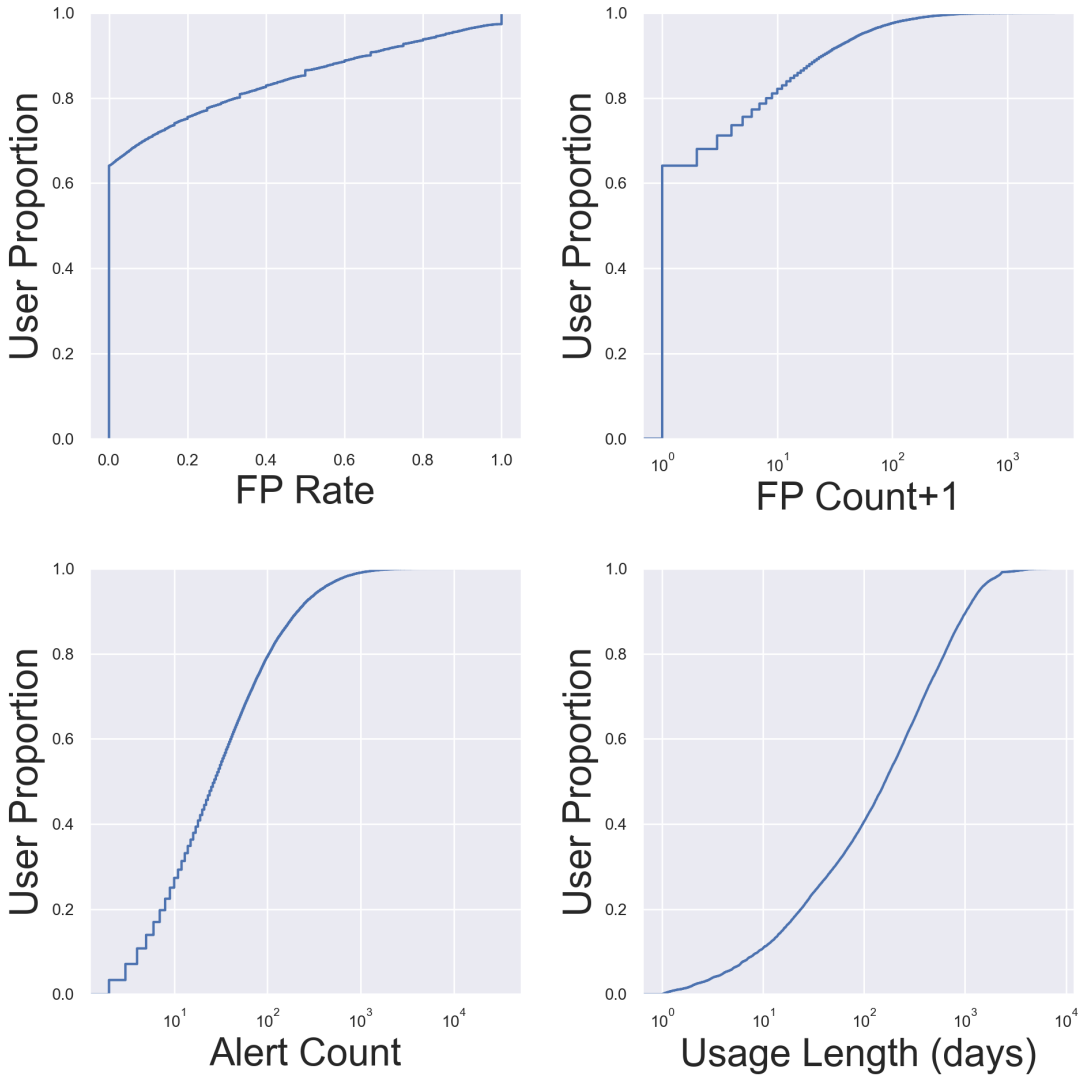
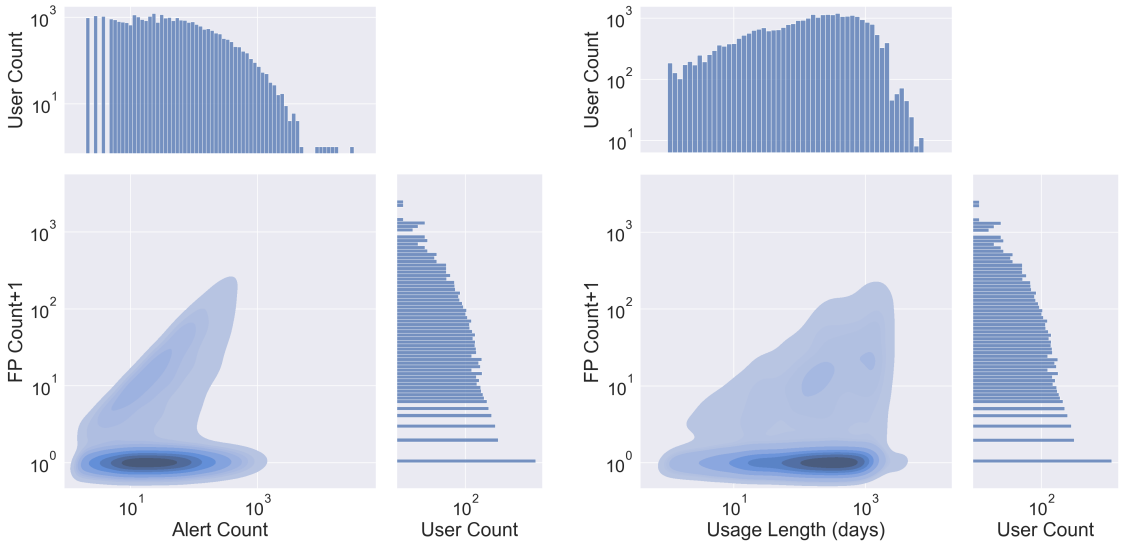


Figure 2-2: The empirical cumulative distribution function (ECDF) of FP rate, FP count, total alert count, and total usage length in the raw data. FP count is incremented by 1 for visualization on the logarithm scale.



(a) (b)

Figure 2-3: Density contours of (a) alert count versus FP count and (b) usage length versus FP count.

- `Bounded`: The bound of each user is determined by the first and the last FP alert. It naturally does not exist for users with a 0% FP rate. Quantities suffixed `Bounded` are calculated within such bounds. For example, `Usage Length Bounded` represents the time elapsed between the first and the last FP alert, which is less than or equal to total usage length; `TP Count Bounded` represents the number of TP alerts between the first and the last FP alert, which is less than or equal to total TP count total.

- `FP Interval SD`: For each user, the standard deviation of intervals between FP alerts in days.

- `FP Interval SD Normalized`: For each user, the intervals between FP alerts are normalized to [0, 1] by subtracting the minimum and then dividing by the maximum.

Figure 2-5 shows the density contours of (a) `Alert Count Bounded` versus FP count and (b) `Usage Length Bounded` versus FP count, respectively. They have successfully removed the bottom eclipse clusters, and the general trend shows a positive correlation. Figure 2-6 shows the distributions of (a) `FP Interval SD` and (b) `FP Interval SD Normalized`. A clear outlier cluster at `FP Interval SD Normalized` > 0.6 is detected.

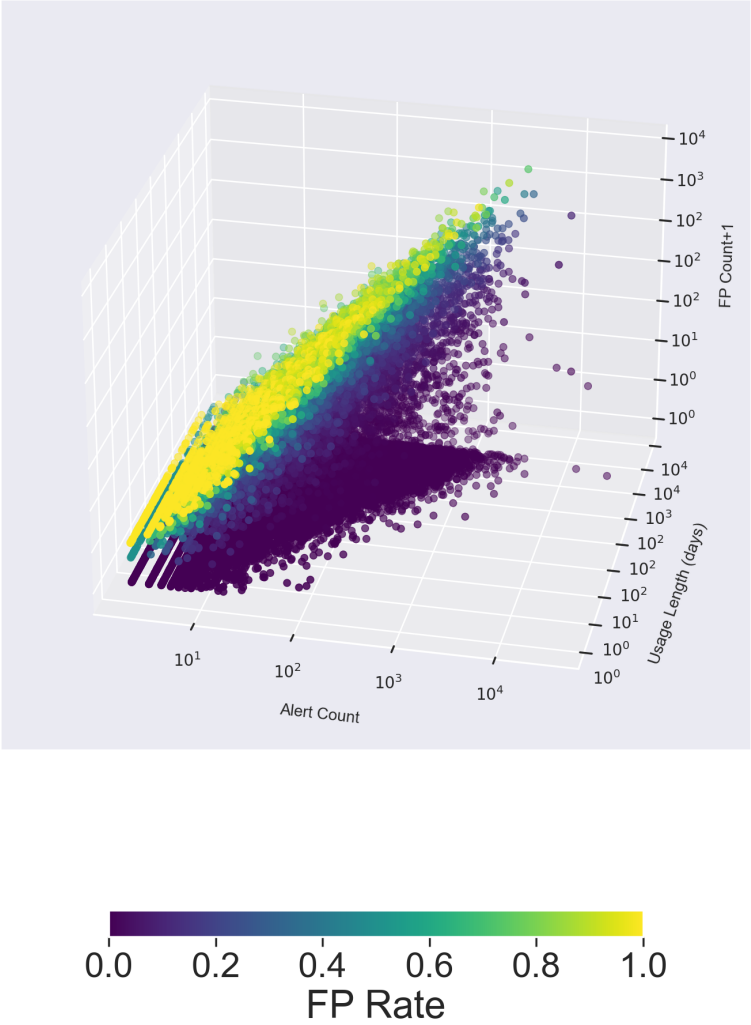
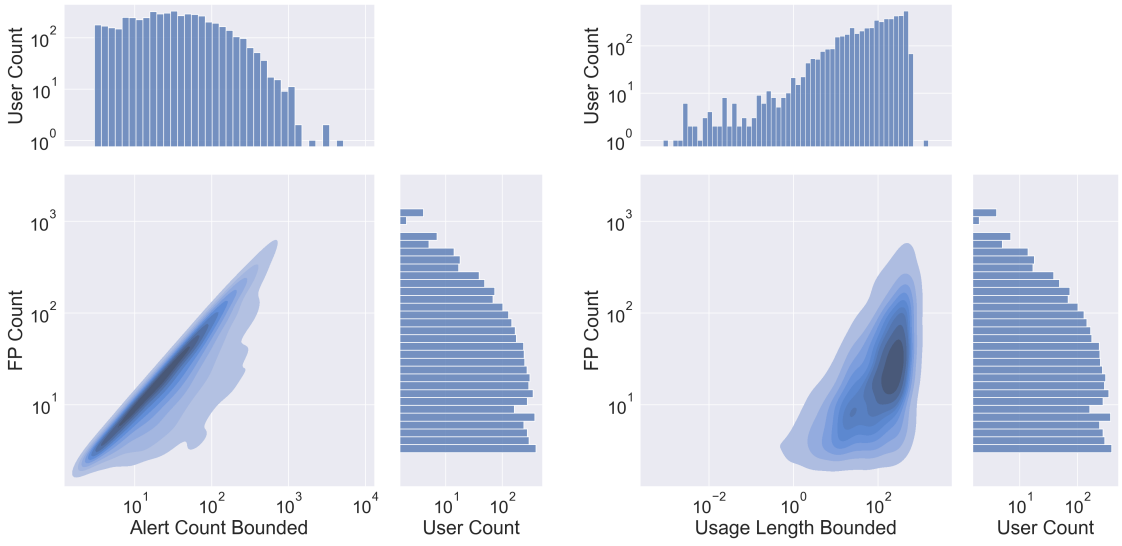
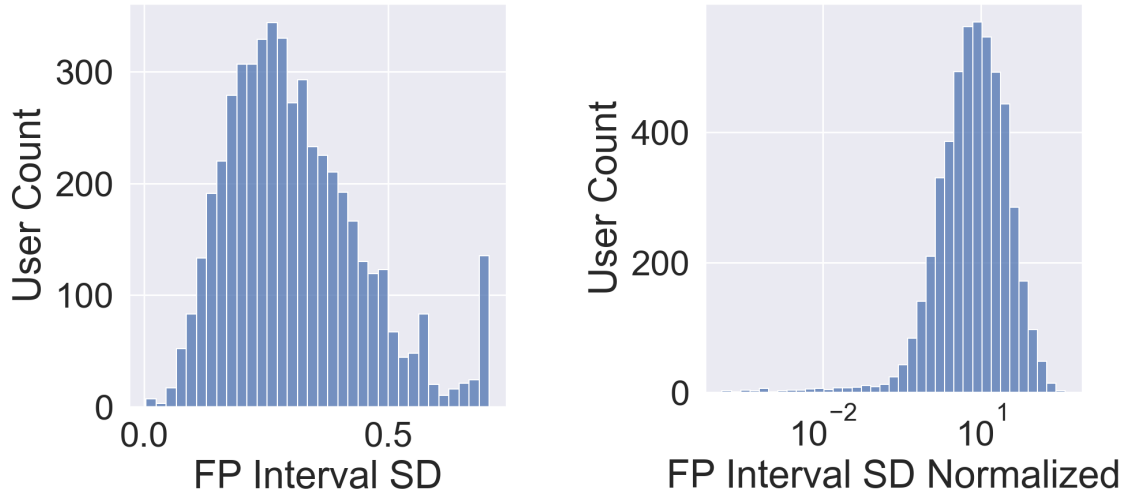


Figure 2-4: 3-dimensional scatter plot of users.



(a) (b)

Figure 2-5: Density contours of (a) `Alert Count Bounded` versus FP count and (b) `Usage Length Bounded` versus FP count.



(a) (b)

Figure 2-6: Distributions of (a) `FP Interval SD` and (b) `FP Interval SD Normalized`.

Criteria	Threshold
`FP Interval SD Normalized`	<0.6
`Usage Length Bounded`	>3 months
`FDA-cleared Algorithm`	Since February 1st, 2018
`TP Count Bounded`	>20

Table 2.2: User compliance criteria

Hence, `FP Interval SD Normalized` reflects the degree to which each user responds with FP alerts regularly regardless of the absolute FP count or the length of intervals in between. The upper bound of `FP Interval SD Normalized` is $\sqrt{0.5}$. It happens when two FP alerts are infinitely close to each other in the beginning, and one FP alert lies at the end. We observed outlier users with `FP Interval SD Normalized` greater than 0.6 in Figure 2-6, reflecting poor usage compliance. `Usage Length Bounded` serves as a strict screening criterion that reflects long device usage and ensures the user was *actively* responding with FP alerts during that time. Low `FP Interval SD Normalized` and high `Usage Length Bounded` together select users who regularly responded with FP alerts over a long time, regardless of the absolute count of FP, which may be subject to individual conditions and, therefore hard to set a threshold.

Moreover, GTCS alerts from non-patient cases, including internal testing, prototyping and engineering, and commercial demo accounts, are discarded. Alerts are deliberately triggered in these cases with external forces or programs mimicking GTCS occurrences. Some alerts were produced by older versions of the detection algorithm before FDA clearance in 2018.

Finally, the screening criteria for compliant usage are shown in Table 2.2.For the rest of the thesis, only TP GTCS alerts are studied, and "TP" is omitted for conciseness. These TP alerts (i) must be detected by the FDA-cleared algorithm after January 2018 and (ii) must be produced by compliant users who regularly reported FP over `Bounded` periods. We simultaneously check the timestamps of alerts and the firmware version of the device in case the user did not update the algorithm after the clearance. If an alert after February 1st, 2018, did not use the updated firmware, it is also discarded.

2.2.3 Rest Detection and Sleep Reports

Sleep data of patients are produced by a published rest detection algorithm on the wrist-band [46]. It is an actigraphy-based algorithm and has been validated with manually scored electroencephalography-based PSG (PSG-EEG) on a clinically diverse population. For each sleep episode detected, the quantities below are provided:

- Timestamp of bed onset in UTC: going to bed but not necessarily sleep onset.
- Timestamp of wake onset in UTC
- Sleep latency (minutes): The time between bed onset and sleep onset.
- Total time in bed (minutes)
- Number of interruptions after sleep onset
- Duration of wakefulness after sleep onset (minutes)

For the longitudinal sleep report of each patient, we first converted the timestamps to the patient local timezone (all data are stored in UTC on the cloud server). Next, naps are identified as non-overnight sleep episodes followed by a longer overnight sleep episode within 12 hours and excluded from further analyses.

Subsequently, we computed the following metrics:

- Timestamp of sleep onset = Timestamp of bed onset - Sleep latency
- Total sleep time (TST) = Total time in bed - Sleep latency - Duration of wakefulness after sleep onset
- Sleep efficiency (SE) = Total sleep time / Total time in bed
- Sleep fragmentation index (SFI): For actigraphy-based sleep detection [2], SFI represents the restlessness during the sleep episode expressed as a percentage. The SFI is calculated as the sum of two percentages: the proportion of 1-minute windows after the sleep onset that are mobile (activity count ≥ 2) and the proportion of all the sleep interruptions after the sleep onset that are ≤ 1 minute in duration. A higher SFI

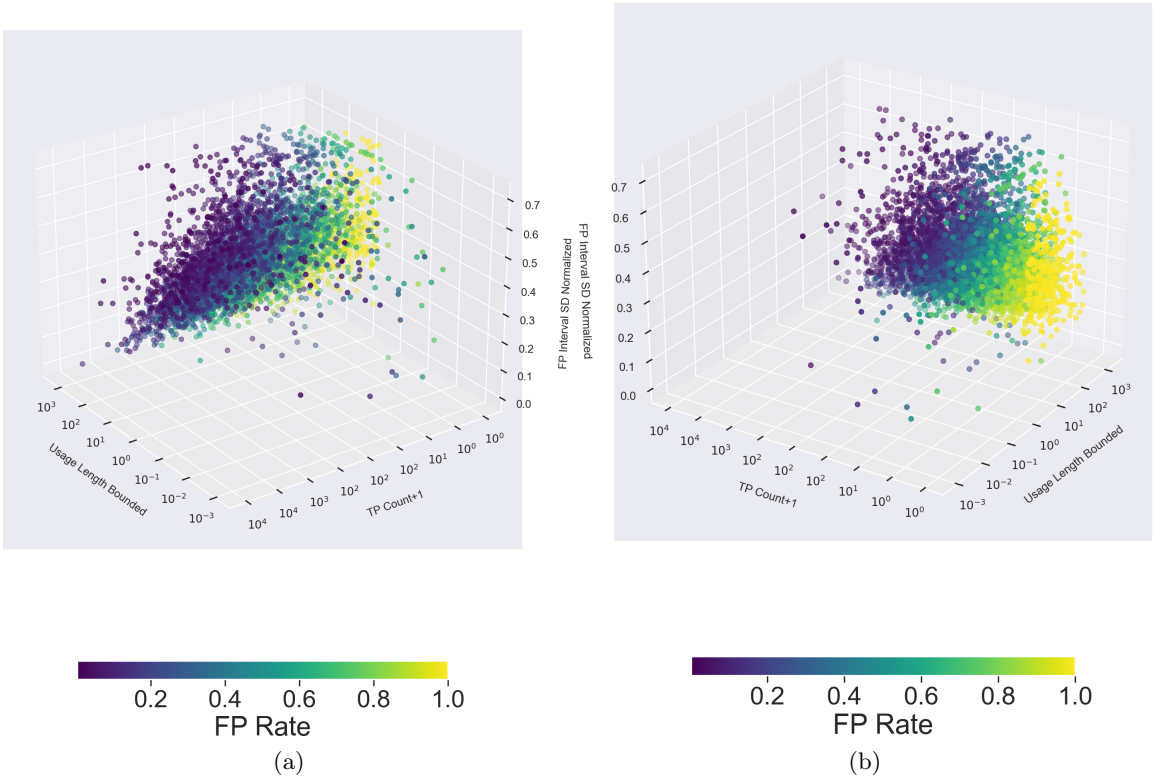


Figure 2-7: 3-dimensional scatter plots from two different angles of users after applying the compliance screening criteria.

represents a more disturbed sleep episode, and *vice versa*.

2.3 Dataset Statistics

In total, 2,932 users passed the above compliance screening. Descriptive statistics are expressed as the median ± standard deviation (SD) for this section. Figure 2-7 shows the 3-dimensional scatter plot of users after applying the above screening criteria. The color of the dot represents the FP rate. After excluding naps, off-wrist periods, and users with low compliance (in terms of FP alerts), 262,796 GTCS events (123 ± 218) and 348,593 sleep episodes (164 ± 121) were detected. Consistent with Stirling et al. [53], we include 2,132 patients with at least three months of data each.

Sleep Metric	Thresholds
TST (hours)	5, 6, 7
SE	0.7, 0.8, 0.9
SFI	0.1, 0.2, 0.3

Table 2.3: Shared thresholds of TST, SE, and SFI for the patient cohort.

2.4 Sleep Parameters and Next-day GTCS Risk

In this section, we investigate the relationship between next-day GTCS events and TST, SE, and SFI of the previous sleep episode. We first define shared thresholds for the cohort on TST, SE, and SFI, respectively. Table 2.3 shows the thresholds used. While population-level conclusions are not uninteresting, given the personal nature in sleep habits, we acknowledge that future work should emphasize more on individual sleep parameters modulating next-day GTCS risk. For each metric and each threshold, we group the sleep episodes of a patient as inferior or normal. For instance, for the 5-hour cutoff of TST, periods ≤ 5 hours are inferior, and periods > 5 hours are normal. TST of 5, 6, and 7 hours are all below the age-dependent sleep duration recommendations of the National Sleep Foundation (U.S.A.) and 24-hour movement guidelines (Canada) [10]. For the 0.2 cutoff of SFI, periods with SFI ≤ 0.2 are normal because lower SFI represents fewer sleep fragments. For SE, higher than the threshold means normal sleep. Most importantly, the patient's data were analyzed for next-day GTCS only if they had at least 30 inferior sleep episodes and at least 30 normal sleep episodes for a given threshold.

Next, we compute two risk ratios per patient:

- $R_1 = \frac{\text{Number of normal sleep episodes followed by} \geq 1 \text{ GTCS within the next 24 hours}}{\text{Number of normal sleep episodes}}$
- $R_2 = \frac{\text{Number of inferior sleep episodes followed by} \geq 1 \text{ GTCS within the next 24 hours}}{\text{Number of inferior sleep episodes}}$

After visual inspection, two-tailed paired t-tests were used to assess if the population-level difference between the two risk ratios, $R_2 - R_1$, is significant with Bonferroni correction [23]. The magnitude of effect [14] (MoE) is computed as the average of $R_2 - R_1$ over all patients, representing clinical significance. Unlike Stirling et al. [53], this approach avoids the one-to-one caveat and may provide a fair estimation of changes in next-day GTCS risk.

In addition to hard thresholds shared by the cohort, we also test patient-specific thresholds for each metric. Specifically, the threshold is set to the lowest 5^{th} percentile for TST and SE and the highest 5^{th} percentile for the SFI of each patient. The constraint of at least 30 inferior sleep and at least 30 normal sleep still applies, and the calculation of $R_2 - R_1$ is the same as above.

To draw population-level conclusions on sleep modulating seizure risk, nocturnal events should be processed carefully. As stated by Stirling et al. [53], their counter-intuitive result shows that "when patients slept less than the 25^{th} percentile, there was a slight (2%) reduction in the odds of a seizure in the following 48 h." They also recognized the reason as "undersleep was suggestive of nights without nocturnal seizures, and nocturnal seizures were found to strongly increase the risk of seizures in the following 48 h." [53] To avoid such phenomena, sleep episodes containing nocturnal GTCS are excluded from the denominators of R_1 and R_2. For numerators of R_1 and R_2, eligible next days may contain nocturnal GTCS, but the previous sleep must be GTCS-free. In such a way, R_1 and R_2 quantify only the unidirectional modulation of sleep on GTCS risk.

Besides, we am interested in the sleep regularity of patients leading to GTCS events. One established metric for quantifying the day-to-day sleep regularity of patients is Sleep Regularity Index (SRI) [41, 16]:

$$SRI = \frac{1 + \frac{1}{T-\tau}\int_0^{T-\tau} s(t)s(t+\tau)dt}{2} \tag{2.1}$$

where $s(t) = 1$ for awake and $s(t) = -1$ for asleep. The period τ is set to 24 hours. For any two time points separated by 24 hours, SRI estimates the probability of the patient being in the same state (asleep or awake). As pointed out by Fischer et al. [16], SRI is a "whole-signal metric" calculated over a period of time. Thus, it is different from a single parameter extracted from one night of sleep, such as TST, SE, and SFI. In this chapter, we consider the SRI over 2, 3, and 7-day windows consistent with Stirling et al. [53].

It is worth noticing that patient sleep reports usually only contain true-positive sleep, and periods without detected sleep may be caused by off-wrist or improper device usage, not necessarily wakefulness. Therefore, true-positive wakefulness is identified as the intersection of non-sleep and on-wrist periods. Off-wrist periods are recorded by the wristband firmware.

2 days of sleep-wake data is the minimum requirement for SRI computation, and we only consider days with less than 4.8 hours (20% of 24 hours) of off-wrist periods.

The analysis procedure of SRI and changes in GTCS risk is similar to that for TST, SE, and SFI, but we only consider personalized cutoff, i.e., SRI in the lowest 5^{th} percentile of each patient. 2, 3, and 7-day SRI are obtained by a moving window with 1-day overlaps, respectively. Three sets of R_1, R_2, and MoE are obtained for SRI over 2, 3, and 7 days.

2.4.1 Sleep Duration

We evaluated 5, 6, and 7 hours as TST thresholds. Significantly elevated GTCS risk was found for TST < 6 ($p < 0.001$, two-tailed paired t-test with Bonferroni correction) and TST < 7 hours ($p < 0.001$, two-tailed paired t-test with Bonferroni correction). Figure 2-8 shows the distributions of $(R_2 - R_1) \times 100\%$ at various TST thresholds where significant results are labeled with $*$. All distributions are empirically Gaussian, justifying the use of stricter paired t-tests over Wilcoxon rank-sum tests. A small number of subjects on the right tail are extremely sensitive to short TST. The red vertical line marks the population average $R_2 - R_1$ (MoE). Figure 2-8d shows the distribution of TST over the cohort. Periods less than 4 hours disappeared, perhaps because of nap exclusion.

Table 2.4 shows MoE for shared and personal (the lowest 5^{th} percentile) TST thresholds along with p-values, t-statistic, average R_1, average R_2, and degree of freedom (DoF). The personalized threshold was also significant ($p = 0.007$, two-tailed paired t-test with Bonferroni correction). All average MoE are small, reflecting marginal clinical significance. These results suggest that reduced sleep of fewer than 6 hours, 7 hours, and patient-specific 5^{th} percentile significantly increase next-day GTCS risk by a marginal amount.

2.4.2 Sleep Fragmentation

We evaluated 0.1, 0.2, and 0.3 as SFI thresholds. Significantly elevated GTCS risk was found for SFI > 0.2 ($p < 0.001$, two-tailed paired t-test with Bonferroni correction) and SFI > 0.3 ($p < 0.001$, two-tailed paired t-test with Bonferroni correction). Figure 2-9 shows the distributions of $(R_2 - R_1) \times 100\%$ at various SFI thresholds where significant results are labeled with $*$. All distributions are also empirically Gaussian, justifying the use of paired t-tests. The red vertical line marks the population average $R_2 - R_1$ (MoE). Figure 2-9d

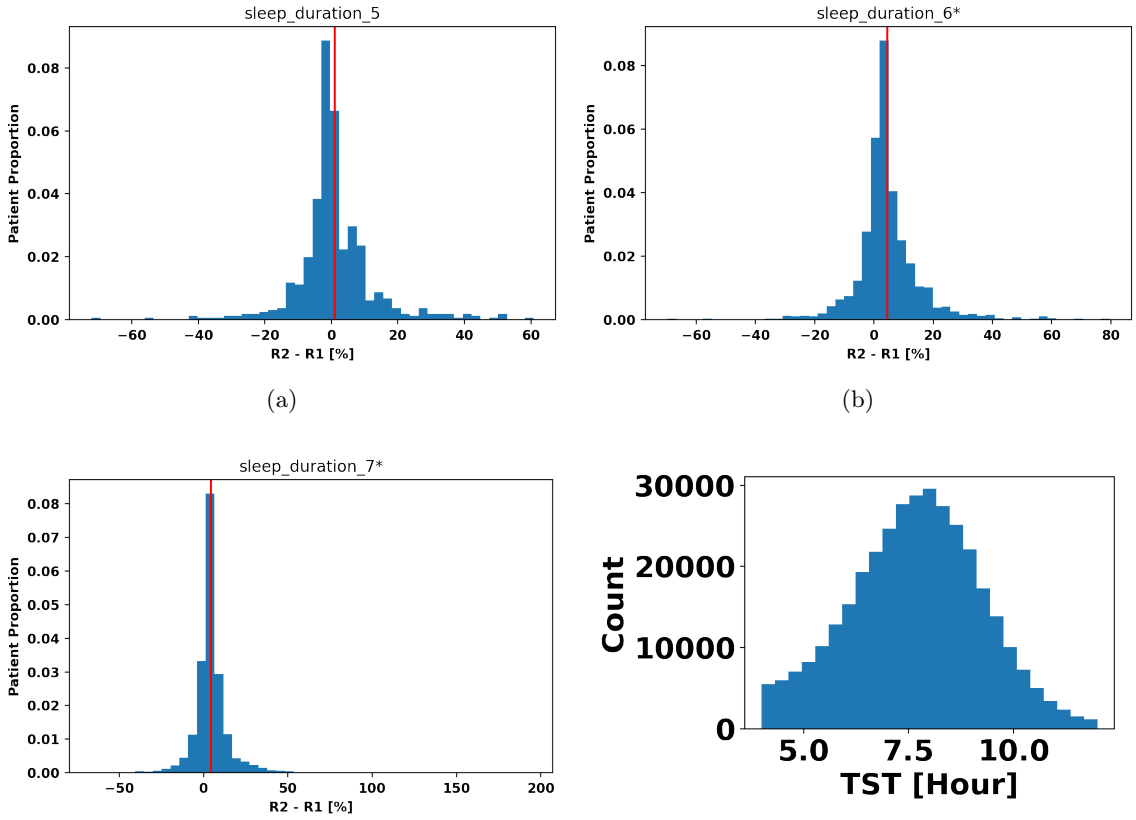


Figure 2-8: Distribution of $(R_2 - R_1) \times 100\%$ at various TST thresholds where significant results are labeled with $*$. The red vertical line marks the population average $R_2 - R_1$ (MoE). (d) Distribution of TST of all patients. R_1 : next-day GTCS risk associated with normal TST. R_2 : next-day GTCS risk associated with short TST.

Threshold	p-value	$R_2 - R_1$ (%)	t-statistic	Average R_1	Average R_2	DoF
5 hours	0.036	1.044	-2.093	0.146	0.157	606
6 hours	<0.001	4.493	-14.805	0.140	0.185	1314
7 hours	<0.001	4.475	-16.702	0.133	0.177	1780
Personalized	0.007	3.267	-11.191	0.144	0.176	1677

Table 2.4: Total Sleep Time (TST): Two-tailed paired t-tests with thresholds of 5 hours, 6 hours, 7 hours, and the lowest 5^{th} percentile personalized TST. Bonferroni correction is done to set the significance level to 0.05/4 [23]. Significant results are found for 6 hours, 7 hours, and personalized cutoff. R_1 : next-day GTCS risk associated with normal TST. R_2 : next-day GTCS risk associated with short TST.

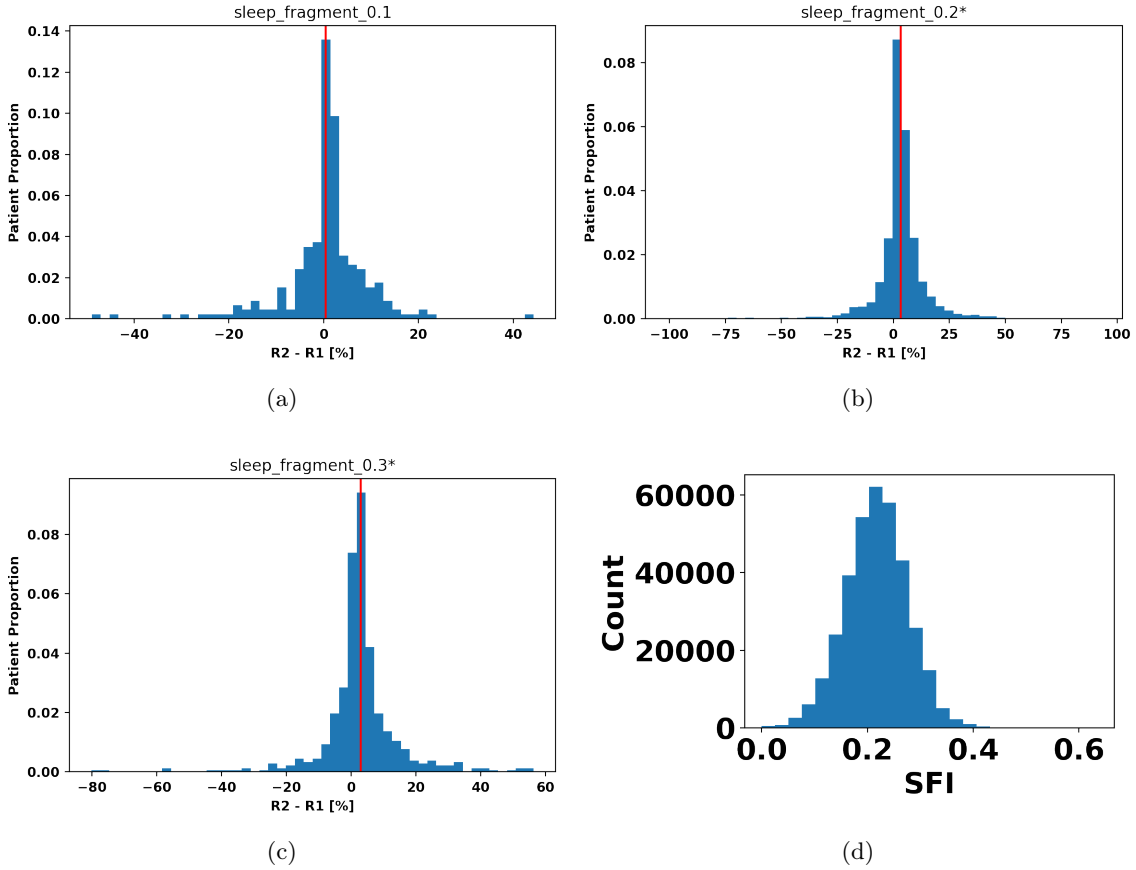


Figure 2-9: Distribution of $(R_2 - R_1) \times 100\%$ at various SFI thresholds where significant results are labeled with $*$. The red vertical line marks the population average $R_2 - R_1$ (MoE). (d) Distribution of SFI of all patients. R_1 : next-day GTCS risk associated with normal SFI. R_2 : next-day GTCS risk associated with high SFI.

shows the distribution of SFI over the cohort.

Table 2.5 shows MoE for shared and personal (the highest 5^{th} percentile) SFI thresholds along with other statistics. The highest 5^{th} percentile personalized threshold was not significant ($p = 0.06$, two-tailed paired t-test with Bonferroni correction). For significant results, all average MoE are small, reflecting marginal clinical significance. Figure 2-9 and Table 2.5 together suggest that a sleep fragmentation index of greater than 0.2 or 0.3 significantly increases next-day GTCS risk by a marginal amount.

Threshold	p-value	$R_2 - R_1$ (%)	t-statistic	Average R_1	Average R_2	DoF
0.1	0.528	0.376	-0.632	0.121	0.124	244
0.2	<0.001	3.276	-11.854	0.142	0.175	1692
0.3	<0.001	2.922	-6.300	0.149	0.178	1669
Personalized	0.06	1.003	-2.517	0.127	0.137	669

Table 2.5: Sleep Fragmentation Index (SFI): Two-tailed paired t-tests with thresholds of 0.1, 0.2, 0.3, and the highest 5^{th} percentile SFI. Bonferroni correction is done to set the significance level to 0.05/4. Significant results are found for SFI greater than 0.2 and 0.3. R_1 : next-day GTCS risk associated with normal SFI. R_2 : next-day GTCS risk associated with high SFI.

Threshold	p-value	$R_2 - R_1$ (%)	t-statistic	Average R_1	Average R_2	DoF
0.7	0.373	-0.822	0.895	0.158	0.150	110
0.8	0.852	0.074	-0.186	0.154	0.154	870
0.9	0.121	0.439	-1.553	0.143	0.147	1768
Personalized	<0.001	3.861	-13.456	0.132	0.171	1797

Table 2.6: Sleep Efficiency (SE): Two-tailed paired t-tests with thresholds of 0.7, 0.8, 0.9, and the lowest 5^{th} percentile personalized SE. Bonferroni correction is done to set the significance level to 0.05/4. Significant result was only found for personalized SE thresholds. R_1 : next-day GTCS risk associated with normal SE. R_2 : next-day GTCS risk associated with low SE.

2.4.3 Sleep Efficiency

No significantly elevated risk was found for SE with thresholds of 0.7, 0.8, and 0.9. The distributions of $(R_2 - R_1) \times 100\%$ at various SE thresholds are also empirically Gaussian, see Figure 2-10. Figure 2-10d shows the distribution of SE over the cohort.

Table 2.6 shows MoE for shared and personal (the lowest 5^{th} percentile) SE thresholds along with other statistics. The personalized threshold was significant ($p < 0.001$, two-tailed paired t-test with Bonferroni correction). Such results suggest that sleep efficiency is highly patient-specific, and low efficiency on a personal level significantly increases next-day GTCS risk by a small amount. This may be attributed to large sleep latency after bed onset, such as insomnia.

2.4.4 Sleep Regularity

Table 2.7 shows MoE for 2-day, 3-day, and 7-day personal SRI thresholds, along with other statistics. The distribution of $(R_2 - R_1) \times 100\%$ also follow an empirically Gaussian distri-

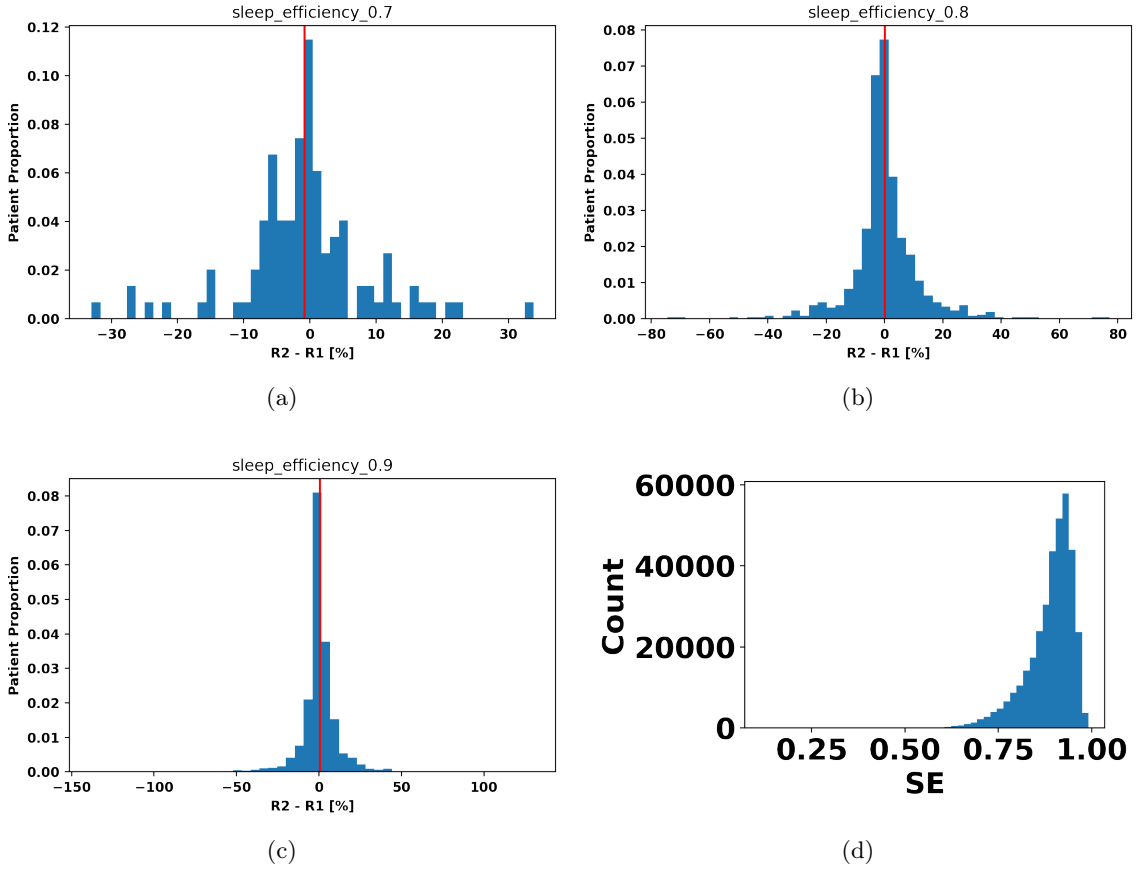


Figure 2-10: The distribution of $(R_2 - R_1) \times 100\%$ at various SE thresholds. The red vertical line marks the population average $R_2 - R_1$ (MoE). (d) Distribution of SE of all patients. R_1 : next-day GTCS risk associated with normal SE. R_2 : next-day GTCS risk associated with low SE.

SRI	p-value	$R_2 - R_1$ (%)	statistic	Average R_1	Average R_2
2-day	0.006	3.130	10337.5	0.144	0.175
3-day	0.139	0.175	2414.0	0.148	0.150
7-day	0.231	-0.038	159788.5	0.153	0.152

Table 2.7: Sleep Regularity Index (SRI): Two-tailed paired t-tests with the lowest 5^{th} percentile SRI threshold over 2, 3, and 7 days. Bonferroni correction is done to set the significance level to 0.05/3. Significant result is only found for 2-day SRI. R_1 : next-day GTCS risk associated with normal SRI. R_2 : next-day GTCS risk associated with low SRI.

bution, so two-tailed paired t-tests were used. The personalized threshold was significant ($p = 0.006$, two-tailed paired t-test with Bonferroni correction) for 2-day SRI. Since 2-day SRI essentially measures only the deviation of the current sleep-wake cycle from the previous day, this result may be confounded by previous sleep's TST and SFI studied above. Replicating and reconciling with the sleep and wake onset regularity metrics proposed by Stirling et al. [53] in this GTCS cohort remains future work.

Chapter 3

Cycles in Electrodermal Activity and Generalized Tonic-Clonic Seizures

In this chapter, we focus on retrospective analyses of cycles in wrist EDA from patients diagnosed with GTCS, validating previous claims on a larger patient cohort. Same as the previous chapter, GTCS attacks are detected by the FDA-cleared multi-modal algorithm embedded in the wearable device. Wavelet methods are used to extract and examine robust circadian and multi-day cycles in EDA. As EDA may be influenced by peripheral temperature measured on wrists (TEMP), e.g., increasing skin temperature coexisting with sweating, we also consider EDA weighted by TEMP. Phase-locking properties of GTCS events detected by wearable devices are assessed by Hilbert spectral analyses. In the end, we provide a simple benchmark for evaluating phase-locking in addition to previously used statistical tests [30, 20], showcasing how *simulated* GTCS reports may frequently phase-lock to cycles of certain periods in weighted and unweighted EDA signals.

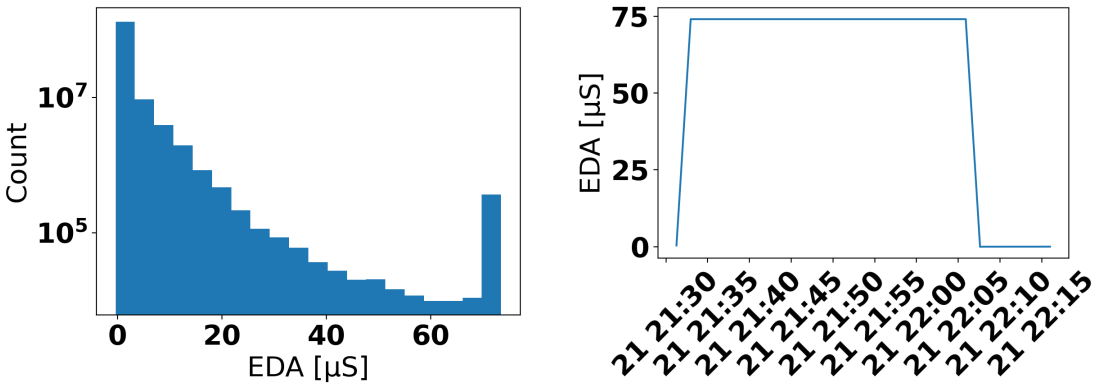
3.1 Previous Literature

Karoly et al. provided a review of historical evidence of multi-day seizure cycles [29]. They have also demonstrated that these cycles do not always sync with the female menstrual or lunar cycles [29]. Cohorts from different studies have reported about-monthly seizure cycles across different epilepsy types, genders, and ages [5, 21]. Self-reported online seizure diaries

converged on such persistent multi-day rhythms across large patient populations, where the periodicity runs from one week to 30 days [28].

With advances in intracranial chronic EEG monitoring, Baud et al. discovered multi-day rhythms in IEA [3]. Seizures cluster around the rising period of IEA. The periodicity of multi-day IEA rhythms ranges between 7 to 30 days for both male and female patients. Subsequently, Rao et al. validated the correlation between IEA phases and seizure occurrences [45]. They have also concluded that multi-day cycles are not entrained by any known external cue, conjecturing an endogenous mechanism. Subsequently, Karoly et al. illustrated that multi-day heart rate (HR) cycles were presented in PWE and healthy control participants, and 10 out of 19 PWE had seizures significantly phase-locked to personal multi-day heart rate cycles [30]. While current studies have only found correlational conclusions, various results have suggested possible causal relationships between hormonal factors and multi-day seizure cycles [29].

EDA has been extensively studied in the epilepsy domain. For instance, wrist EDA has helped improve the specificity of GTCS event onset detection [38]. Rising EDA of large amplitudes are correlated with the duration of post-ictal generalized EEG suppression (PGES) [43]. PGES has been observed in 100% of SUDEP cases monitored by EEG [49]. Consistent with Ryvlin et al. [49], unusually high EDA has been observed in one SUDEP case [42]. EDA may thus be associated with neurological changes during seizures at heightened risk of SUDEP, though other pathways remain possible. Recently, researchers have started looking at long-term rhythmic patterns in EDA of PWE. Gregg et al. discovered that circadian and multi-day cycles exist in physiological signals measured by wristbands, including heart rate, actigraphy (ACC), EDA, and TEMP, and seizures recorded by RNS devices showed significant correlation with particular phases of these cycles [20]. Objectively and subjectively recorded seizures are found to occur at preferred phases of different cycles with similar statistical methods, and the periods of detected physiological cycles are often patient-specific [29]. Converging evidence in separate published studies implies the prevalence of circadian and multi-day cycles in physiological signals of PWE. These cycles are discovered by retrospective observational statistics, and investigations of the pathological source of seizure risk modulation are required. Nonetheless, it has been shown that phase information in physiological signals can facilitate better seizure risk forecasting tech-



(a) Empirical distribution of EDA values with outliers

(b) An example of flat-top EDA

Figure 3-1

nology [8].

3.2 Data Preprocessing

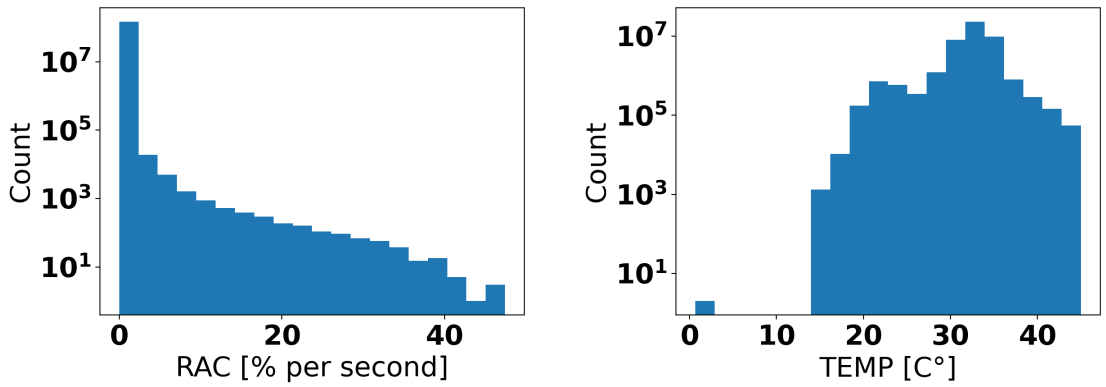
3.2.1 GTCS Alerts

Consistent with the previous chapter, only lead seizures were kept in the GTCS report of each patient. Patients with fewer than 20 GTCS events were excluded from further analyses. Filtering of false-positive alerts and compliant patient screening procedures are the same as Section 2.2.

3.2.2 Physiological Signal from Wearable Device

Chronic wrist EDA is susceptible to movement artifacts, device malfunction, and user non-compliance. To ensure high-quality EDA signals, we followed previously proposed EDA preprocessing standards [6, 35]. In particular, EDA samples less than $0.05\mu S$ were discarded. To set an upper threshold for invalid EDA, we visually inspected the distribution of EDA values, see Figure 3-1a. An abnormal peak $\geq 70\mu S$ was observed, and further visualization revealed a "flat-top" pattern for these high values, possibly due to sensor saturation. Figure 3-1b gives an example of flat-top EDA at this abnormal value. Consequently, all EDA samples greater than $70\mu S$ were removed.

The rate of amplitude change (RAC) is another essential indicator of EDA quality [6, 35].



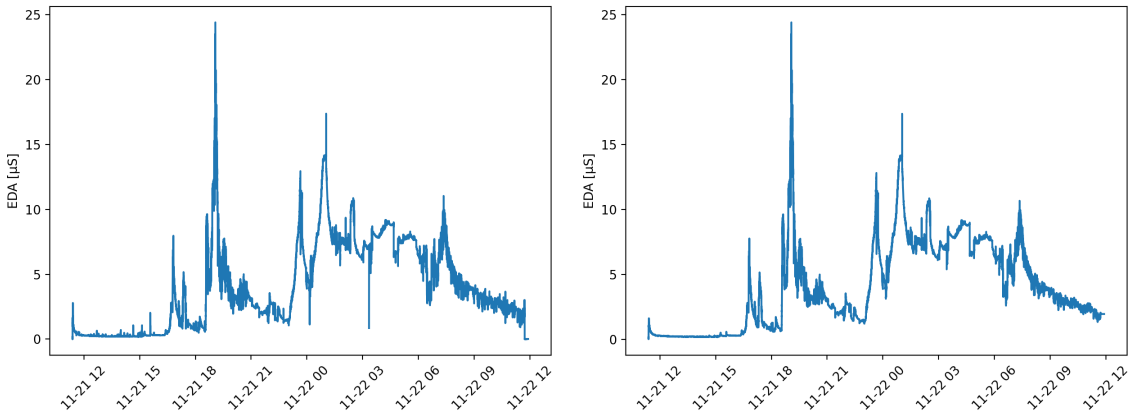
(a) Distribution of RAC in raw EDA signals. Notice the large RAC representing $> 40\%$ change in 1-second windows.

(b) Empirical distribution of TEMP with outliers

Figure 3-2

For all EDA signals, we performed a rolling window of 1 second, and RAC is calculated as the percentage change between the maximum and minimum EDA within each window. Figure 3-2a shows the distribution of RAC in raw EDA signals, noticing the large RAC $> 40\%$ within 1 second. In the case of ascending, windows with RAC ≥ 0.2 were removed. In the case of descending, windows with RAC ≥ 0.1 were removed. The previous EDA outlier removal was performed before RAC cleaning in order to avoid false-positive RAC windows caused by exceptionally high or low EDA. Concurrent TEMP was also cleaned by simple thresholds [6, 20]. The distribution of TEMP is shown in Figure 3-2b, and only values between $29°C$ to $40°C$ were preserved. Notably, TEMP is converted from the digital-to-analog (DAC) value on Embrace 2 wristbands to Celsius, and the conversion is done by a lookup table provided by the temperature sensor manufacturer. We have retrieved the correct table for different firmware versions, but outliers still exist and need explanations.

For concurrent samples of EDA and TEMP, if one modality was deemed invalid, both were removed. Also, physiological signals during GTCS events are removed, and the onset and duration are provided in the alerts. In the end, EDA and TEMP samples within the first minute after putting on the device were discarded to account for sensors warming up. Figure 3-3 showcases the effect of the proposed artifact removal for an example 24-hour period. Additional details on batch downloading from cloud servers (more than 10TB), memory-efficient parsing of long-term signals, and fixing signal timestamp drifts are omitted.



(a) Before EDA artifact removal

(b) After EDA artifact removal

Figure 3-3

After removing invalid EDA and TEMP values, missing gaps are interpolated. Following [3], for a given period T (see Section 3.4), only gaps shorter than 20% of T were interpolated. For example, only gaps less than 4.8 hours were filled for wavelet analysis of 24 hours. We utilized a moving window of T centered at each missing sample to interpolate. For a missing sample, it is filled by the median EDA within the window of length T, ignoring other missing values in the same window if presented. There have been several attempts and discussions of interpolation by taking the average of all samples from the same clock hour [30, 20], e.g., missing values at 8 P.M. are filled by the average of other EDA values at 8 P.M. Yet, it could modify circadian and other possible rhythms in the signal, so we refrained from this practice.

3.2.3 Weighted Electrodermal Activity

With cleaned EDA and TEMP from above, we also analyzed weighted EDA. Since skin conductance may be influenced by skin temperature, e.g., sweating when overheating, it is interesting to study EDA variations in the absence of TEMP changes. As skin conductance is controlled by the sympathetic nervous system, skin conductance responses (SCR) may appear when experiencing "fight or flight" responses without temperature variations. TEMP signals are first up-sampled from 1Hz to 4Hz by linear interpolation. For each patient, we normalized the TEMP signal to $[0, 1]$ by (i) subtracting the minimum TEMP value and then

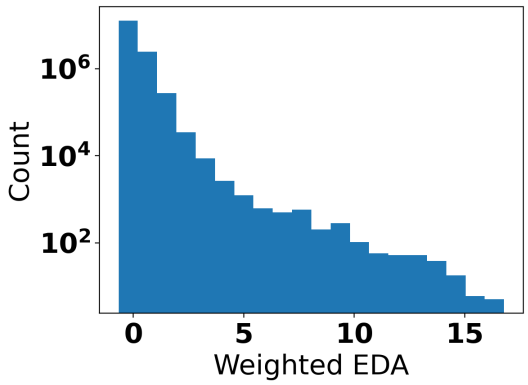


Figure 3-4: The distribution of the weighted EDA values from the same subjects in Figure 3-1 and Figure 3-2.

(ii) dividing the whole signal by the new maximum TEMP value.

Weighted EDA is computed as the original EDA multiplied by 1− normalized TEMP at the same timestamp. The distribution of the weighted EDA values from the same subject in Figure 3-1 and Figure 3-2 is shown in Figure 3-4. Please notice that weighted EDA was obtained after cleaning the original EDA and TEMP.

3.3 Dataset Statistics

After signal cleaning and interpolation, only the longest segment without missing values from each patient was included for further analyses. Physiological signal segments shorter than three months were discarded. Since the amount of valid data varies with the wavelet period T, we report here patients with at least one cycle being analyzed. In total, 69,719 days of physiological signals and 26,280 concurrent lead GTCS alerts were available from 1,797 patients. Physiological signal length ranges from 3 months to 3.7 years (median: 2.7 years). The number of lead GTCS alerts ranges from 20 to 177 (median: 121). Following the definition of data completeness to quantify patient compliance [6], the completeness score ranges from 0.43 to 0.96 (median: 0.78) in the cohort. Detailed breakdowns of the demographic information of the patient population are not available due to industry privacy regulations.

3.4 Cycle Extraction with Wavelet Analysis

3.4.1 Background

A Morlet wavelet is defined by a simple periodic wave damped by a Gaussian:

$$\psi_0(t) = e^{i\omega_0 t} e^{-t^2/2}, \tag{3.1}$$

where $\omega_0 = 6$ in our analysis. The peak frequency f_0 on its discrete Fourier transform is given by $\frac{6}{2\pi} = \frac{3}{\pi}$, and $T_0 = \frac{1}{f_0}$, i.e., the Fourier period. To scale the wavelet to a desired period T, we define $sT_0 = T$, and thus the scaled wavelet is given by

$$\psi(t) = e^{i\omega_0 \frac{t}{s}} e^{-\left(\frac{t}{s}\right)^2/2} \tag{3.2}$$

Figure 3-5 showcases how the Morlet wavelet was scaled to desired frequencies.

3.4.2 Wavelet Analysis of Physiological Signal

For each segment, the signal was down-sampled to 1 sample per 5 minutes and then z-scored [30, 20]. Before down-sampling, an anti-aliasing low-pass finite impulse response (FIR) filter was applied to the 4Hz EDA signal with a cutoff frequency of $\frac{4}{1/(5\times60)}$. The order of the FIR filter was set to $20 \times \frac{4}{1/(5\times60)}$, and we used a Hamming window. We used the `SciPy` [58] package [1]. The power of the rhythm of period T in the signal was obtained by a Morlet wavelet ($\omega_0 = 6$) transform scaled to T. Consistent with Baud et al. [3] and Karoly et al. [30], we investigate 89 periods T (scales) with ascending spacing: 1.2 hours between 2.4 and 31.2 hours, 2.4 hours between 33.6 and 48 hours, 4.8 hours between 2.2 and 4 days, 12 hours between 4.5 and 10 days, and 24 hours between 11 and 45 days. If a given T is longer than 25% of the segment length, the wavelet analysis is terminated. It is required to observe at least four cycles of any given T.

The scaled wavelet $\psi(t)$ in Equation 3.2 was then convolved with the down-sampled signal to obtain the power spectrum. Figure 3-6 illustrates the wavelet transform process of an example patient with six months of EDA where all small gaps are interpolated. For visualization purposes, hourly average z-scored EDA are plotted in Figure 3-6a, but the actual

[1] `https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.decimate.html`

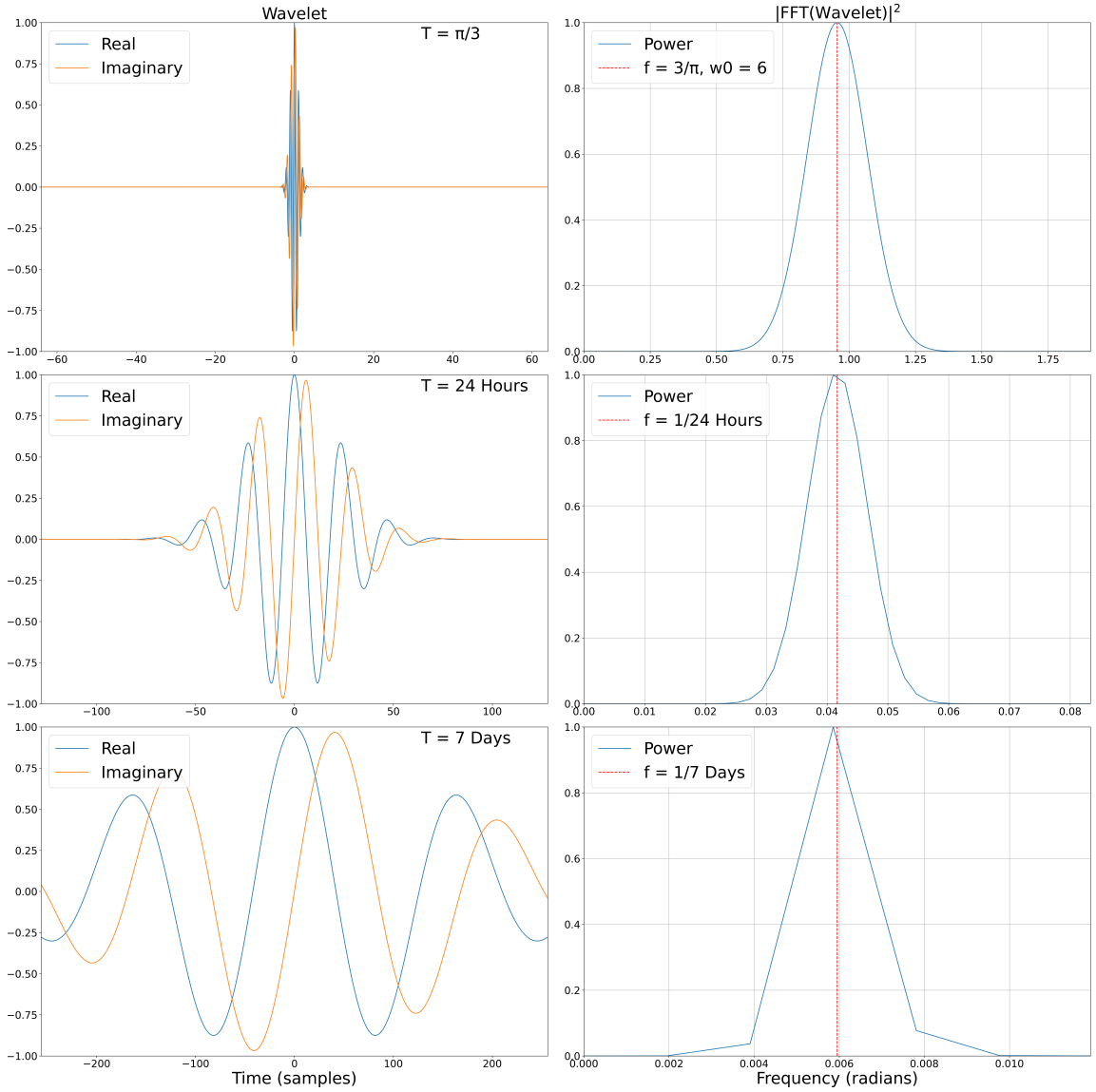


Figure 3-5: Morlet wavelets scaled to different frequencies. Top to bottom: $T = \frac{\pi}{3}$, 24 hours, and 7 days. Left column: wavelets in the time domain. Right column: The power spectrum of wavelets after Fourier transforms with peaks at $\frac{1}{T}$. Power index on the y-axis is normalized (divided by the maximum power) for visualization.

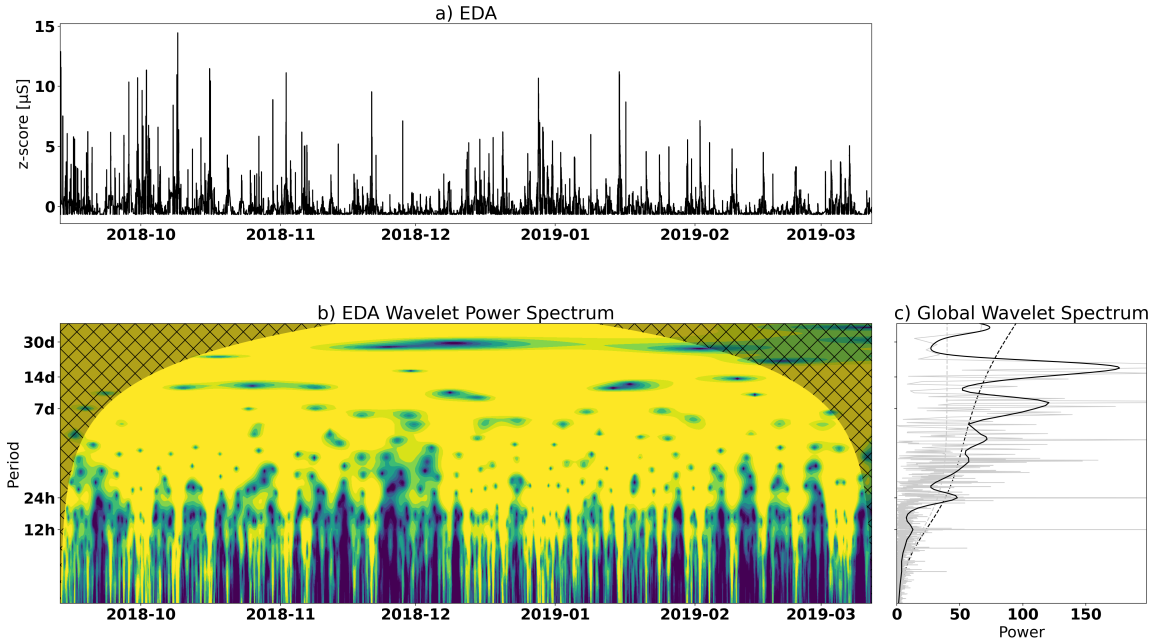


Figure 3-6: The wavelet transform of ambulatory wrist EDA from an example patient. a): Hourly average z-scored EDA values. b): Wavelet power spectrum across the 89 scales. Regions crossed out represent the CoI. c): Time-average global power spectrum. The solid gray line shows the Fast Fourier Transform (FFT) of the EDA for comparison.

wavelet was convolved with the original z-scored EDA. Figure 3-6b is the corresponding wavelet power spectrum across the 89 scales. Zeros were padded on both ends of the finite-length segment for the wavelet transform. The cone of influence (CoI), i.e., regions where the estimated power was diminished due to the edges of the signal, is given by $\sqrt{2} \times T$ [56] and crossed out in the figure. We excluded all CoI in further analyses of phase-locking. In Figure 3-6c, the solid black line shows the time-average global power spectrum, and the solid gray line shows the Fast Fourier Transform (FFT) of the EDA. This comparison reaffirms that wavelets are more suitable for analyzing frequencies in non-stationary signals than FFT.

3.4.3 Significant Rhythm Finding

In order to determine, for each patient and each signal modality, if a peak in the time-average spectrum is significant, we adopt a red noise simulation method [56]. In particular, the lag-1 autocorrelation of the original signal is used to simulate a red-noise process. On the time-average power spectrum, simulated red noise produces an upper 95% CI for each frequency, and this bound is used to threshold peaks in the spectrum, i.e., only peaks above the 95%

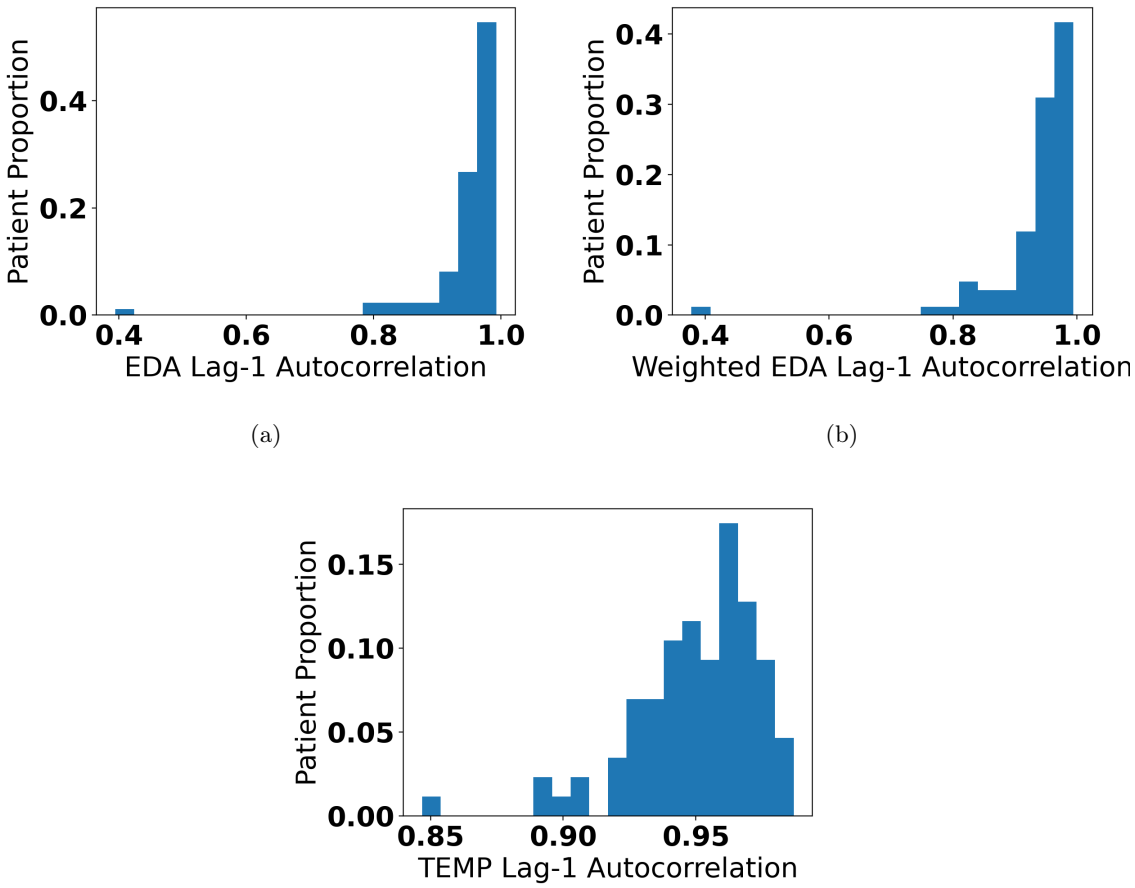


Figure 3-7: Distribution of lag-1 autocorrelation α (x-axis) for (a) EDA, (b) weighted EDA, and (c) TEMP over the cohort.

CI are deemed significant. Figure 3-7 below shows the distribution of lag-1 autocorrelation α for (a) EDA (range 0.4-0.98; median: 0.96), (b) weighted EDA (range 0.38-0.99; median: 0.97), and (c) TEMP (range 0.85-0.98; median: 0.95) over the cohort.

In Figure 3-6c, the dashed black line shows the upper 95% CI for different frequency components in the EDA of the example patient. For comparison, the dashed gray line shows the 95% CI based on a white-noise process (autocorrelation of zero), which was used by Gregg et al. [20]. The red-noise process produces a higher and more realistic CI than the white-noise process for physiological signals. Peaks are defined as local maxima by comparing every three adjacent values. Only the highest is kept for any neighboring peaks within $\frac{T}{3}$. The same wavelet transform and peak finding were also applied on weighted EDA and

	EDA	Weighted EDA
At least 1 cycle	96%	92%
More than 1 cycle	82%	81%
24-hour cycle	94%	90%
7-day cycle	70%	70%
28 to 32-day cycles	66%	64%

Table 3.1: Percentages of the patient population with significant cycles detected in wrist EDA and weighted EDA.

TEMP signals to identify frequency components with significant power, see Figure 3-8 from the same example patient. We used `Python 3.10` and the `SciPy` [58] package for wavelet analyses and peak finding.

Before looking into the relationship between GTCS occurrence and significant cycles discovered in EDA, we inspect the distribution of time-average power spectra over the patient cohort. Figure 3-9a shows the global average EDA and weighted EDA spectra over the 1,797 patients from Section 3.3. Shaded areas represent ± 1 standard deviation. Dashed lines represent the population-average upper 95% CI for EDA and weighted EDA power spectra, respectively. The upper 95% CI was calculated for each patient based on personal lag-1 autocorrelation (see Figure 3-7), but ± 1 standard deviation is not displayed in the figure for visual clarity. The power index was estimated for each T as the square root of the time-average absolute value of complex wavelet coefficients. Given certain long periods, some patients might not have enough data for the analysis (e.g., three months of data but $T = 45$ days) and were thus ignored for that T. In Figure 3-9b, circadian (24-hour), weekly (7-day), and about-monthly (28 to 32-day) EDA and weighted EDA cycles are robust and shared across the patient cohort. The median ratio of weekly to circadian peak amplitude was 1.8 (range 0.7-3.7; > 1 in 1,310 patients), demonstrating similar power. Not much difference was found between the global power spectra of EDA and weighted EDA in Figure 3-9, reflecting that the effect of TEMP on EDA rhythms may be marginal.

Table 3.1 shows percentages of the patient population with significant cycles detected in wrist EDA and weighted EDA ($>$ 95%CI, red noise simulations). Following [3], in order to assess the consistency of physiological rhythms over time, the Pearson coefficient was computed between the time-average spectrum (Figure 3-6c) and each vertical slice (timestamp) of the

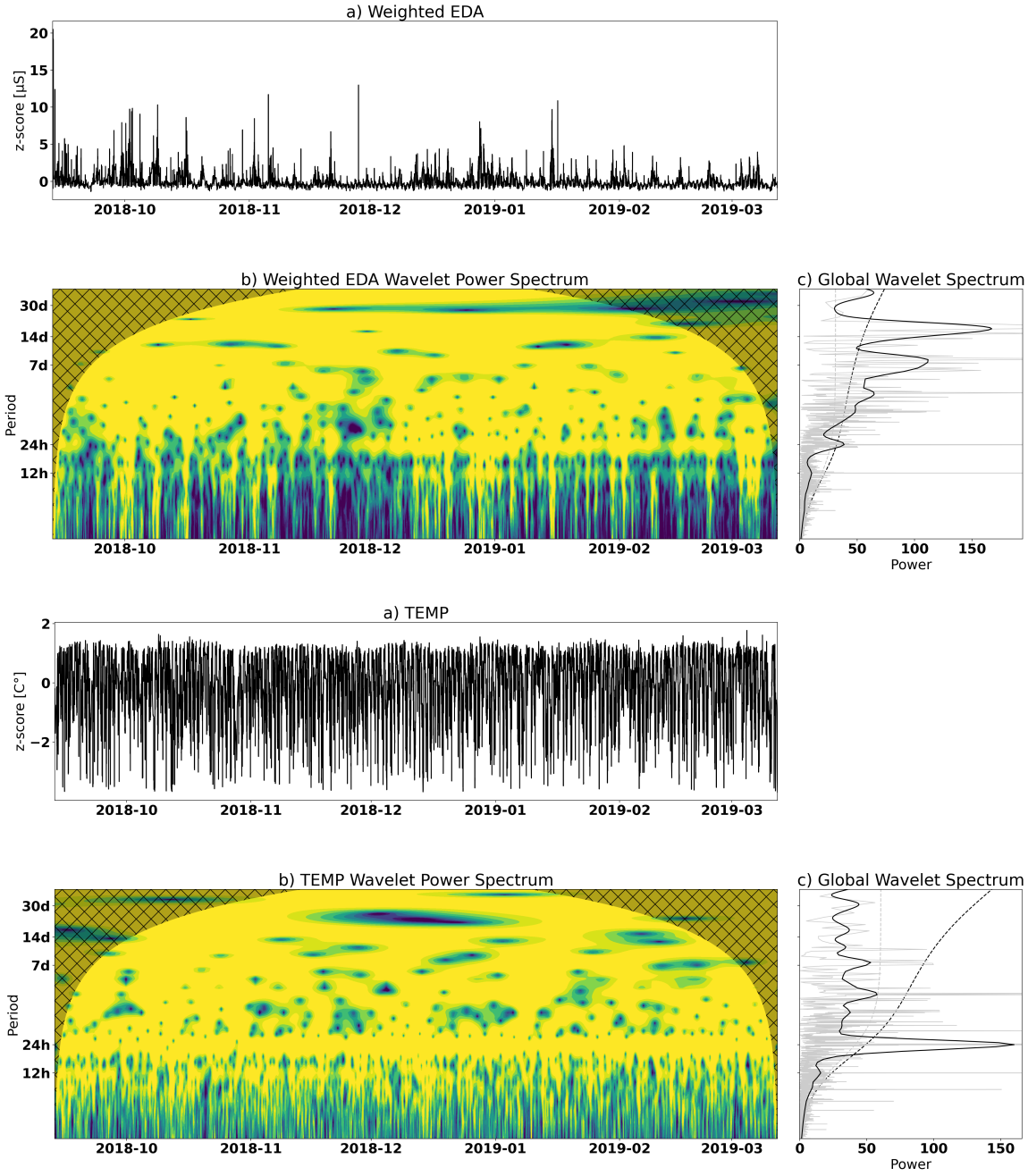


Figure 3-8: The wavelet transform and time-average spectrum of weighted EDA and TEMP from the same patient in Figure 3-6. The solid gray line shows the Fast Fourier Transform (FFT) for comparison.

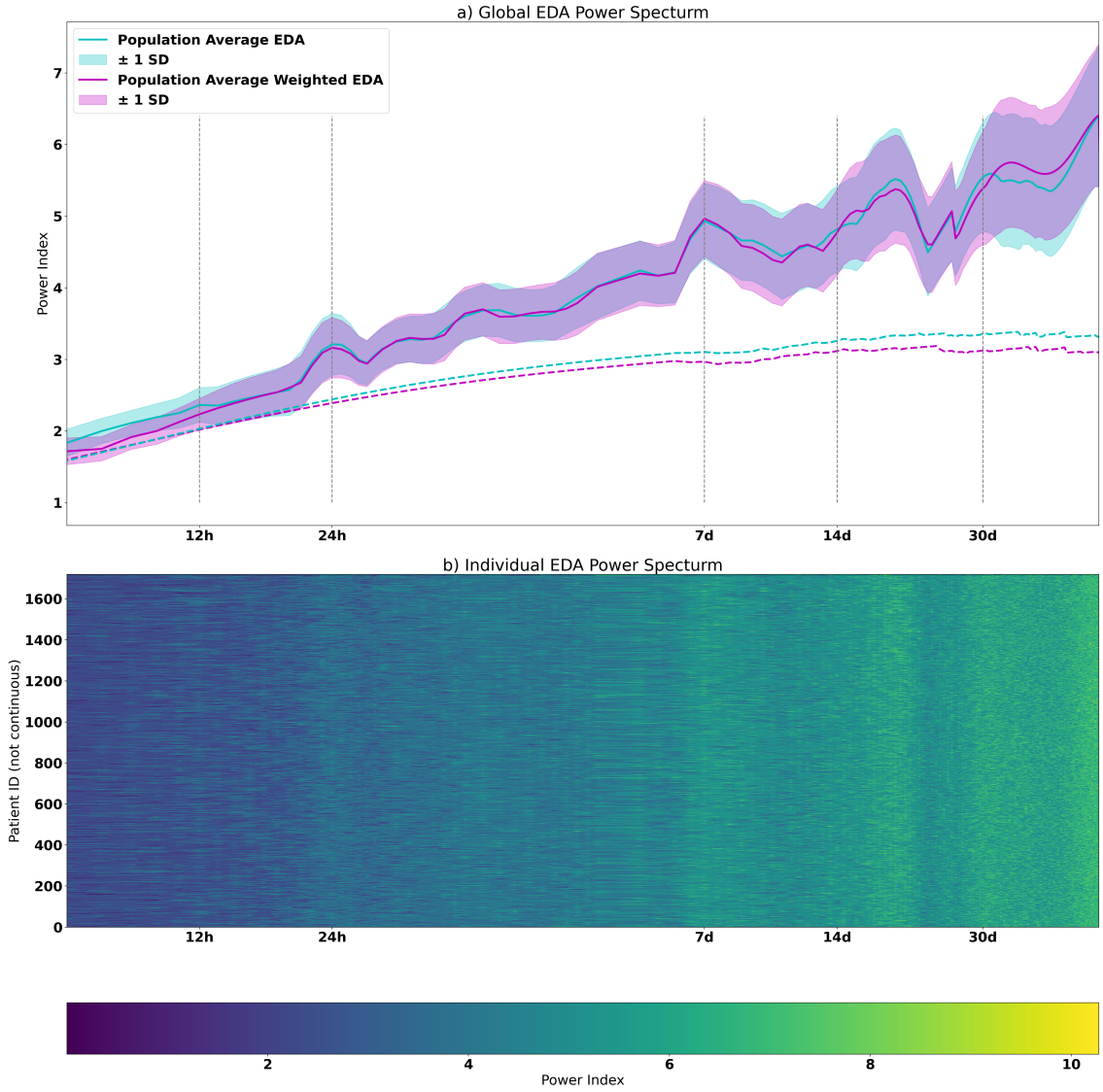


Figure 3-9: EDA and weighted EDA power spectra of the cohort. a): Global average spectrum over the 1,797 patients from Section 3.3. Shaded areas represent ±1 standard deviation. Dashed lines represent the population-average upper 95% CI for EDA and weighted EDA power spectra, respectively. b): Individual EDA spectra. Circadian, weekly, and about-monthly cycles are shared across patients.

original wavelet spectrum (Figure 3-6b). After that, the lag-1 autocorrelation is computed for this array of Pearson coefficients of each patient. The median autocorrelation is 0.73 (range 0.23-0.91; < 0.5 in 89 patients), suggesting cycles are not robust over time for a small number of patients. These 89 patients were excluded from further analyses.

In contrast, Figure 3-10a shows the global average TEMP spectrum over the 1,797 patients from Section 3.3. Only the circadian cycle in TEMP was prevalent across the cohort, see Figure 3-10b.

3.5 GTCS Alerts Phase-locking to Physiological Cycles

3.5.1 Instantaneous Phase Estimation

We now estimate the phases of GTCS events using the Hilbert Transform in the frequency domain. Given a significant frequency $\frac{1}{T}$, a second-order zero-phase Butterworth band-pass filter with a cutoff at $\frac{1}{T} \pm 33.3\%$ was applied to the physiological signal to obtain the corresponding component $x(t)$. Similar to Baud et al. [3] and Karoly et al. [30], if the bands of multiple $x(t)$ overlap, only the strongest was kept for analyses. We denote the FFT of the real-valued $x(t)$ as

$$X(f) = \mathcal{F}\{x(t)\} \tag{3.3}$$

Next, we obtain $Y(f)$ as the result of applying the frequency response of the Hilbert Transform to $X(f)$:

$$Y(f) = \begin{cases} -j \cdot X(f) & \text{for } f > 0 \\ 0 & \text{for } f = 0 \\ j \cdot X(f) & \text{for } f < 0 \end{cases} \tag{3.4}$$

This creates a phase shift of $-\frac{\pi}{2}$ for positive frequencies and $\frac{\pi}{2}$ for negative frequencies. Subsequently, the Hilbert Transform, $\hat{x}(t)$, of $x(t)$ is given by the inverse FFT:

$$\hat{x}(t) = \mathcal{F}^{-1}\{Y(f)\} \tag{3.5}$$

The analytic signal $x_a(t)$ is obtained by combining the original $x(t)$ and its Hilbert Transform $\hat{x}(t)$:

$$x_a(t) = x(t) + j\hat{x}(t) \tag{3.6}$$

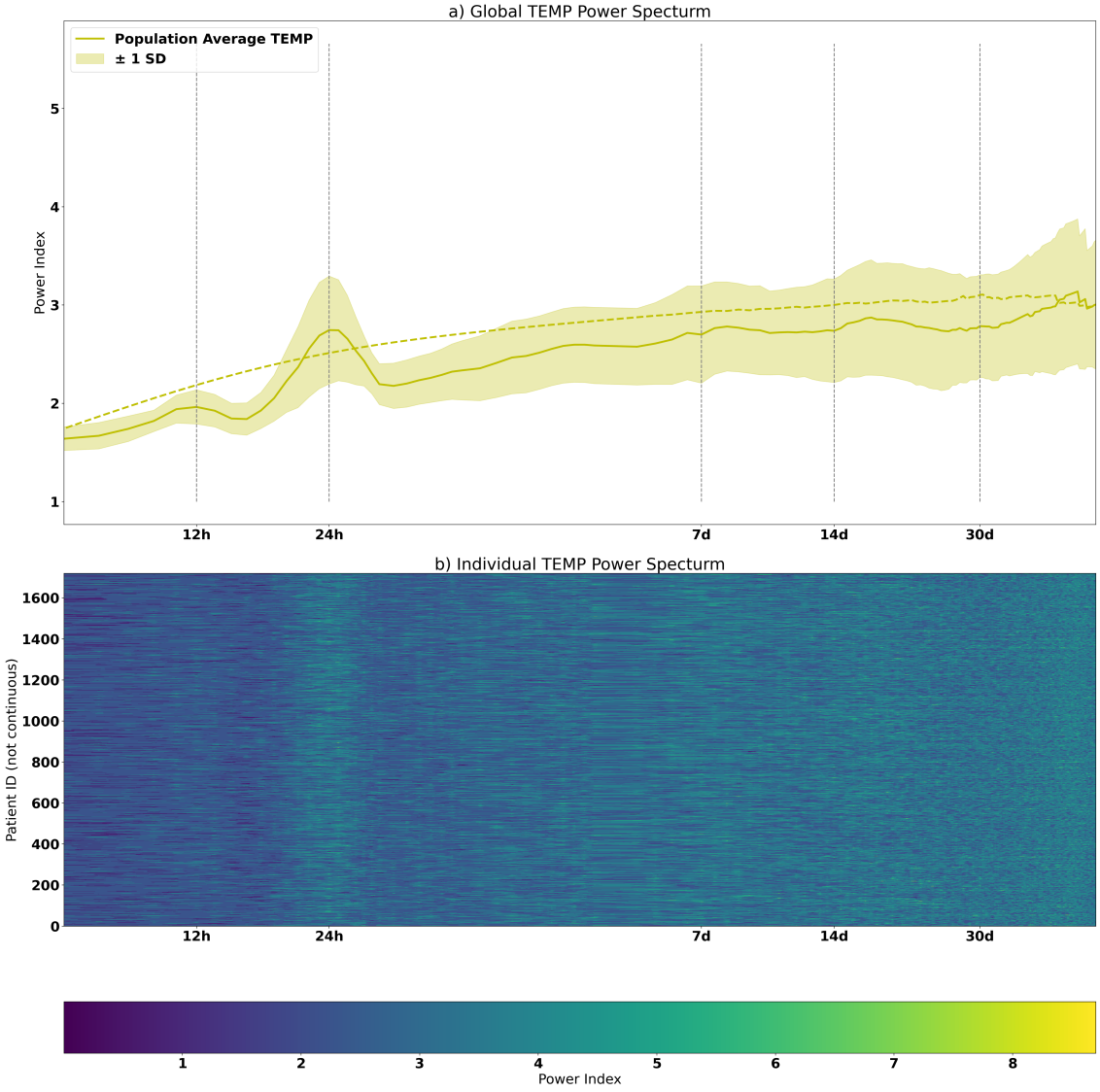


Figure 3-10: TEMP power spectra of the cohort. a): Global average spectrum over the 1,797 patients from Section 3.3. Shaded areas represent ±1 standard deviation. The dashed line represents the population-average upper 95% CI for TEMP. b): Individual TEMP spectra. Only the circadian cycle is shared across patients.

In the frequency domain, this addition corresponds to $X(f)+jY(f)$. The negative frequency components of $\hat{x}(t)$ are the conjugates of the positive frequency components of $x(t)$, and *vice versa*. Hence, the negative frequency components cancel out in $x_a(t)$, and amplitudes of the positive frequency components are doubled. Finally, suppose a patient had N GTCS events detected at $t_1, t_2, ..., t_N$, the corresponding instantaneous phases $\phi_1, \phi_2, ..., \phi_N$ of $x_a(t)$ can be computed by the argument function:

$$\phi_n = \arg[x_a(t_n)] \tag{3.7}$$

We used the Hilbert Transform function [2] from the `SciPy` [58] package.

Figure 3-11 and Figure 3-12 demonstrate the alignments of GTCS events with circadian, weekly, and about-monthly components in the EDA and weighted EDA of an example patient, respectively. Figure 3-13 shows the alignment of GTCS events with circadian, weekly, and about-monthly components in the TEMP of the same patient for comparison.

3.5.2 GTCS Phase-locking

Thus, for a particular period T and the band-pass filtered signal, a mean resultant vector can be calculated for each patient by averaging $\phi_1, \phi_2, ..., \phi_N$:

$$\frac{1}{N}\sum_{n=1}^{N} e^{i\phi_n} \tag{3.8}$$

The magnitude of the mean resultant vector quantifies the degree of "phase-locking" of GTCS timing with respect to the cycle of period T in physiological signals. It measures the distribution of phases at which GTCS occurred. For the rest of the thesis, we denote this magnitude as the phase-locking value (PLV), and it ranges from 0 ($\phi_1, \phi_2, ..., \phi_N$ are evenly distributed on $[0, 2\pi]$) to 1 ($\phi_1, \phi_2, ..., \phi_N$ are the same).

We used the Omnibus (Hodges-Ajne) test to calculate statistical significance for non-uniform angular distribution. The null hypothesis asserts a uniform circular distribution of $\phi_1, \phi_2, ..., \phi_N$ over $[0, 2\pi]$. In comparison to the commonly used Rayleigh test, the Omnibus test is non-

[2] `https://docs.scipy.org/doc/scipy/reference/generated/scipy.signal.hilbert.html`

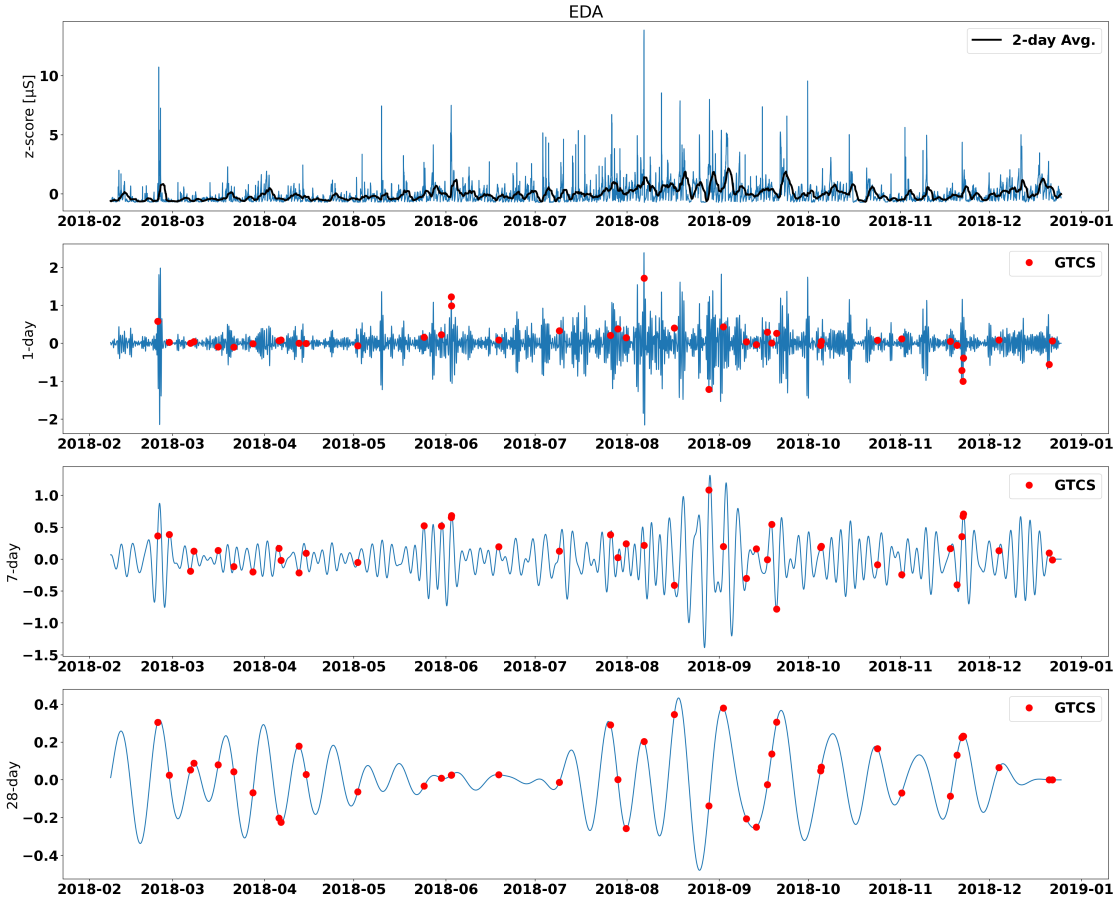


Figure 3-11: GTCS events aligning with EDA cycles. From the top to the bottom row: the z-scored EDA and two-day average, the 24-hour component from the Butterworth band-pass filter, the 7-day component, and the 28-day component.

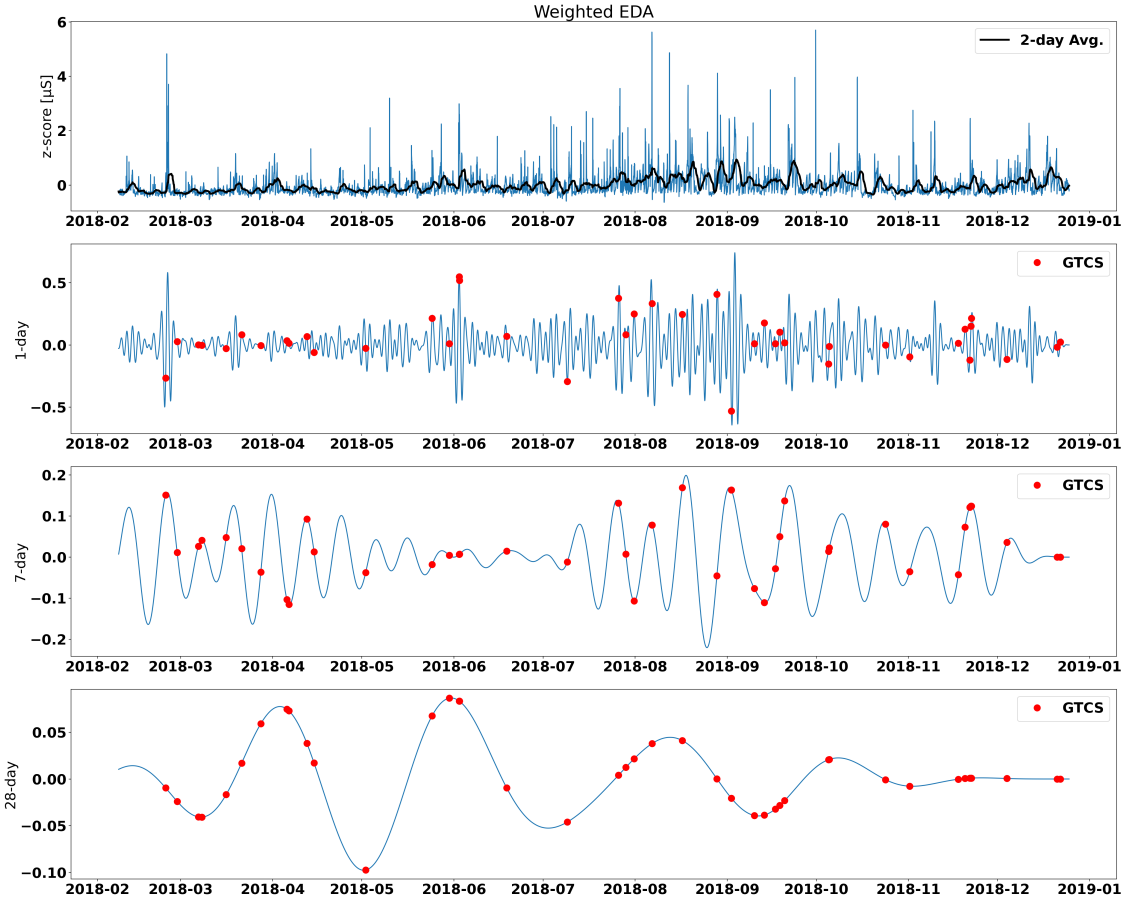


Figure 3-12: GTCS events aligning with weighted EDA cycles. From the top to the bottom row: the z-scored EDA and two-day average, the 24-hour component from the Butterworth band-pass filter, the 7-day component, and the 28-day component.

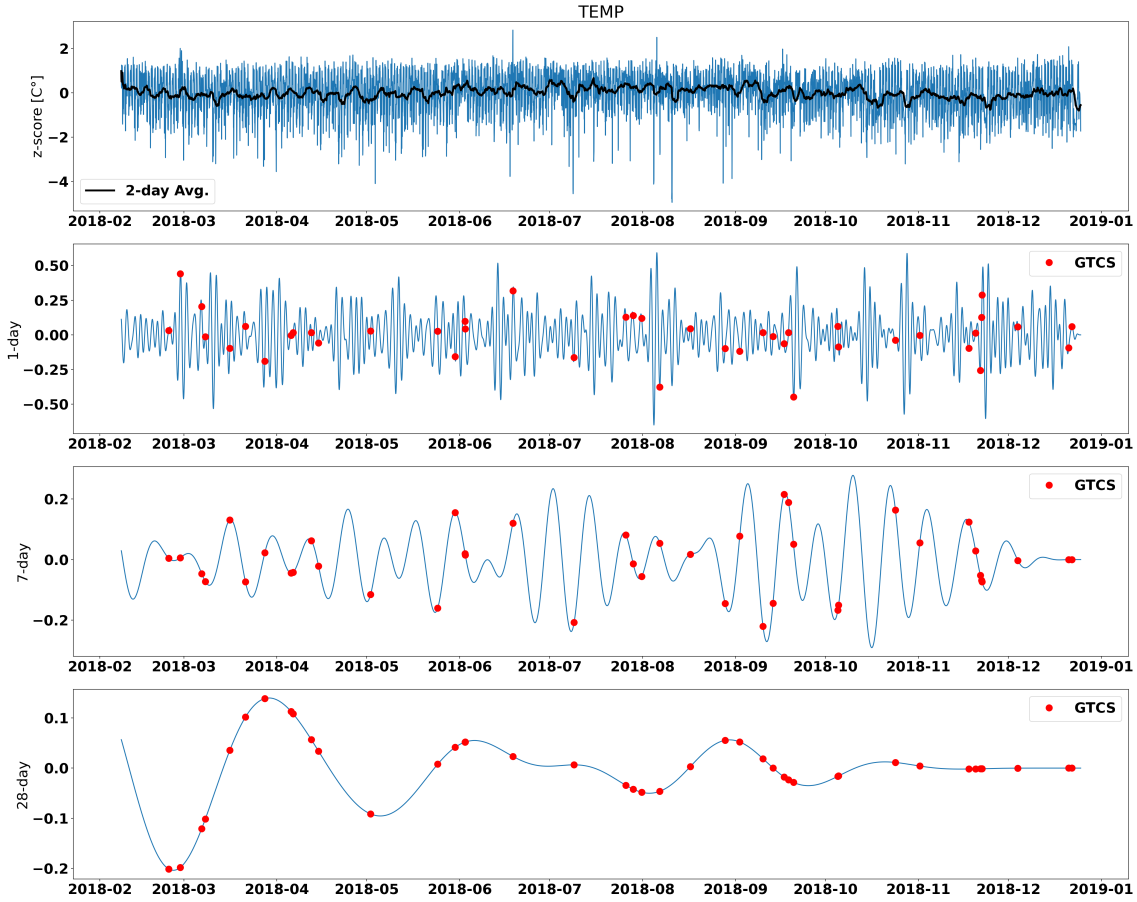


Figure 3-13: GTCS events aligning with TEMP cycles. From the top to the bottom row: the z-scored TEMP and two-day average, the 24-hour component from the Butterworth band-pass filter, the 7-day component, and the 28-day component.

	GTCS <>EDA	GTCS <>Weighted EDA
At least 1 cycle	86%	81%
More than 1 cycle	73%	70%
24-hour cycle	82%	75%
7-day cycle	21%	14%
28 to 32-day cycles	22%	18%

Table 3.2: Percentages of the patient population with GTCS occurrences significantly phase-locked to cycles detected in wrist EDA and weighted EDA.

parametric and weaker. Still, it does not require GTCS phases to follow a von Mises (circular normal) distribution, which is a strong assumption often unmet by the majority of the patients and periods analyzed [3, 28]. Figure 3-14 shows the angular distribution of GTCS phases of the example patient from Figure 3-11 for T = (a) 24 hours, (b) 7 days, and (c) 28 days. For visualization, GTCS phases were binned into 12 equal-sized bins on polar coordinates ($\frac{\pi}{6}$ each), forming a polar histogram where the bar height represents the number of GTCS events. Also, the mean resultant vectors are enlarged by multiplying the maximum bar height in each plot for visualization. Opaque bars on the polar plots mean the null hypothesis of the Omnibus test was not rejected with a Bonferroni correction, i.e., uniform circular distribution. For this patient, GTCS is significantly phase-locking to all 24-hour, 7-day, and 28-day EDA cycles. Significantly phase-locking also exists in 24-hour and 7-day weighted EDA rhythms and 28-day TEMP rhythm. For future work, we should scrutinize how GTCS events of this subject are phase-locked to the trough of the 28-day TEMP rhythm.

Table 3.2 shows percentages of the patient population with GTCS occurrences significantly phase-locked to cycles detected in wrist EDA and weighted EDA ($p < 0.05$, Omnibus test with a Bonferroni correction). GTCS phase-locking is only assessed for significant physiological cycles extracted above. More than 80% of the patients have at least one cycle detected in EDA or weighted EDA. GTCS phase-locking to circadian EDA or weighted EDA cycles is more prevalent than phase-locking to weekly or about-monthly cycles.

Across patients, the average PLV (length of the resultant vector) was similar between circadian (mean: 0.37) and multi-day (mean: 0.36) EDA rhythms ($p = 0.27$, Wilcoxon test). This suggests that GTCS phase-locking to multi-day EDA cycles is as strong as to circadian

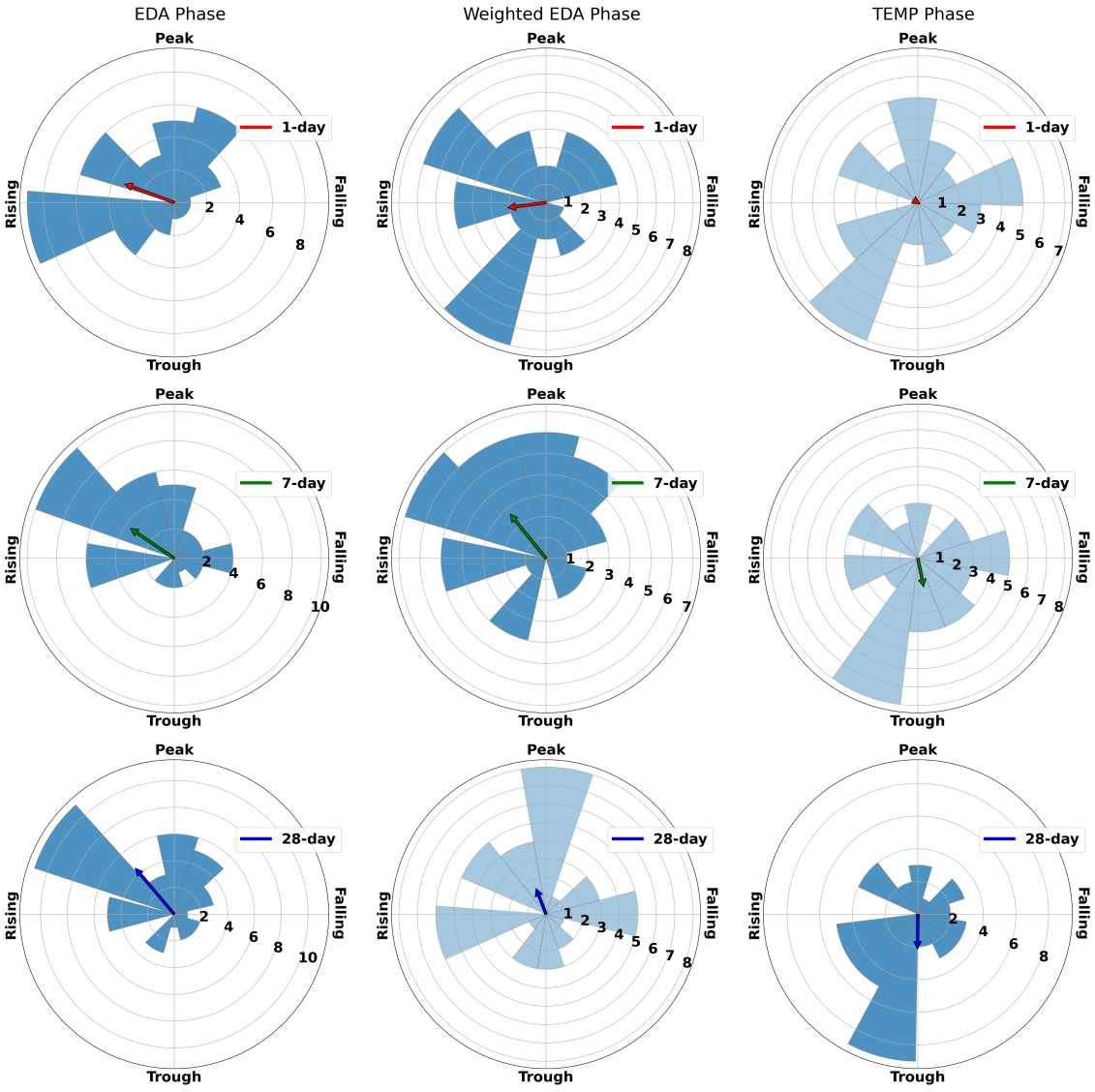
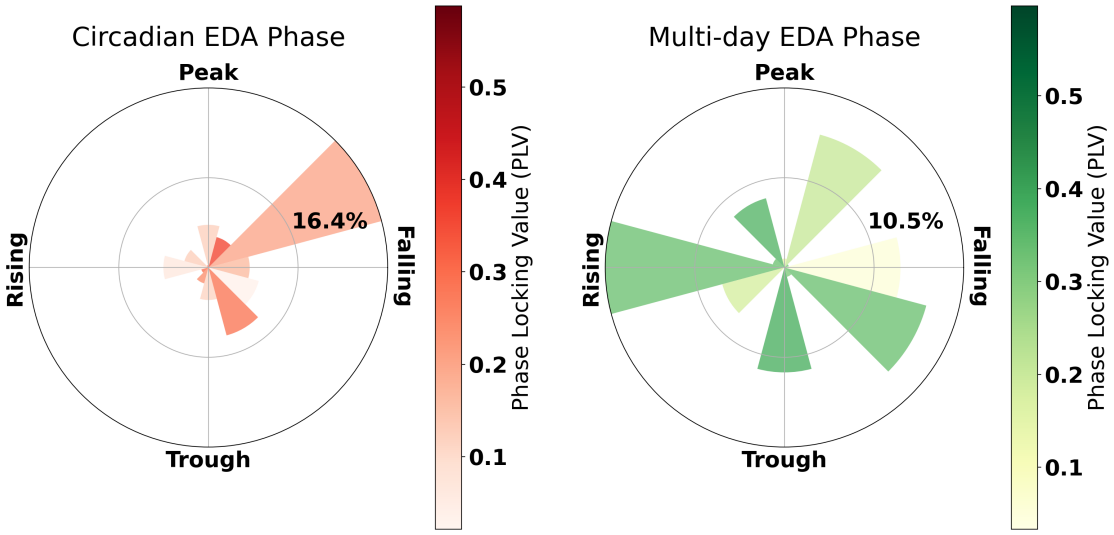


Figure 3-14: Angular distribution of GTCS phases of the example patient from Figure 3-11 for T of (a) 24 hours, (b) 7 days, and (c) 28 days. Opaque bars on the polar plots mean the null hypothesis of the Omnibus test was not rejected.

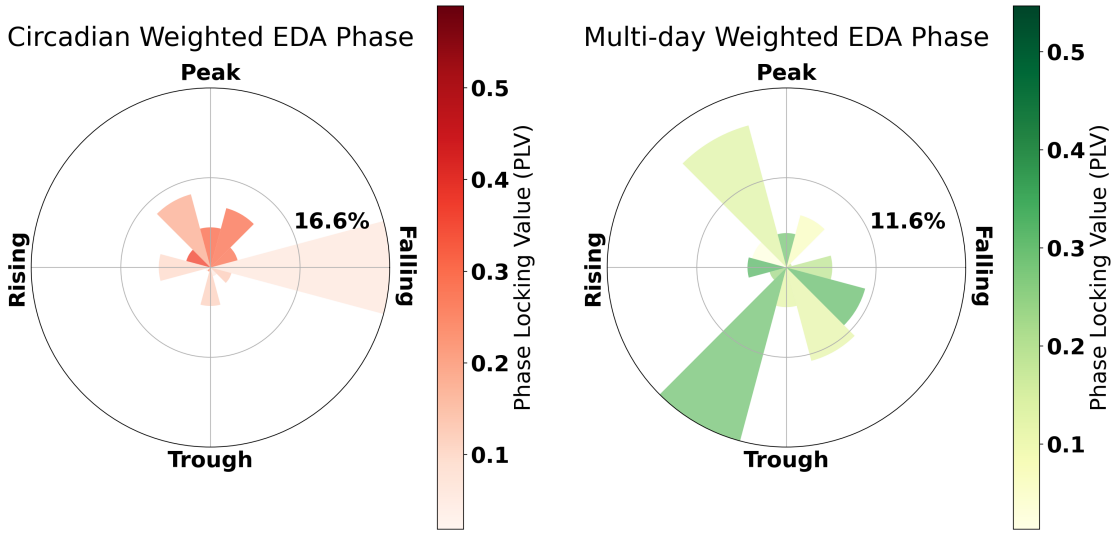


(a) (b)

Figure 3-15: GTCS occurrences with respect to phases in EDA cycles. Each bar on the polar coordinates represents the percentage of patients whose mean resultant vector lies within the $\pi/6$ range. All bars sum to 100%. The inner circle marks half of the highest bar.

EDA cycles. Moreover, the population average angular distributions of $\phi_1, \phi_2, ..., \phi_N$ are significantly different between circadian and multi-day EDA rhythms ($p = 0.007$, Kuiper's test). As presented in Figure 3-15a, GTCS occurrences tend to cluster on a narrow range of phases (pre-falling) of circadian EDA cycles. In Figure 3-15b, GTCS occurrences are more scattered across phases of multi-day EDA cycles, suggesting that the phase-locking phenomena over the long term are more patient-specific. Here, the shortest EDA cycle with $T > 24$ hours is selected for each patient as the multi-day rhythm. Each bar on the polar coordinates represents the percentage of patients whose mean resultant vector lies within the $\pi/6$ range. All bars sum to 100%. The inner circle marks half of the highest bar. This finding is consistent with the results of EDA cycles by Gregg et al. [20], though the proposed tonic-phasic component separation (based on low-pass filtering) may induce nuanced differences. In contrast, for intracranial IEA, GTCS events are coupled to multi-day IEA rhythms over a narrow (rising) range of phases [3].

For circadian cycles in EDA weighted by TEMP, the population-level distribution of mean resultant vectors did not alter drastically, and the biggest patient cluster is shifted to the falling phase, see Figure 3-16a. Examining this shift is required for future work. For multi-day cy-



(a) (b)

Figure 3-16: GTCS occurrences with respect to phases in weighted EDA cycles. Each bar on the polar coordinates represents the percentage of patients whose mean resultant vector lies within the $\pi/6$ range. All bars sum to 1. The inner circle marks half of the highest bar.

cles in EDA weighted by TEMP, the population-level distribution of mean resultant vectors is still scattered. Still, notable patient clusters are formed at different phases, comparing Figure 3-15b and Figure 3-16b. The average PLV was similar between circadian (mean: 0.28) and multi-day (mean: 0.29) EDA rhythms ($p = 0.23$, Wilcoxon test). Nonetheless, PLV for weighted EDA are both significantly lower than their EDA counterparts ($p < 0.05$, Wilcoxon test). The population average angular distributions of $\phi_1, \phi_2, ..., \phi_N$ are significantly different between circadian and multi-day weighted EDA rhythms ($p = 0.006$, Kuiper's test).

3.5.3 Physiological Signal Phase-locking to Clock Time

For $T = 24$ hours, we also explore how the peaks of each signal modality align with the 24-hour clock time on the population level. Figure 3-17 shows the mean resultant vectors (average over the whole cohort) of peaks of four signal modalities. The peak of the 24-hour TEMP cycle predominately happens a little past 0 A.M. ($p < 0.001$, Omnibus test). A potential explanation may be the thermal regulation of human bodies during sleep, and peripheral temperature usually increases as core body temperature drops [51]. The peak of

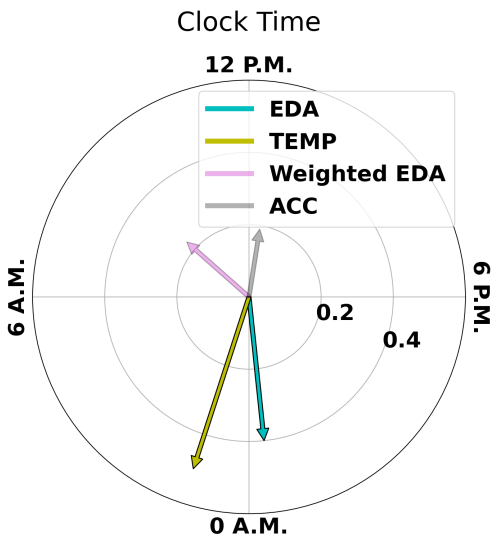


Figure 3-17: Mean resultant vectors (average over the patient cohort) of 24-hour peaks of four signal modalities aligning with clock time. Opaque arrows represent insignificant alignment failing Omnibus tests.

the 24-hour EDA cycle mostly happens close to 0 A.M. ($p < 0.001$, Omnibus test), possibly due to EDA sleep storms [50]. Peaks of weighted EDA ($p = 0.14$, Omnibus test) and ACC ($p = 0.06$, Omnibus test) did not show any preferred alignment with clock time. ACC is quantified by an existing method [3] to capture activity intensity.

3.6 Rhythmic Patterns in GTCS Alerts

So far, the above analyses validated previous findings of seizure occurrence phase-locking to cycles in physiological signals with a large patient cohort. However, previous studies either only explored cycles in physiological signals [30, 20] or cycles in seizure reports [28] but not both. In this section, we look for rhythmic patterns in GTCS alerts from the same cohort. Such investigation will serve as a complement to the above wavelet analyses, and it will also facilitate a simple benchmark for evaluating phase-locking properties. Karoly et al. studied patients who self-reported at least 100 seizures [28], and we also only consider patients with more than 100 GTCS detected below. A subset of 877 patients are selected from the cohort of Section 3.3, and their longitudinal GTCS alert reports range from 1.5 to 3.7 years (median: 3.1 years).

[3] https://support.empatica.com/hc/en-us/articles/202028739-How-is-the-acceleration-data-formatted-in-E4-conne

3.6.1 Cycles in GTCS Alerts

As stated by Karoly et al. [28], mean resultant vectors and PLV can be calculated solely based on seizure reports. In particular, GTCS phases $\phi_1, \phi_2, ..., \phi_N$ can be obtained by $2\pi \times (t_n \mod T)$ where t_n is the time elapsed leading to each GTCS event. It is worth noticing that self-reported seizure diaries are prone to missed events and personal bias. Following Karoly et al. [28] in this section, we analyze T with increasing spacing: 6 hours between 6 to 24 hours and one day between 2 to 32 days. The PLV is computed for longitudinal GTCS alerts of each patient according to Equation 3.8. Figure 3-18 shows the raster plot of PLV from all patients. The y-axis shows only the patient ID and is not continuous. Despite the difference in the nature of seizures recorded (objective alerts from wearable device versus subjective seizure diary), the population-level distribution of PLV in this cohort coincides with the one reported by Karoly et al. [28] where circadian and weekly cycles are predominant. Such validation may inform future seizure forecasting technology to produce weekly and monthly forecasts separately.

3.6.2 Monthly GTCS Frequency

In addition, we perform some common statistical analyses to better describe GTCS events detected by wearable devices. The distribution of monthly GTCS rate over the cohort ($>$ 100 GTCS detected per patient) is shown in Figure 3-19. The cohort has a median number of GTCS per month of 3.8 (range 0.1-28). For a subset of patients with high rates (> 4 GTCS events per month), the median rate is 7. Despite the difference in the objective and subjective seizures recorded, these rates are similar to previously reported large-scale seizure diaries [15]. These rates reflect the heterogeneity of seizure frequency between patients.

3.6.3 *L*-relationship

Counting the number of seizures self-reported by each patient every month, the linear relationship between the logarithm of mean and standard deviation of seizure count per month has been affirmed by a previous large-scale study of seizure diary [18]. It is denoted as the L-relationship, and the linear coefficient r lies in $[7.3, 8.3]$. Figure 3-20 complied with this claim with a fitted linear regression line over the cohort (Pearson $r = 0.73$, $p < 0.001$). While seizure frequency may vary drastically across age or gender groups, stratified analyses

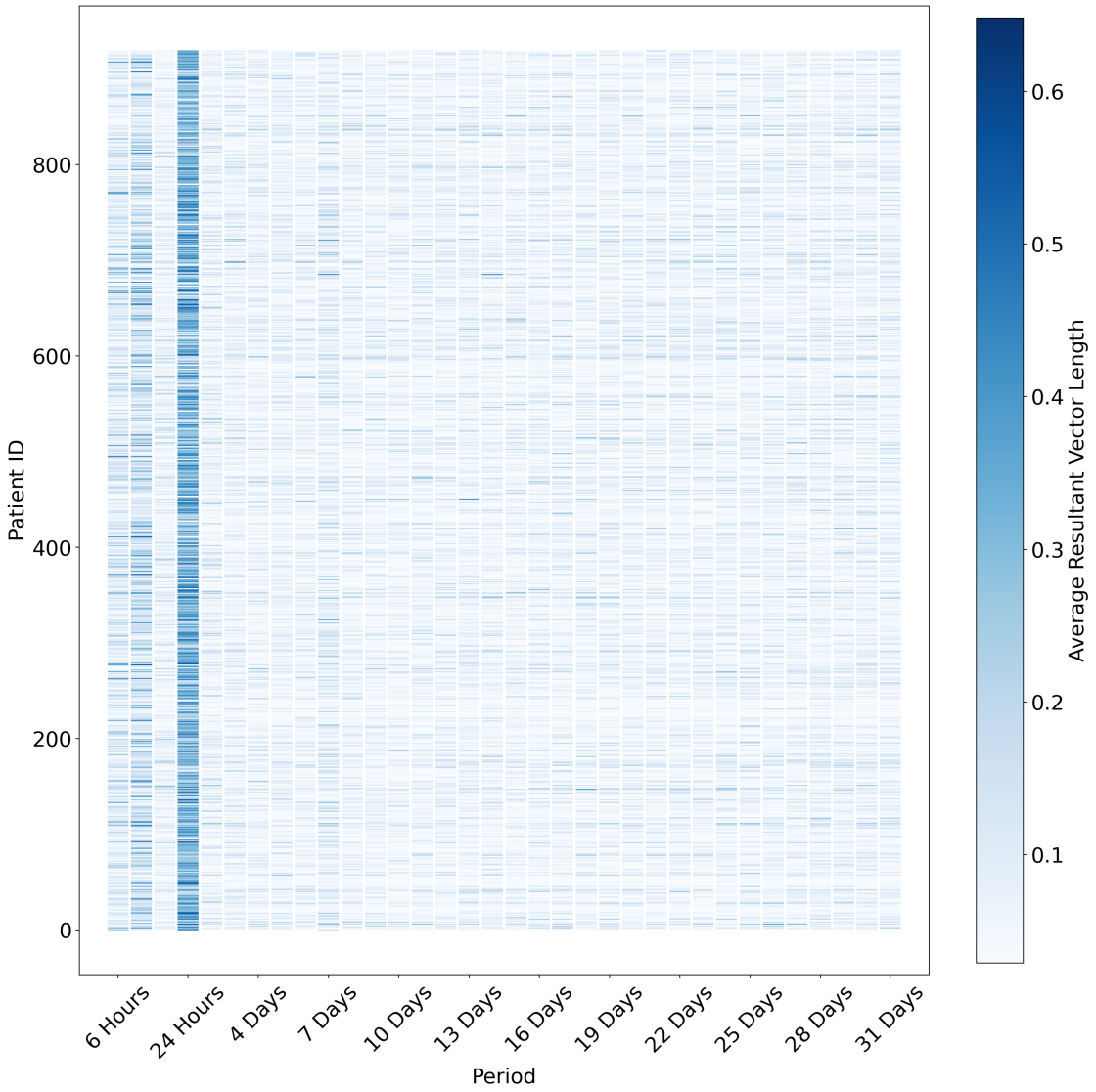


Figure 3-18: Raster plot of PLV estimated from longitudinal GTCS alerts detected by the Embrace 2 wristband of each patient.

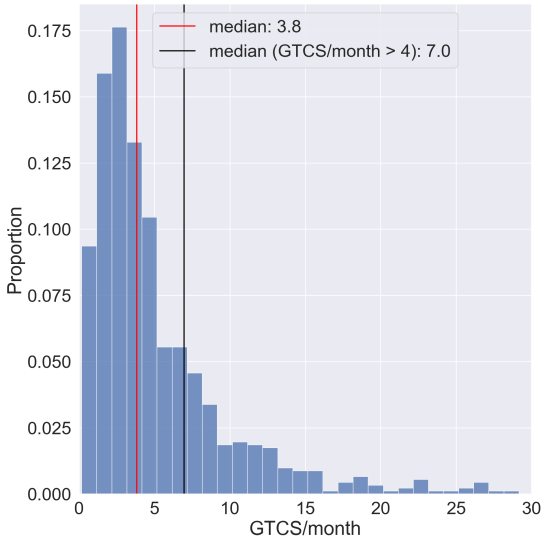


Figure 3-19: Distribution of monthly GTCS rate in the patient population. Each patient has at least 100 GTCS detected.

based on age, gender, or other demographic information are not available due to privacy regulations.

3.7 Rethinking Phase-locking with Simulated GTCS

In this section, we ask a simple question: "Given a series of GTCS events phase-locking to a physiological rhythm, what is the chance of a series of *random* GTCS events also phase-locking to the rhythm?" This question becomes non-trivial when investigating rhythms of large T and binned phases of lower granularity. Intuitively, given three months of EDA, a significant monthly EDA rhythm, and four binned phases (rising, peak, falling, and trough), the chance of N simulated GTCS occurring at the same phase is not negligible, especially when N is small. It challenges the significance of PLV computed based on the above procedure, as PLV was supposed to shed light on latent continuous GTCS risk modulation. This question is a facet of simultaneously and *independently* assessing the presence of rhythm of T in physiological signals and seizure reports, as mentioned above. It is possible that a rhythm of T is discovered in EDA but not found in the GTCS report, and GTCS phases are still phase-locking to the $\frac{1}{T}$ frequency component in EDA. The phase-locking significance established by Omnibus tests merely states that the distribution of angles is not circular-uniform.

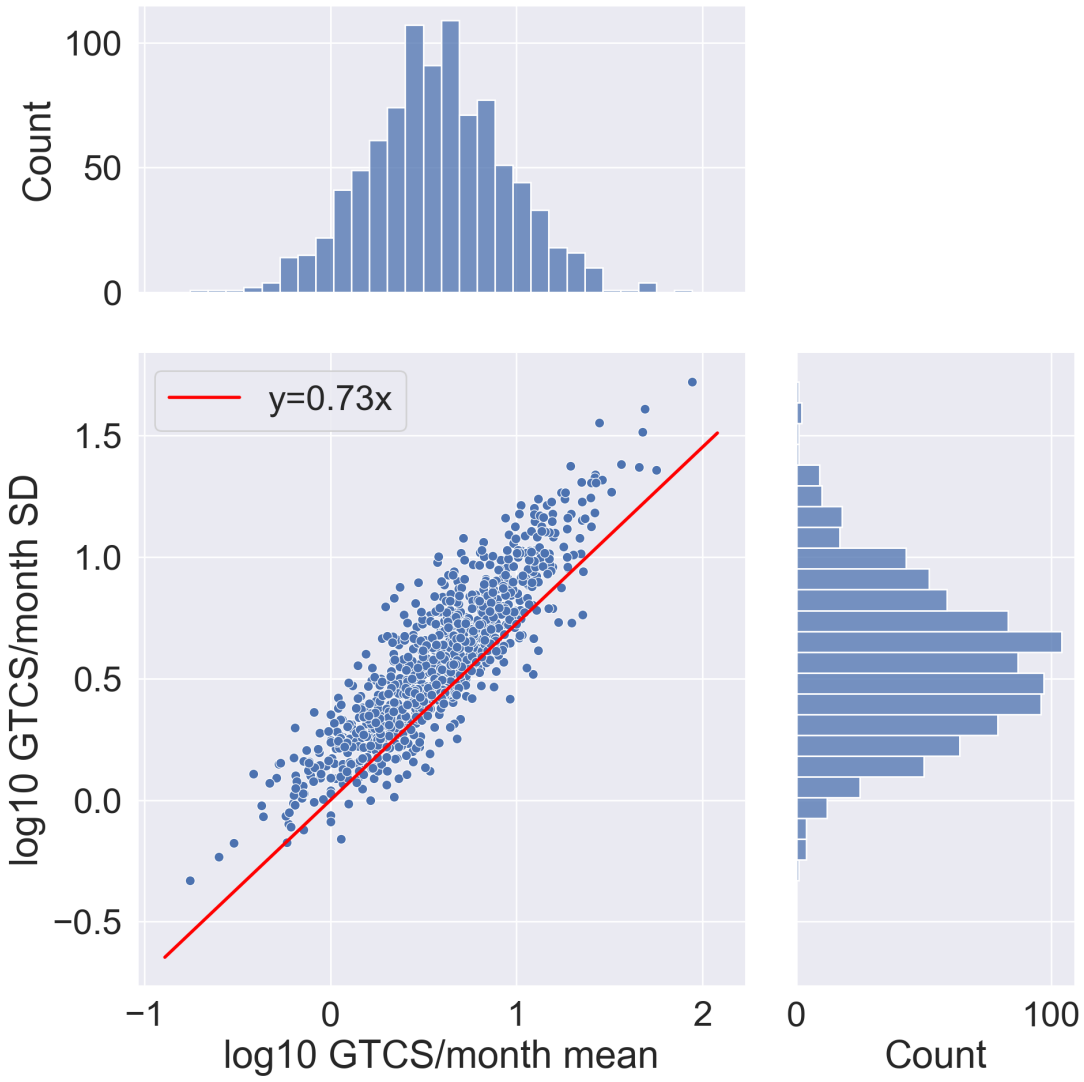


Figure 3-20: *L*-relationship between the logarithm of mean and standard deviation of GTCS count per month. This relationship is consistent between objective GTCS alerts and subjective seizure diaries [18].

Ultimately, we propose a simple benchmark to augment circular-uniformity tests when evaluating PLV, inspired by Karoly et al. [28]. For a patient with a significant physiological rhythm of T and N GTCS events detected, a negative binomial distribution is fitted to describe the time intervals between GTCS events via maximum likelihood estimation. Previous literature found that negative binomial distributions perform better than Poisson distributions at characterizing point seizure events from patient diaries [28, 55]. The negative binomial distribution is often used to describe the number of "failures" before the nth "success" in a sequence of Bernoulli trials. The probability mass function of a negative binomial distribution is given by:

$$P(k_n|n,p) = \binom{k_n+n-1}{k_n}(1-p)^{k_n}p^n \text{ where } k_n, n, p > 0 \tag{3.9}$$

Here, n is the number of GTCS events, p is the probability of GTCS, and k_n is the number of hours before observing n GTCS events. Assuming that each event happened independently, the likelihood of observing a specific GTCS report of length N is the product of the individual events. The log-likelihood function is then given by:

$$\begin{aligned}\mathcal{L}(n,p) &= \sum_{n=1}^{N} \log P(k_n|n,p) \\ &= \sum_{n=1}^{N} \log\left(\binom{k_n+n-1}{k_n}(1-p)^{k_n}p^n\right)\end{aligned} \tag{3.10}$$

Our goal is to find n and p that maximize this objective. This optimization problem has no closed-form solution. Instead, we solved it numerically using iterative methods. We used the `statsmodels` package [4] in Python [52] with the Broyden-Fletcher-Goldfarb-Shanno method for optimization.

Next, 10^4 simulated GTCS reports are drawn from the fitted negative binomial distribution of the patient, each containing N attacks. A PLV is calculated for each period T in each simulated report. These simulations did not incorporate any rhythmic modeling and created an upper 95% CI for PLV at period T. Finally, we calculate the percentage of patients with PLV greater than the upper 95% CI at each T, and Table 3.3 below shows the result. Numbers in parentheses represent the percentage change of patient count from Table 3.2.

[4] `https://www.statsmodels.org/dev/generated/statsmodels.discrete.discrete_model.NegativeBinomial`

	GTCS <>EDA	GTCS <>Weighted EDA
At least 1 cycle	52% (-34%)	55% (-26%)
More than 1 cycle	48% (-25%)	44% (-26%)
24-hour cycle	21% (-61%)	34% (-41%)
7-day cycle	7% (-14%)	9% (-5%)
28 to 32-day cycles	10% (-12%)	8% (-10%)

Table 3.3: Percentage of patients with PLV greater than the upper 95% CI generated by simulations. Numbers in parentheses represent the percentage change of patients from Table 3.2.

All changes are negative, meaning the 95% CI from simulated GTCS events sets a stricter significance level for evaluating phase-locking. This finding complies with Leguia et al. [33]: While circular distributions of seizures tend to occur in temporal clusters, passing non-uniform statistical tests does not necessarily entail seizure cycles. More rigorous testing is needed to scrutinize phase-locking properties.

Chapter 4

Conclusions and Future Work

4.1 Summary of Findings

In this thesis, we discussed two types of rhythms associated with GTCS events detected by smart wristbands: (i) sleep-wake cycles and next-day GTCS risk and (ii) multi-day rhythms in EDA and GTCS phase-locking. All sleep data, physiological signals, and GTCS occurrences are recorded by smart wristbands in ambulatory settings. The conclusions are drawn from over 1,000 to 2,000 patients diagnosed with GTCS, depending on the availability of specific analyses. All GTCS events are detected by an FDA-cleared algorithm and from compliant users who reported false-positive alerts on a regular basis.

In Chapter 2, we studied how sleep duration, efficiency, fragmentation, and regularity are related to next-day GTCS risk. Sleep parameters are estimated by published activity-based algorithms [46]. Data preprocessing involved filtering for compliant users (false positive alerts) and valid sleep episodes. By comparing ratios between normal and inferior sleep followed by GTCS and GTCS-free days, the four sleep parameters showed various relationships with next-day GTCS risk. For sleep duration, elevated GTCS risk was found for TST less than 6 hours, 7 hours, and the lowest 10^{th} percentile personal threshold. For sleep fragmentation, next-day GTCS risk is correlated with SFI greater than 0.2 and 0.3. For sleep efficiency, the lowest 10^{th} percentile personal threshold is associated with elevated next-day GTCS risk. For sleep regularity, the lowest 10^{th} percentile personal SRI over 2 and 3 days prior to GTCS events showed significant results. The magnitude of elevated risk was

marginal for all sleep parameters across the patient cohort. Potential confounding factors are discussed, and both population and personal thresholds are tested.

In Chapter 3, we studied the prevalence of multi-day physiological rhythms over the patient cohort, patient-specific periodicity, and GTCS detected by wristbands phase-locking to EDA rhythms. Data quality checks included cleaning physiological signals, such as artifact removal, and user compliance screening, the same as in the previous chapter. EDA rhythms are estimated by established wavelet analyses [3, 20, 30], and we also considered EDA weighted by TEMP in an effort to eliminate confounding effects of peripheral temperature. On the population level, the prevalence of long-term EDA rhythms is consistent with previous literature [20]. GTCS phase-locking to long-term EDA was also discovered, but the periodicity and preferred phases of occurrence are patient-specific. Ultimately, we discovered that the population statistics are similar between objective GTCS alerts detected by wristbands and large-scale seizure diaries, including monthly frequency and rhythmic patterns. We proposed a simple simulation method to evaluate the significance of the PLV estimated.

4.2 Future Work

There are a considerable number of topics worth investigating to augment the analyses in this thesis. From a signal-processing perspective, we can adapt better EDA artifact removal and interpolation methods to process long-term physiological data, such as supervised machine-learning approaches [54] and deep-learning multi-modal auto-encoders [27]. Also, given the flexibility of wavelet analysis, other waveforms can be tailored to extract rhythms from EDA signals. The Morlet wavelet may not be the optimal choice, depending on the intrinsic shape of the long-term signal. The fitness of waveforms can be assessed by the greedy matching pursuit method. For determining significant peaks in individual spectra, a surrogate time series approach may be more rigorous than a simple red noise process. For testing significant GTCS phase-locking, Leguia et al. have found that the commonly used statistical tests are too sensitive and easy to significance [33]. Instead, a surrogate data strategy may be more helpful.

Additionally, we have not explored the influence of naps on next-day GTCS events. All naps

are excluded from analyses in Chapter 2, and their compensation effect remains inconclusive. For sleep regularity, other validated methods may be suitable for studying this cohort, such as Composite Phase Deviation (CPD) [41, 16]. CPD can be measured daily, characterizing circadian disruptions. It is closely related to the bedtime and wake time onset irregularity in [53] because it measures how much the current sleep deviates away from the habitual schedule of each patient. Particularly, for each patient, a chronotype $\overline{MS}$ is first determined by the average mid-sleep time point adjusted for the day of the week and any shift work schedule. For day i with mid-sleep time MS_i, $\Delta\text{Reference} = \overline{MS} - MS_i$ measures the change w.r.t. the individualized average. The deviation from previous day is defined as $\Delta\text{Day-to-day} = MS_{i-1} - MS_i$. Then CPD for the day i is denoted by

$$CPD_i = \sqrt{\Delta\text{Reference}^2 + \Delta\text{Day-to-day}^2} \tag{4.1}$$

GTCS risk after circadian disruption quantified by CPD may be distinguishable between patients resilient to sudden changes in sleep conditions and patients with less tolerance. It will also be interesting to replicate the bed and wake onset regularity measurements by Stirling et al. [53] and compare these to the SRI and CPD.

Most importantly, we have not integrated the proposed sleep or rhythm parameters into GTCS forecasting algorithms to assess the actual predictive power. For example, phase-phase analyses were conducted to assess if combined circadian and multi-day peaks will increase seizure risk [20, 3], and such phase information has been incorporated into deep-learning seizure forecast models [7]. Likewise, sleep duration, efficiency, and fragmentation from the previous day may serve as input features to data-driven next-day GTCS risk forecasts. While we have explored individual sleep parameters above, combinations of multiple parameters may provide greater information. To what extent sleep parameters can provide useful information for seizure forecasts is a critical yet unsolved problem.

Finally, the underlying source that modulates multi-day EDA rhythms is unclear. While GTCS risks are shown to be phase-locked to certain phases in long-term rhythms, a causal relationship is lacking. Several studies have proposed possible endogenous drivers for long-term rhythms in physiological signals [29, 45], but their relationship with external multi-day periods, such as lunar cycles, is not fully understood. Currently, it is still unknown how to

perform randomized controlled interventions on multi-day cycles of patients with epilepsy.

Bibliography

[1] Christopher T Anderson, Thomas K Tcheng, Felice T Sun, and Martha J Morrell. Day–night patterns of epileptiform activity in 65 patients with long-term ambulatory electrocorticography. *Journal of Clinical Neurophysiology*, 32(5):406–412, 2015.

[2] G Aubert-Tulkens, C Culee, K Harmant-Van Rijckevorsel, and DO Rodenstein. Ambulatory evaluation of sleep disturbance and therapeutic effects in sleep apnea syndrome by wrist activity monitoring. *American Review of Respiratory Disease*, 136(4):851–856, 1987.

[3] Maxime O Baud, Jonathan K Kleen, Emily A Mirro, Jason C Andrechak, David King-Stephens, Edward F Chang, and Vikram R Rao. Multi-day rhythms modulate seizure risk in epilepsy. *Nature communications*, 9(1):1–10, 2018.

[4] Carl W Bazil, Douglas Short, David Crispin, and Wei Zheng. Patients with intractable epilepsy have low melatonin, which increases following seizures. *Neurology*, 55(11):1746–1748, 2000.

[5] Nicholas A Bercel. The periodic features of some seizure states. *Annals of the New York Academy of Sciences*, 1964.

[6] Sebastian Böttcher, Solveig Vieluf, Elisa Bruno, Boney Joseph, Nino Epitashvili, Andrea Biondi, Nicolas Zabler, Martin Glasstetter, Matthias Dümpelmann, Kristof Van Laerhoven, et al. Data quality evaluation in wearable monitoring. *Scientific reports*, 12(1):21412, 2022.

[7] Benjamin Brinkmann, Ewan Nurse, Mona Nasseri, Pedro Viana, Philippa Karoly, Tal Pal Attia, Nicholas Gregg, Boney Joseph, Caitlin Grzeskowiak, Matthias Dumpelmann, et al. We-120. seizure forecasting and detection with wearable devices and subcutaneous EEG-a practical seizure gauge. *Clinical Neurophysiology*, 141:S62–S63, 2022.

[8] Benjamin Brinkmann, Ewan Nurse, Pedro Viana, Mona Nasseri, Levin Kuhlmann, Philippa Karoly, Tal Pal Attia, Nicholas Gregg, Boney Joseph, Matthias Dumpelmann, et al. Seizure forecasting and detection with wearable devices and subcutaneous EEG-outcomes from the my seizure gauge trial (pl4. 001), 2023.

[9] Paul R Carney, Sachin S Talathi, Dong-Uk Hwang, and William Ditto. 19 circadian regulation of neural excitability in temporal lobe epilepsy. *Epilepsy: The Intersection of Neurosciences, Biology, Mathematics, Engineering, and Physics*, page 289, 2016.

[10] Jean-Philippe Chaput, Caroline Dutil, and Hugues Sampasa-Kanyinga. Sleeping hours: what is the ideal number and how does age impact this? *Nature and science of sleep*, pages 421–430, 2018.

[11] Katrina L Dell, Daniel E Payne, Vaclav Kremen, Matias I Maturana, Vaclav Gerla, Petr Nejedly, Gregory A Worrell, Lhotska Lenka, Filip Mivalt, Raymond C Boston, et al. Seizure likelihood varies with day-to-day variations in sleep duration in patients with refractory focal epilepsy: A longitudinal electroencephalography investigation. *EClinicalMedicine*, 37:100934, 2021.

[12] Jonas M den Heijer, Willem M Otte, Eric van Diessen, Jolien S van Campen, E Lorraine Hompe, Floor E Jansen, Marian Joels, Kees PJ Braun, Josemir W Sander, and Maeike Zijlmans. The relation between cortisol and functional connectivity in people with and without stress-sensitive epilepsy. *Epilepsia*, 59(1):179–189, 2018.

[13] Tyler S Durazzo and Hitten P Zaveri. 37 seizure prediction and the circadian rhythm. *Ivan Osorio• Hitten P. Zaveri*, page 489, 2011.

[14] James H Dwyer. Analysis of variance and the magnitude of effects: A general approach. *Psychological Bulletin*, 81(10):731, 1974.

[15] Victor Ferastraoaru, Daniel M Goldenholz, Sharon Chiang, Robert Moss, William H Theodore, and Sheryl R Haut. Characteristics of large patient-reported outcomes: where can one million seizures get us? *Epilepsia open*, 3(3):364–373, 2018.

[16] Dorothee Fischer, Elizabeth B Klerman, and Andrew JK Phillips. Measuring sleep regularity: theoretical properties and practical usage of existing metrics. *Sleep*, 44(10):zsab103, 2021.

[17] Birgit Frauscher and Jean Gotman. Sleep, oscillations, interictal discharges, and seizures in human focal epilepsy. *Neurobiology of disease*, 127:545–553, 2019.

[18] Daniel M Goldenholz, Shira R Goldenholz, Robert Moss, Jacqueline French, Daniel Lowenstein, Ruben Kuzniecky, Sheryl Haut, Sabrina Cristofaro, Kamil Detyniecki, John Hixson, et al. Is seizure frequency variance a predictable quantity? *Annals of clinical and translational neurology*, 5(2):201–207, 2018.

[19] Daniel M Goldenholz, Kshitiz Rakesh, Kush Kapur, Marina Gaínza-Lein, Ryan Hodgeman, Robert Moss, William H Theodore, and Tobias Loddenkemper. Different as night and day: patterns of isolated seizures, clusters, and status epilepticus. *Epilepsia*, 59(5):e73–e77, 2018.

[20] Nicholas M Gregg, Tal Pal Attia, Mona Nasseri, Boney Joseph, Philippa Karoly, Jie Cui, Rachel E Stirling, Pedro F Viana, Thomas J Richner, Ewan S Nurse, et al. Seizure occurrence is linked to multiday cycles in diverse physiological signals. *Epilepsia*, 2023.

[21] GwenvronM Griffiths and J Tylor Fox. Rhythm in epilepsy. *The Lancet*, 232(5999):409–416, 1938.

[22] Caitlin L Grzeskowiak and Sonya B Dumanis. Seizure forecasting: patient and caregiver perspectives. *Frontiers in neurology*, 12, 2021.

[23] Winston Haynes. Bonferroni correction. *Encyclopedia of systems biology*, pages 154–154, 2013.

[24] Dale C Hesdorffer, Torbjorn Tomson, Emma Benn, Josemir W Sander, Lena Nilsson, Yvonne Langan, Thaddeus S Walczak, Ettore Beghi, Martin J Brodie, W Allen Hauser, et al. Do antiepileptic drugs or generalized tonic-clonic seizure frequency increase SUDEP risk? a combined analysis. *Epilepsia*, 53(2):249–252, 2012.

[25] Wytske A Hofstra, Marijke CM Gordijn, Job van der Palen, Renate van Regteren, Bertine E Grootemarsink, and Al W de Weerd. Timing of temporal and frontal seizures in relation to the circadian phase: a prospective pilot study. *Epilepsy research*, 94(3):158–162, 2011.

[26] Wytske A Hofstra, Bertine E Grootemarsink, Rianneke Dieker, Job Van Der Palen, and Al W De Weerd. Temporal distribution of clinical seizures over the 24-h day: a retrospective observational study in a tertiary epilepsy clinic. *Epilepsia*, 50(9):2019–2026, 2009.

[27] Natasha Jaques, Sara Taylor, Akane Sano, and Rosalind Picard. Multimodal autoencoder: A deep learning approach to filling in missing sensor data and enabling better mood prediction. In *2017 Seventh International Conference on Affective Computing and Intelligent Interaction (ACII)*, pages 202–208. IEEE, 2017.

[28] Philippa J Karoly, Daniel M Goldenholz, Dean R Freestone, Robert E Moss, David B Grayden, William H Theodore, and Mark J Cook. Circadian and circaseptan rhythms in human epilepsy: a retrospective cohort study. *The Lancet Neurology*, 17(11):977–985, 2018.

[29] Philippa J Karoly, Vikram R Rao, Nicholas M Gregg, Gregory A Worrell, Christophe Bernard, Mark J Cook, and Maxime O Baud. Cycles in epilepsy. *Nature Reviews Neurology*, 17(5):267–284, 2021.

[30] Philippa J Karoly, Rachel E Stirling, Dean R Freestone, Ewan S Nurse, Matias I Maturana, Amy J Halliday, Andrew Neal, Nicholas M Gregg, Benjamin H Brinkmann, Mark P Richardson, et al. Multiday cycles of heart rate are associated with seizure likelihood: An observational cohort study. *EBioMedicine*, 72:103619, 2021.

[31] Mary Langdon-Down and W Russell Brain. Time of day in relation to convulsions in epilepsy. *The Lancet*, 213(5516):1029–1032, 1929.

[32] Marc G Leguia, Ralph G Andrzejak, Christian Rummel, Joline M Fan, Emily A Mirro, Thomas K Tcheng, Vikram R Rao, and Maxime O Baud. Seizure cycles in focal epilepsy. *JAMA neurology*, 78(4):454–463, 2021.

[33] Marc G Leguia, Vikram R Rao, Jonathan K Kleen, and Maxime O Baud. Measuring synchrony in bio-medical timeseries. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 31(1):013138, 2021.

[34] A Molina-Carballo, A Munoz-Hoyos, M Sanchez-Forte, J Uberos-Fernández, F Moreno-Madrid, and D Acuña-Castroviejo. Melatonin increases following convulsive seizures may be related to its anticonvulsant properties at physiological concentrations. *Neuropediatrics*, 38(03):122–125, 2007.

[35] Mona Nasseri, Ewan Nurse, Martin Glasstetter, Sebastian Böttcher, Nicholas M Gregg, Aiswarya Laks Nandakumar, Boney Joseph, Tal Pal Attia, Pedro F Viana, Elisa Bruno, et al. Signal quality and patient experience with wearable devices for epilepsy management. *Epilepsia*, 61:S25–S35, 2020.

[36] Marcus Ng and Milena Pavlova. Why are seizures rare in rapid eye movement sleep? review of the frequency of seizures in different sleep stages. *Epilepsy research and treatment*, 2013, 2013.

[37] Lino Nobili, Birgit Frauscher, Sofia Eriksson, Steve Alex Gibbs, Peter Halasz, Isabelle Lambert, Raffaele Manni, Laure Peter-Derex, Paola Proserpio, Federica Provini, et al. Sleep and epilepsy: A snapshot of knowledge and future research lines. *Journal of Sleep Research*, page e13622, 2022.

[38] Francesco Onorati, Giulia Regalia, Chiara Caborni, Matteo Migliorini, Daniel Bender, Ming-Zher Poh, Cherise Frazier, Eliana Kovitch Thropp, Elizabeth D Mynatt, Jonathan Bidwell, et al. Multicenter clinical assessment of improved wearable multimodal convulsive seizure detectors. *Epilepsia*, 58(11):1870–1879, 2017.

[39] Frederick L Patry. The relation of time of day, sleep, and other factors to the incidence of epileptic seizures. *American Journal of Psychiatry*, 87(5):789–813, 1931.

[40] Milena K Pavlova, Steven A Shea, Frank AJL Scheer, and Edward B Bromfield. Is there a circadian variation of epileptiform abnormalities in idiopathic generalized epilepsy? *Epilepsy & Behavior*, 16(3):461–467, 2009.

[41] Andrew JK Phillips, William M Clerx, Conor S O'Brien, Akane Sano, Laura K Barger, Rosalind W Picard, Steven W Lockley, Elizabeth B Klerman, and Charles A Czeisler. Irregular sleep/wake patterns are associated with poorer academic performance and delayed circadian and sleep/wake timing. *Scientific reports*, 7(1):1–13, 2017.

[42] Rosalind W Picard, Matteo Migliorini, Chiara Caborni, Francesco Onorati, Giulia Regalia, Daniel Friedman, and Orrin Devinsky. Wrist sensor reveals sympathetic hyperactivity and hypoventilation before probable SUDEP. *Neurology*, 89(6):633–635, 2017.

[43] M-Z Poh, T Loddenkemper, C Reinsberger, NC Swenson, S Goyal, JR Madsen, and Rosalind W Picard. Autonomic changes with seizures correlate with postictal eeg suppression. *Neurology*, 78(23):1868–1876, 2012.

[44] Mark Quigg. Circadian rhythms: interactions with seizures and epilepsy. *Epilepsy research*, 42(1):43–55, 2000.

[45] Vikram R Rao, Marc G Leguia, Thomas K Tcheng, and Maxime O Baud. Cues for seizure timing. *Epilepsia*, 62:S15–S31, 2021.

[46] Giulia Regalia, Giulia Gerboni, Matteo Migliorini, Matteo Lai, Jonathan Pham, Nirajan Puri, Milena K Pavlova, Rosalind W Picard, Rani A Sarkis, and Francesco Onorati. Sleep assessment by means of a wrist actigraphy-based algorithm: agreement with polysomnography in an ambulatory study on older adults. *Chronobiology International*, 38(3):400–414, 2021.

[47] Roland Renzel, Christian R Baumann, and Rositsa Poryazova. EEG after sleep deprivation is a sensitive tool in the first diagnosis of idiopathic generalized but not focal epilepsy. *Clinical Neurophysiology*, 127(1):209–213, 2016.

[48] Kyle C Rossi, Jalyoung Joe, Monica Makhija, and Daniel M Goldenholz. Insufficient sleep, electroencephalogram activation, and seizure risk: re-evaluating the evidence. *Annals of Neurology*, 87(6):798–806, 2020.

[49] Philippe Ryvlin, Lina Nashef, Samden D Lhatoo, Lisa M Bateman, Jonathan Bird, Andrew Bleasel, Paul Boon, Arielle Crespel, Barbara A Dworetzky, Hans Høgenhaven, et al. Incidence and mechanisms of cardiorespiratory arrests in epilepsy monitoring units (mortemus): a retrospective study. *The Lancet Neurology*, 12(10):966–977, 2013.

[50] Akane Sano, Rosalind W Picard, and Robert Stickgold. Quantitative analysis of wrist electrodermal activity during sleep. *International Journal of Psychophysiology*, 94(3):382–389, 2014.

[51] Julienne Aldeano Sarabia, Marıa Angeles Rol, P Mendiola, and Juan Antonio Madrid. Circadian rhythm of wrist temperature in normal-living subjects: A candidate of new index of the circadian system. *Physiology & behavior*, 95(4):570–580, 2008.

[52] Skipper Seabold and Josef Perktold. statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*, 2010.

[53] Rachel E Stirling, Cindy M Hidajat, David B Grayden, Wendyl J D'Souza, Katrina L Dell, Ray Boston, Logan D Schneider, Ewan Nurse, Dean Freestone, Mark J Cook, et al. Sleep and seizure risk in epilepsy: Bed and wake times are more important than sleep duration. *medRxiv*, pages 2022–08, 2022.

[54] Sara Taylor, Natasha Jaques, Weixuan Chen, Szymon Fedor, Akane Sano, and Rosalind Picard. Automatic identification of artifacts in electrodermal activity data. In *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1934–1937. IEEE, 2015.

[55] Joseph J Tharayil, Sharon Chiang, Robert Moss, John M Stern, William H Theodore, and Daniel M Goldenholz. A big data approach to the development of mixed-effects models for seizure count data. *Epilepsia*, 58(5):835–844, 2017.

[56] Christopher Torrence and Gilbert P Compo. A practical guide to wavelet analysis. *Bulletin of the American Meteorological society*, 79(1):61–78, 1998.

[57] Jolien S Van Campen, E Lorraine Hompe, Floor E Jansen, Demetrios N Velis, Willem M Otte, Fia Van De Berg, Kees PJ Braun, Gerhard H Visser, Josemir W Sander, Marian Joels, et al. Cortisol fluctuations relate to interictal epileptiform discharges in stress sensitive epilepsy. *Brain*, 139(6):1673–1679, 2016.

[58] Pauli Virtanen, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C J Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods*, 17:261–272, 2020.

[59] Melodie R Winawer, Jerry Shih, Erin S Beck, Jessica E Hunter, Michael P Epstein, and EPGP Investigators. Genetic effects on sleep/wake variation of seizures. *Epilepsia*, 57(4):557–565, 2016.