Neural Fragility as an EEG Marker of the Seizure Onset Zone

Adam Li[1,9], Chester Huynh[1,9], Zachary Fitzgerald[2], Iahn Cajigas[3], Damian Brusko[3], Jonathan Jagid[3], Angel Claudio[4], Andres Kanner[4], Jennifer Hopp[5], Stephanie Chen[5], Jennifer Haagensen[5], Emily Johnson[6], William Anderson[6], Nathan Crone[6,8], Sara Inati[7], Kareem Zaghloul[7], Juan Bulacio[2], Jorge Gonzalez-Martinez[8,*], and Sridevi V. Sarma[1,9,*]

[1]Department of Biomedical Engineering, Johns Hopkins University, Baltimore, United States
[2]Epilepsy Center, Cleveland Clinic, Cleveland, United States
[3]Department of Neurosurgery, University of Miami Miller School of Medicine, Miami, United States
[4]Department of Neurology, University of Miami Miller School of Medicine, Miami, United States
[5]Neurology, University of Maryland Medical Center, Baltimore, United States
[6]Neurology, Johns Hopkins Hospital, Baltimore, United States
[7]Surgical Neurology Branch, NINDS, NIH, Bethesda MD
[8]Neurosurgery, and Epilepsy Center, University of Pittsburgh Medical Center, Pittsburgh, United States
[9]Institute for Computational Medicine, Johns Hopkins University, Baltimore, United States
[*]Co-senior authors

November 14, 2020

One Sentence Summary: Neural fragility, an intracranial EEG biomarker for the seizure onset zone in drug-resistant epilepsy, predicts surgical outcomes with high accuracy.

Abstract

Over 15 million epilepsy patients worldwide do not respond to drugs. Successful surgical treatment requires complete removal, or disconnection of the seizure onset zone (SOZ), brain region(s) where seizures originate. Unfortunately, surgical success rates vary between 30%-70% because no clinically validated biological marker of the SOZ exists. We develop and retrospectively validate a new EEG marker - *neural fragility.* We validate this new marker in a retrospective analysis of 91 patients by using neural fragility of the annotated

[1]Also reachable at adam2392 at gmail dot com. Also reachable at the NCSL website: https://sarmalab.icm.jhu.edu/. Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD, 21218

SOZ as a metric to predict surgical outcomes. Fragility predicts 43/47 surgical failures with an overall prediction accuracy of 76%, compared to the accuracy of clinicians being 48% (successful outcomes). In failed outcomes, we identify fragile regions that were untreated. When compared to 20 EEG features proposed as SOZ markers, fragility outperformed in predictive power and interpretability suggesting neural fragility as an EEG fingerprint of the *SOZ*.

1 Introduction

Over 15 million epilepsy patients worldwide and 1 million in the US suffer from drug-resistant epilepsy (DRE) [1, 2, 3, 4, 5]. DRE is defined as continued seizures despite adequate trials of two tolerated appropriately chosen anti-epileptic drugs [6]. DRE patients have an increased risk of sudden death and are frequently hospitalized, burdened by epilepsy-related disabilities, and the cost of their care is a significant contributor to the $16 billion dollars spent annually in the US treating epilepsy patients [7, 8]. Approximately 50% of DRE patients have focal DRE, where a specific brain region or regions, termed the epileptogenic zone (EZ), is necessary and sufficient for initiating seizures and whose removal (or disconnection) is necessary for complete abolition of seizures [9, 10, 11, 12]. The EZ encompasses the clinically identified seizure onset zone (SOZ) and early propagation zone (EPZ). The brain regions associated with the *SOZ* demonstrate the earliest electrophysiological changes during a seizure event, and in general precede the clinical onset of seizures; and the EPZ regions are involved at the time of the earliest clinical (semiological) manifestations during a seizure event. Successful surgical and neuromodulatory treatments can stop seizures altogether or allow them to be controlled with medications [13, 14, 15], but outcomes for both treatments critically depend on accurate localization of the *SOZ*.

Localizing the *SOZ* also relies on accurate placement of electrodes such that they cover the EZ, and the ability to identify abnormalities in the iEEG channels that may correlate to the *SOZ* with the naked eye. Unfortunately, even the most experienced clinicians are challenged because epilepsy is fundamentally a network disease, which cannot be entirely defined by the current methods of localization. Abnormal connections across several channels may constitute a more effective marker of the *SOZ* [16]. Localization thus lends itself to a data-driven network-based computational approach and several EEG algorithms have been proposed to localize the *SOZ* from recordings. Many entail investigations of the spectral power in each iEEG channel including high frequency oscillations [17, 18, 19, 20, 21, 22, 23], but these approaches do not consider network properties of the brain because they treat each EEG channel independently. Others have proposed graph-based analysis of iEEG [24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35], but these approaches fail to identify internal network properties that cause seizures to occur in the first place.

In this study, we propose a new EEG marker of the *SOZ*, neural fragility, of brain regions (conceptually described in Figure 1). To create our fragility marker, we first build a personalized dynamic model of the brain network from observed iEEG signals. The model is generative in that it can accurately reconstruct the patient's iEEG recordings [36, 37]. Using this model, we then calculate which network nodes are imbalanced, meaning they have more excitatory or less inhibitory influence on the network and thus can trigger seizures. Neural fragility measures the degree to which a node is imbalanced [38]. To evaluate neural fragility as a marker of the SOZ, we conducted a retrospective study using iEEG data from 91 patients treated across 5 epilepsy centers: Johns Hopkins Hospital, National Institute of Health, Cleveland Clinic, University of Maryland Medical Center and Jackson Memorial Hospital, University of Miami. In the study population, all DRE patients underwent invasive iEEG monitoring followed by surgical resection or laser ablation of the *SOZ* (44 success and 47 failure outcome). We demonstrate that neural

fragility is higher/lower in electrode contacts within clinically annotated SOZs for success/failure patients. In addition, we compare fragility of iEEG nodes to 6 frequency-based and 14 graph theoretic features in a 10-fold nested-cross validation. Neural fragility has an area under the curve (AUC) discrimination score of 0.88 +/- 0.064, which is 13% better compared to the next best feature. In addition, it has a high degree of interpretability, which we demonstrate by computing an interpretability ratio suggesting that the spatiotemporal heatmaps of neural fragility are a robust iEEG fingerprint of the SOZ that can incorporate seamlessly into the clinical workflow.

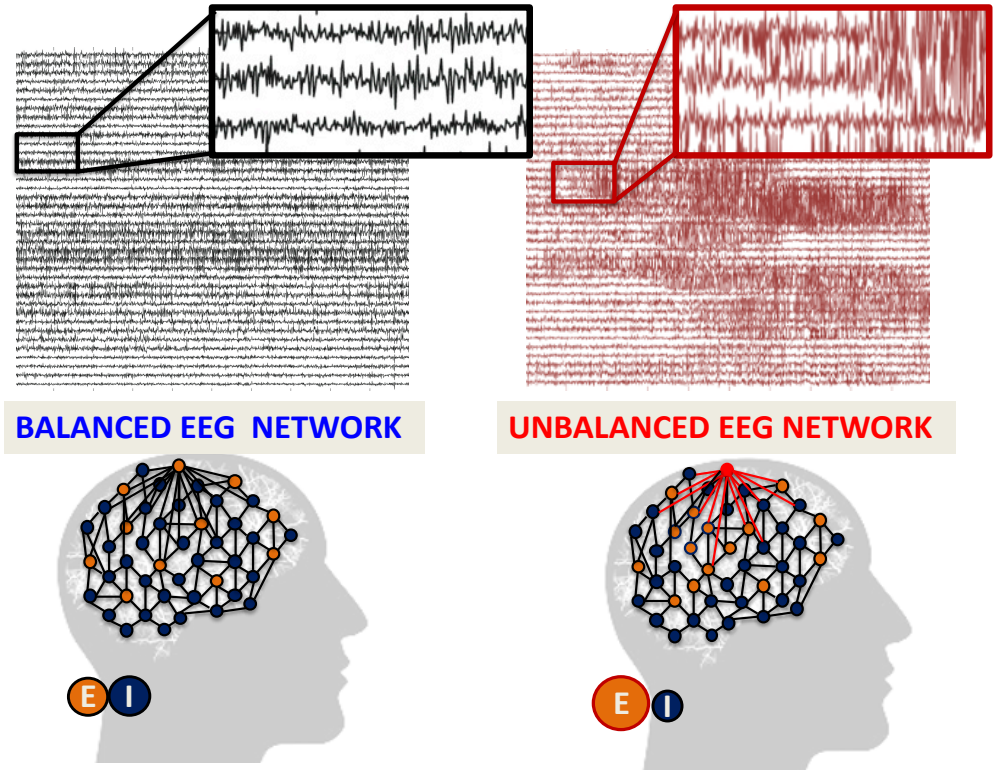


Figure 1: Neural fragility intuition (Top) iEEG traces in between seizures (left) and during a seizure (right). **(Bottom)** network schematic showing change in connectivity (right) in fragile node that causes seizure. Describes qualitatively the concept of neural fragility in the context of a dynamical iEEG network, with nodes representing excitatory (E) and inhibitory (I) population of neurons. From a dynamical systems point of view, such imbalance arises from a few fragile nodes causing instability of the network in the form of over-excitation, or under-inhibition. We define fragility of a network node to be the minimum perturbation applied to the node's functional connectivity to its neighbors before rendering the network unstable [38, 36]. In system theory, stable systems return to a baseline condition when a node is perturbed. In contrast, unstable systems can oscillate and grow when a node is perturbed. In the context of epilepsy, a fragile node is one that requires a smaller perturbation to lead to seizure activity.

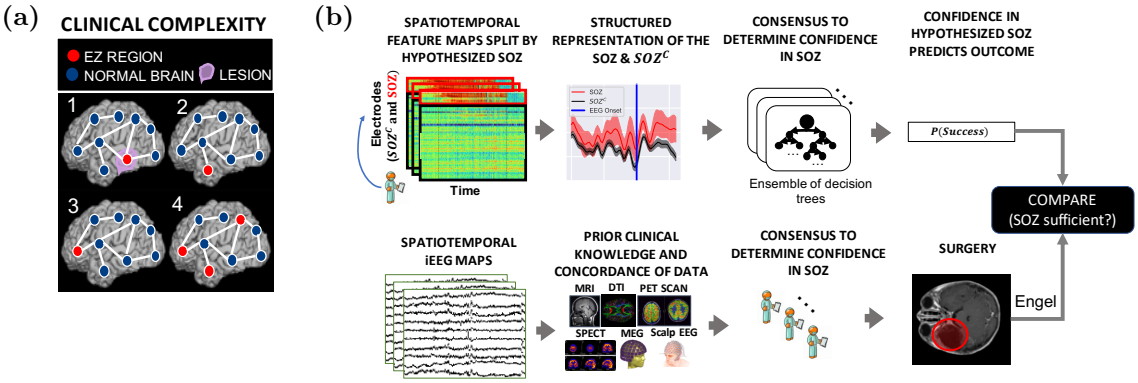


Figure 2: Clinical complexity and our experimental paradigm (a) Schematic of the difficulty of different epilepsy etiologies that might arise in DRE patients. Since there is no biomarker for the EZ and it is never observed directly, the network mechanisms that cause seizures are complex. Case clinical complexity ordered by increasing localization difficulty: lesional (1), focal temporal (2), focal extratemporal (3), and multi-focal (4) that are present in the dataset. These four categories simplify the possible EZ presentations, but provide a broad categorization of simple to complex cases observed in the clinic. **(b)** Is a schematic of our experimental design starting with iEEG data. **(top row)** We evaluate various representations of iEEG in the form of spatiotemporal heatmaps, create a partitioned summary of the clinically annotated SOZ around seizure onset, feed them into a Random Forest classifier and computed a probability of success (i.e. a confidence score) in the clinically hypothesized SOZ. The probability was then compared with the actual outcome of patients. **(bottom row)** Shows a simplified analogous workflow that clinicians take to evaluate their confidence in a proposed SOZ localization resulting in a surgery. During invasive monitoring, clinicians identify the SOZ from iEEG patterns (e.g. spiking/rhythmic activity). An estimate of the true EZ is then created using a combination of the clinical SOZ hypothesis and non-invasive findings. When possible, subsequent surgical resection or laser ablation, generally including the SOZ along with a variable extent of additional tissue, is performed. Post-operatively, patients are followed for 12+ months and categorized as either success, or failure as defined in Section 6.2. This translates to clinical outcomes measured by Engel or ILAE score. In order for a feature to be an accurate representation of the underlying epileptic phenomena, the following assumptions are made. As a result of seizure freedom, one can assume that the clinically hypothesized SOZ was sufficient, and the probability of success has a high value. In contrast, if seizures continue, then the SOZ was not sufficient and the probability should have a low value.

2 Results

There is no clinically validated biomarker of the SOZ, yet it forms an important component for the location of the underlying EZ [39, 40, 41, 42, 43, 44]. This presents a serious challenge for clinicians to accurately localize the SOZ and has led to surgical success rates to vary between 30-70% despite large brain regions being removed. With no biomarker available, clinicians are driven to perform extensive evaluations with neuroimaging, clinical testing, and visual inspection of scalp electroencephalography (EEG) recordings (Figure 2). When non-invasive tests are inconclusive, patients undergo intracranial monitoring, during which intracranial EEG (iEEG) electrodes are either placed directly on the cortex or implanted into the brain (example in Supplemental Figure 7). iEEG provides high temporal resolution data that enables clinicians to visually detect abnormal activity, such as spikes and high frequency bursts, in between seizures (interictal) and during seizures (ictal). Specifically, clinicians attempt to identify the electrodes involved in the SOZ and early spread [45, 46, 47, 11, 12]. Surgical resection is then performed on the basis of this hypothesis, unless the region overlaps with eloquent cortex [48].

We analyze all patient's iEEG using fragility and 20 other baseline features, resulting in spatiotemporal heatmaps per feature for every patient. The baseline features include spectral power in various frequency bands (e.g. delta band 1-4 Hz) and specific graph measures of bivariate correlation measures (e.g. eigenvector centrality and degree of correlation and coherence matrices), which have been previously reported in the literature to correlate to the SOZ [27, 26, 28, 29, 30, 25, 31, 35, 49]. We consider all of these as potential EEG features that can represent the data in different ways, potentially useful for SOZ localization.

2.1 Description of Neural Fragility

Neural fragility is a paradigm shift in the EEG analytics space. It is a concept based on the conjecture that focal seizures arise from a few fragile nodes, i.e., the SOZ, which renders the cortical epileptic network on the brink of instability. We begin with an intuitive explanation of neural fragility in Figure 1 (for a two-node quantitative example, see Supplemental Figure 11). When one observes iEEG data during interictal, or preictal periods (Figure 1, top left), activity recorded from each channel is noisy and hovers around a baseline value. In contrast, when one observes iEEG data during a seizure event (Figure 1, top right), activity (i) grows in amplitude, (ii) oscillates, and (iii) spreads in the brain. From a dynamical systems perspective, the iEEG network has switched from a stable (non-seizure) to an unstable (seizure) network. The only difference between the left and right iEEG network in Figure 1 is the connection strengths representing the dynamical interactions between a few channels, i.e., the SOZ. Our conjecture is that small changes in connection strengths at SOZ nodes cause an imbalance in connectivity between inhibitory (I) and excitatory (E) populations (nodes) in the region. Either inhibition is decreased and/or excitation is increased; thus, if the SOZ is perturbed then over excitation can occur manifesting in a seizure.

2.2 Fragility heatmap highlights the clinical SOZ in successful patients

To qualitatively assess the usefulness of fragility in localizing electrodes of interest, we first look at specific examples of patients with varying outcomes and epilepsy etiologies (see Figure 2). We analyzed with fragility and demonstrate how it may provide additional information for SOZ localization. In Figure 3, we show an example of three different patients with differing surgical treatments, outcomes, Engel class and clinical complexity along with their fragility heatmaps and corresponding raw iEEG data (for full clinical definitions; see Supplementary Excel table).

In Figure 3a, the red electrode labels on the y-axis correspond to the clinical SOZ electrodes; note that the red electrodes are typically a subset of the resected region. This figure shows the period 10 seconds before and after electrographic seizure onset, indicated by the black dashed line. In Patient_1, the clinical SOZ shows a high degree of fragility, even before seizure onset, which is not visibly clear in the raw EEG. The heatmap for Patient_1 also captures the propagation of seizure activity (Supplementary Figure 15). This patient had a successful surgery, and so we can assume the epileptogenic tissue within the clinical resection likely contained the SOZ and EPZ regions; it is likely the clinicians correctly localized the EZ. When viewing the raw EEG data in Figure 3b (top), Patient_1 has iEEG signatures that are readily visible around seizure EEG onset (Figure 3b). We see high-frequency and synchronized spiking activity at onset that occurs in electrodes that clinicians annotated as SOZ, which correspond to the most fragile electrodes at onset. In addition, the fragility heatmap captures the onset in the **ATT** and **AD** electrodes and early spread of the seizure into the **PD** electrodes. Specifically, **ATT1 (anterior temporal lobe area)** shows high fragility during the entire period before seizure onset (Figure 3). This area was not identified with scalp EEG, or non-invasive neuroimaging. In this patient, the fragility heatmap agreed with clinical visual EEG analysis, identifying the SOZ, which was included in the surgery and led to a seizure free patient.

2.3 Fragility heatmap highlights other possibly epileptic regions outside the SOZ in failed patients

In Figure 3, Patient_26 and Patient_40 both show distinct regions with high fragility that were *not* in the clinically annotated SOZ (or the resected region), and both had recurrent seizures after surgery. Specifically in Patient_26, the **ABT (anterior basal temporal lobe)**, **PBT (posterior basal temporal lobe)** and **RTG29-39 (mesial temporal lobe)** electrodes were highly fragile, but not annotated as SOZ. Patient_40 had laser ablation performed on the electrode region associated with **Q2**, which was not fragile. From seizure onset, many electrodes exhibit the EEG signatures that are clinically relevant, such as spiking and fast-wave activity [46]. In this patient, the **X' (posterior-cingulate)**, **U' (posterior-insula)** and **N'/M'/F' (superior frontal gyrus)** were all fragile compared to the **Q2** electrode, which recorded from a lesion in the right periventricular nodule. Patient_26 had a resection performed in the right anterior temporal lobe region. Clinicians identified the **RAD**, **RHD** and **RTG40/48** electrodes as the SOZ. In the raw EEG data, one can see synchronized spikes and spike-waves in these electrodes, but the patient had seizures continue after resection. In the corresponding fragility heatmap, the **ABT** and the **RTG29-32** electrodes are highly fragile compared to the clinically annotated SOZ region. In the raw EEG shown in Figure 3b, it is not visibly clear that these electrodes would be part of the SOZ. Visual analysis of the EEG was insufficient for Patient_26 and Patient_40, which ultimately led to insufficient localizations and then to failed surgical outcomes. Based on the fragility heatmaps, the fragile regions could be hypothesized to be part of the SOZ, and possibly candidates for resection.

2.4 Fragility outperforms all iEEG features in predicting surgical outcomes

To test the validity of neural fragility and the baseline features as SOZ markers, we investigate each feature's ability to predict surgical outcomes of patients when stratified by the set of SOZ contacts and the rest which we denote as the SOZ complement, SOZ^C. To do so, we trained a Structured Random Forest model (RF) [50, 51, 52] for each feature that takes in a partitioned spatiotemporal heatmap based on the clinically annotated SOZ and generates a probability of

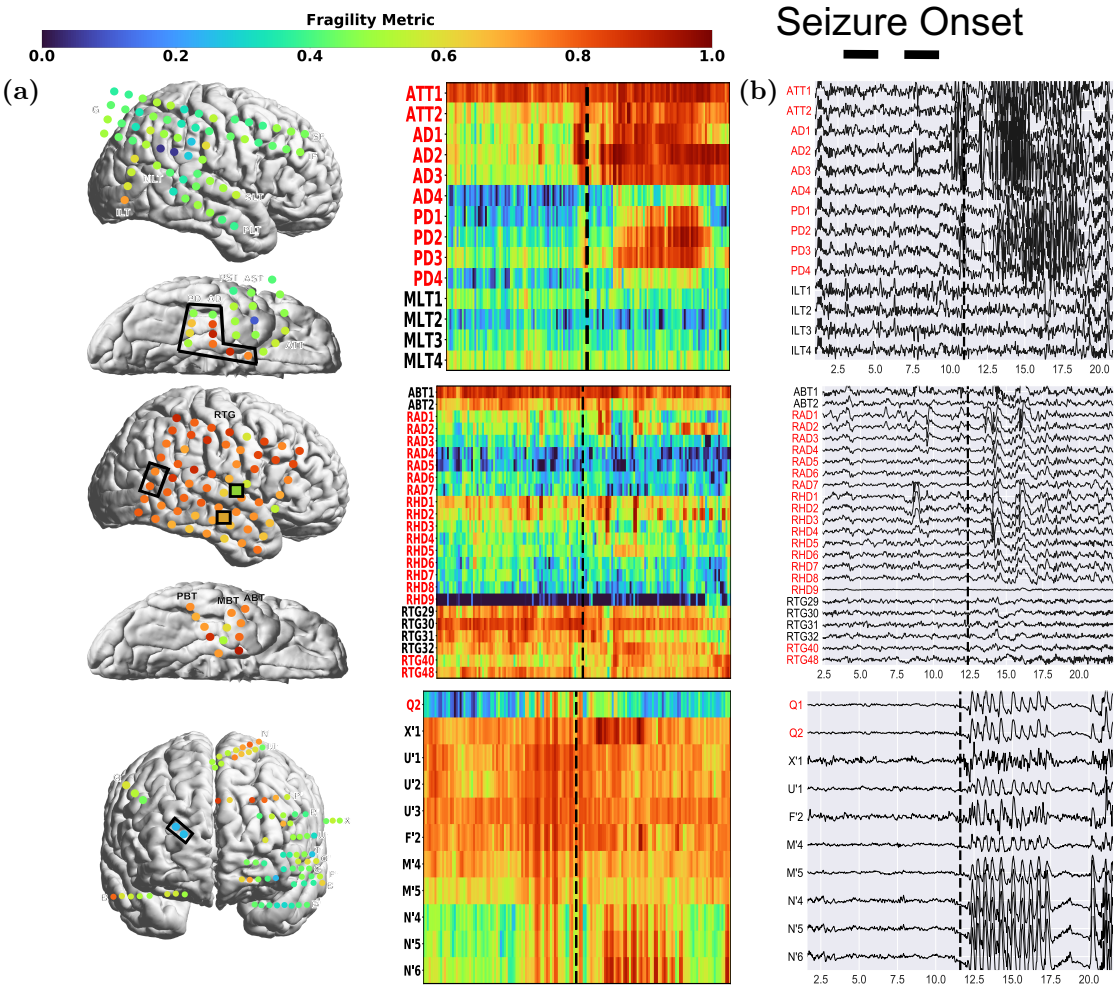


Figure 3: Fragility heatmaps, and corresponding raw EEG traces of successful and failed outcome patients. (a) From top to bottom, Patient_1 (success, NIH treated, CC1, Engel score 1), Patient_26 (failure, JHH treated, CC3, Engel score 4), and Patient_40 (failure, CClinic treated, CC4, Engel score 3) are shown respectively. The color scale represents the amplitude of the normalized fragility metric, with closer to 1 denoting fragile regions and closer to 0 denoting relatively stable regions. **(Left)** Overlaid average neural fragility value of each electrode in the window of analysis we used. Black dark squares represent a depth electrode that is not shown easily on the brain. Black lines outline where the clinicians labeled *SOZ*. Note in Patient_26, RAD and RHD electrodes are denoted by the squares with the color showing the average over the entire electrode. **(Right)** Clinically annotated heatmaps of the implanted ECoG/SEEG electrodes with red y-axis denoting SOZ contacts. The red contacts are also part of the surgical resection in these patients. Data is shown in the turbo colormap. Best seen if viewed in color. **(b)** Corresponding raw EEG data for each patient with electrodes on y-axis and time on x-axis with the dashed white-lines denoting seizure onset. Each shows 10 seconds before seizure onset marked by epileptologists, and 10 seconds after. EEG was set at a monopolar reference with line noise filtered out. Each EEG snapshot is shown at a fixed scale for that specific snapshot that was best for visualization, ranging from 200 uV to 2000 uV.

success - a confidence score in the clinical hypothesis. We test each feature's model on a held out data set by applying a varying threshold to the model's output and computing a receiver operating characteristic (ROC) curve. The ROC curve plots true positive rates versus false positive rates and the area under the curve (AUC) is a measure of predictive power of the feature. The larger the AUC, the more predictive the feature is and thus the more valid it is as an iEEG marker of the SOZ. We also compute the precision (PR) or positive predictive value (PPV) for each feature, which is the proportion of predicted successful or "positive" results that are true positives (actual successful surgeries). In addition, we compute the negative predictive value (NPV) as well. The larger the PPV/NPV, the more predictive the feature is and thus the more valid it is as an iEEG marker of the SOZ.

Since each patient's implantation has a varying number of electrodes, we summarize each feature's distribution of the SOZ and SOZ^C electrodes into quantile statistics over time (for details, see Methods 4). As an input to the RF models, we compute the quantile statistics over time of the SOZ and SOZ^C set for each patient and each feature. This allows each patient's heatmaps to have the same dimensions, which can be input into a machine learning RF model. RF models are attractive because they are non-parametric, are interpretable (they are just a set of decision trees performing a consensus procedure), and are able to handle higher dimensional data better compared to other models, such as Logistic Regression [53]. As an output of the trained RF model, each heatmap gets mapped to a probability that the outcome will be a success. An RF model is tuned for each feature through 10-fold cross-validation (CV; see Methods section for full details), resulting in a uniform and rigorous benchmark against neural fragility on the same set of patients. Note that the "high-gamma" frequency band feature encompasses what some would consider HFOs (i.e. 90-300 Hz) activity.

In terms of AUC (measures discrimination), neural fragility performs the best (Figure 4a). Compared to the top 3 leading baseline features, neural fragility outperforms the beta band (15-30 Hz) power by more than 13%. The AUC of fragility is the highest with a value of 0.88 +/- 0.064, compared to the next best representation, the beta frequency band, with a value of 0.82 +/- 0.040. From the ROC (and PR) curves, we observe that the fragility consistently has a higher sensitivity for the same false positive rate compared to the 20 other feature representations (Supplemental Figure 17). In terms of effect size, neural fragility improves over the beta frequency band with a large [54] effect size of 0.97 Cohen's D with 9 out of 10 folds improving the AUC (Paired Wilcoxon Rank-sum PValue = 0.027 (see Supplemental Figure 17). As a result of the 10-fold CV, we can compare the AUC on the same set of patients in a paired effect size and statistical test improving AUC on 9 out of the 10 CV folds.

In terms of PR (measures PPV), neural fragility also performs the best (Figure 4b). IN terms of average precision, which weighs the predictive power in terms of the total number of patients, neural fragility also obtains an average precision of 0.83 +/- 0.076, which is >5% better than the next best feature. In addition, compared to the clinical surgical success rate of 47%, fragility had a 76% +/- 6% accuracy in predicting surgical outcome, a PPV of 0.903 +/- 0.103 and a NPV of 0.872 +/- 0.136. In terms of PR, neural fragility improves over the beta band power with a medium [54] effect size of 0.57 Cohen's D with 8 out of 10 folds improving the PR (PValue = 0.08).

When we compare the differences in the success probabilities predicted by the models stratified by surgical outcomes, we observe that neural fragility has the highest effect size of 1.507 (1.233-1.763 - 95% confidence interval) (Figure 4c) and lowest pvalue of 6.748e-31 (Figure 4d). In addition to having good discrimination, we compute how well-calibrated the model is, having the success probabilities values reflect true risk strata in the patient population. We quantify how well calibrated the success probability distributions are over the held-out test set. In Supplementary Figure 16, we show that the RF model trained on neural fragility produces well-calibrated success

probability values. A perfectly calibrated model would have a 20% confidence in a set of patients with exactly 20% success outcomes. The Brier-loss (a measure of calibration; 0 being perfect) was 0.162 +/- 0.036, which was a 6.7% improvement compared to the next best feature.

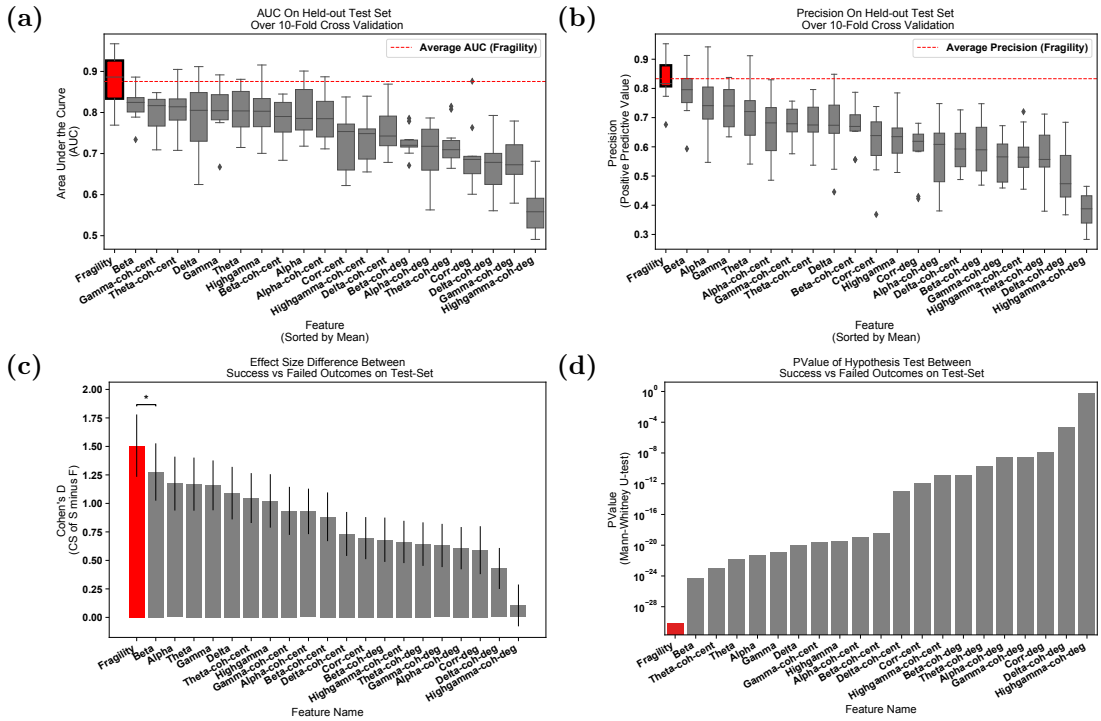


Figure 4: Area under the curve and average precision performance. Specific results for neural fragility are marked in red for each of the panels (a-d). **(a)** Discrimination plot (measured with AUC) shows the relative performance of benchmark feature representations compared to that achieved with neural fragility. **(b)** A similar average-PR curve shows the relative positive predictive value of all features compared with fragility. Average precision is the analagous area under the curve for the PR curve. **(c)** A summary of the Cohen's D effect size measurements between the success and failed outcome CS distributions across all features. The effect size of neural fragility is significantly greater then that of the beta band (alpha = 0.05) **(d)** The corresponding PValues of the effect size differences between success and failed outcomes, computed via the Mann-Whitney U-test.

2.5 Fragility correlates with clinical complexity and treatment outcomes

Neural fragility correlates with treatment outcomes and clinical complexity. Successful outcomes are defined as seizure free at 12+ months post-op (Engel class I and ILAE scores of 1 and 2) and failure outcomes are defined as seizure recurrence (Engel classes 2-4) at 12+ months post-op. In addition, we can categorize patients by their clinical complexity (CC) as follows: (1) lesional, (2) focal temporal, (3) focal extratemporal, and (4) multi-focal (Figure 2) [46, 47]. We stratify the distribution of the success probabilities based on Engel score and show a decreasing trend in success probability as Engel score increases (Figure 5b). The effect sizes when comparing

against Engel class I were 1.067 for Engel class II (P = 4.438e-50), 1.387 for Engel class III (P = 2.148e-69), and 1.800 for Engel class IV (P = 4.476e-74). Although the AUC indirectly implies such a difference would exist between Engel score I (success) and Engel score II-IV (failures), it is reassuring to see that the level of confidence decreases as the severity of the failure increases; Engel score IV patients experience no changes in their seizures, suggesting the true epileptic tissue was not localized and resected, while Engel score II patients experience some changes, suggesting there was portions of the true epileptic tissue that may have been resected. We also compare the success probability distributions with respect to the ILAE score, which is another stratification of the surgical outcomes (Figure 5c). There is a similar trend as seen in the Engel score - as ILAE score increases, the success probability decreases.

We also analyze the success probability with respect to clinical measures of the epilepsy severity, such as the clinical complexity (CC) of the patient, which is a categorization of the etiology of the disease. CC is determined by what type of seizures the patient exhibits rather than the severity of the seizures after-surgery (for more details see Methods 6.2). CC is a factor that can be determined before surgery and one we expect would correlate with failure rate. This is because the higher the CC, the more difficult localization of the SOZ is and hence possibly the more seizure recurrences after surgery. In Figure 5a, we observed that CC1 and CC2 (i.e., lesional and temporal lobe) patients had a similar distribution of success probabilities (Cohen's D = 0.020; MW U-test PValue = 0.799; LQRT PValue = 0.76), while CC3 and CC4 had significantly lower distributions (CC3: [Cohen's D = 0.779; MW U-test PValue = 2.808e-19; LQRT PValue = 0.00]; CC4: [Cohen's D = 1.138; MW U-test PValue = 7.388e-26; LQRT PValue = 0.00]). This trend is not optimized for directly in our discrimination task, but it aligns with clinical outcomes. CC1 and CC2 are comparable, which agrees with current data suggesting that lesional and temporal lobe epilepsy have the highest rates of surgical success [55, 56, 57, 58]. Extratemporal (CC4) and multi-focal (CC4) patients tend to have lower success rates due to insufficient localizations and thus the neural fragility confidence in those SOZ localizations should be low [55, 59, 60, 61]. Finally, we examine the success probability differences based on gender, handedness, onset age and surgery age to show that there are no relevant differences (Supplemental Figure 19).

2.6 Fragility heatmaps are more interpretable than all other EEG feature maps

We can parse the RF model to determine important aspects of the spatiotemporal heatmap that are important in predicting surgical outcome. To do so, we perform permutations on the heatmap inputs to the RF models to measure their relative importances over space and time. We observed that the AUC metric was affected primarily by a combination of the highest (90th quantiles and up) neural fragility values of both the SOZ and SOZ^C contacts (Figure 18). The SOZ^C neural fragility as much as 10 seconds before seizure onset impacted the AUC as did neural fragility of the SOZ right at seizure onset. For both the SOZ and the SOZ^C fragility distributions, the 80th quantile and below did not contribute to the predictive power. In different patients across different clinical centers, the practice of marking seizure onset times can vary due to varying methodologies [47]. As a result, in the SOZ^C, there was some variation in terms of which time point in the heatmap mattered most.

Although strong predictive capabilities of an EEG marker of the SOZ are necessary, it is also important that the marker be presented in an interpretable manner to clinicians. We next show how fragility heatmaps are the most interpretable over all baseline features. In Figure 6a, there are two heatmaps computed for the same seizure event in Patient_01: one is a beta band map (left) and one is a neural fragility map (right). Both maps are normalized across channels and both are computed with similar sliding windows (full details in Online Methods). However, it is

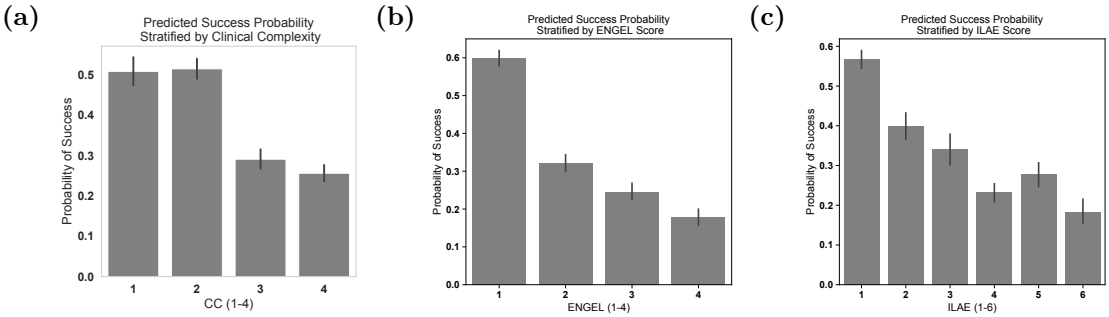


Figure 5: Neural fragility of patients stratified by clinical covariates. (a) Distribution of success probability values per patient stratified by clinical complexity (CC; see Methods 6.2), where lesional (1) and temporal lobe epilepsy (2) patients have similar distributions because they are generally the "easier" patients to treat, whereas extratemporal (3) and multi-focal (4) have lower general probabilities because they are "harder" patients to treat. **(b)** The distribution of the probability values per patient stratified by Engel score. Due to the AUC being high for fragility, it is expected that Engel I has high CS, while Engel II-IV have lower success probability. However, the relative downward trend in the success probabilities from Engel II-IV indicated that neural fragility is present in the clinical SOZ in varying degrees from Engel II-IV. This suggests that it correlates with the underlying severity of failed outcomes. **(c)** A similar distribution for another measure of surgical outcome, the ILAE score, where 1 are considered success and 2-6 are considered failure.

clear that less contacts "stand-out" as pathological in the beta-band heatmap before seizure onset, with the majority of the map being different shades of blue. In contrast, in the fragility heatmap, one contact (**ATT1** from Figure 3) is fragile the entire duration before seizure onset (solid white line), and then it is clear that a few more contacts become fragile after the electrographic onset of the seizure. These fragile areas that "pop-out" in the heat-map as red-hot areas occur in the clinically annotated SOZ and this patient had a successful surgery.

To quantify interpretability, we compute an interpretability ratio: the ratio of the feature values in the 90th quantile between the SOZ and the SOZ^C over the section of data used by the feature's RF model. This measures the contrast that one sees between the extreme values in the SOZ versus the extreme values they see in the SOZ^C. The larger the ratio, then the more contrast the map will have. In Figure 6b, we show that the effect size difference between successful and failed outcome of this interpretability ratio is largest in neural fragility when compared to all baseline features. It is well-known that RF models are scale-invariant [52], so it is plausible that there are portions of heatmaps that distinguish one channel from another and that can be parsed out via a decision tree and not the naked eye, which leads to high AUC in the beta band representation in Figure 4. For example a decision tree can discriminate between 0.30000 and 0.2995, which on a normalized color-scale is difficult to parse by visual inspection. Neural fragility on the other hand, shows marked differences between the clinically annotated SOZ (red electrodes on the y-axis) and the actual fragility values even before seizure onset.

3 Discussion

We demonstrate that neural fragility, a networked-dynamical systems based representation of iEEG is a strong candidate for a highly interpretable iEEG fingerprint of the SOZ. We compared

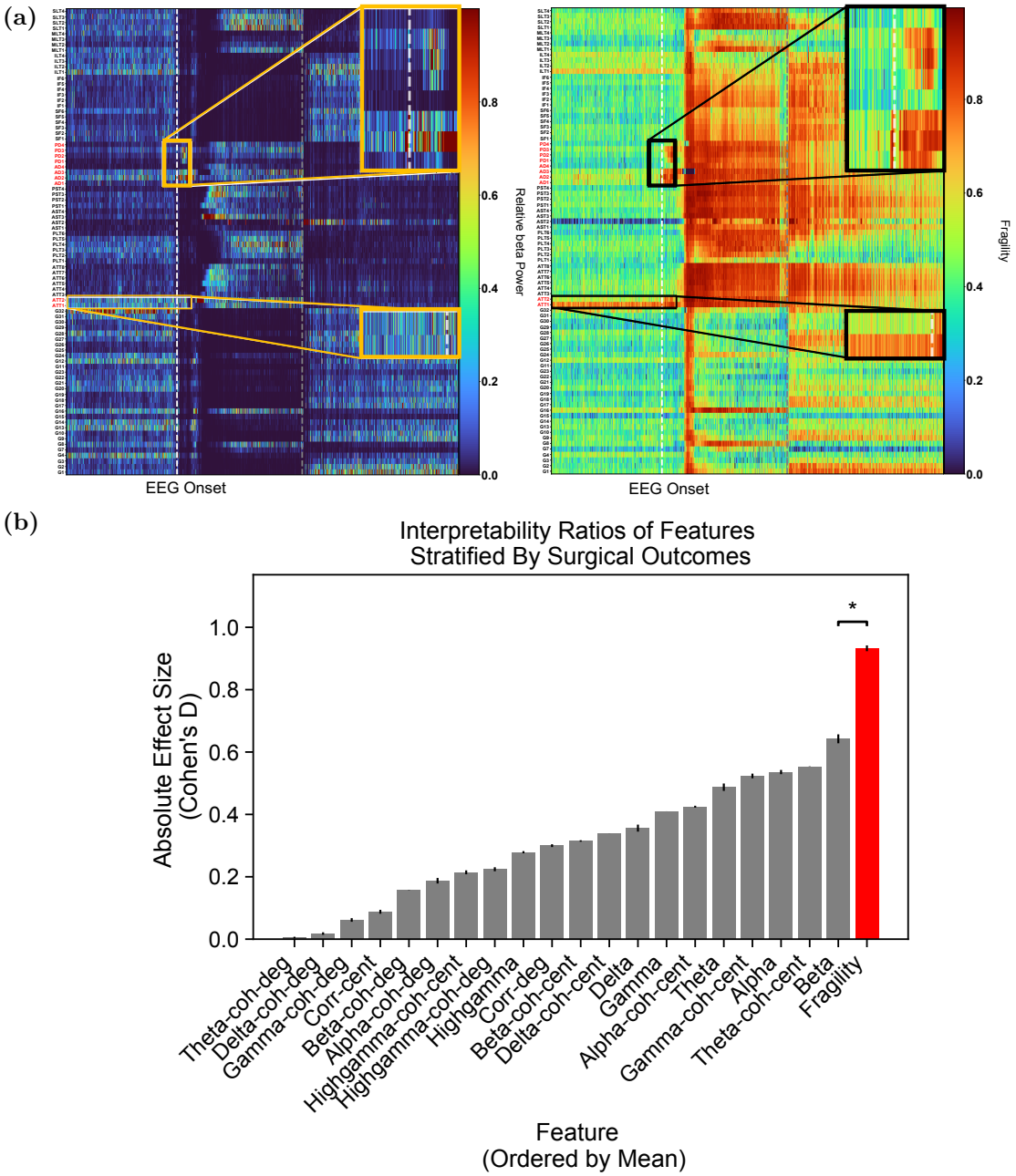


Figure 6: Interpretability ratio of feature heatmaps. (a) Two heatmap examples of a seizure snapshot of Patient_01 (Treated at NIH, ECoG, CC1, Engel I, ILAE 1) with the beta frequency band (left) and the neural fragility heatmap (right). Both colormaps show the relative feature value normalized across channels over time. The black line denotes electrographic seizure onset. **(b)** A bar plot of the interpretability ratio that is defined in Results Section 2.6 computed for every feature. The y-axis shows an effect size difference between the interpretability ratios of success and failed outcomes. The interpretability ratio for each patient's heatmap is defined as the ratio between the feature values in the two electrode sets ($\frac{SOZ}{SOZ^C}$). Neural fragility is significantly greater then the beta band (alpha level=0.05). The black bars denote 1 std.

neural fragility to 20 other popular features using data from 91 patients treated across five centers with varying clinical complexity. Neural fragility performed the best in terms of AUC, PR and interpretability. This is the first study to our knowledge that benchmarks proposed iEEG features in a uniform fashion over the same sets of patients using the best possible parameters of a RF model to compute a confidence measure (i.e. probability of success) of the clinically annotated SOZ. To facilitate reproducibility and further investigation of iEEG features, we made this dataset BIDS-compliant and publicly available through the OpenNeuro website (for details see Methods 6.1).

Challenges in validating iEEG features as SOZ markers Many features have been proposed as potential biomarkers for the SOZ, but none have successfully translated into the clinical workflow [27, 26, 28, 62, 63, 55, 31]. Current limitations for evaluating computational approaches to localization can be largely attributed to i) the lack of ground-truth labels for the true underlying SOZ (it cannot be observed in practice because there is no biomarker), ii) insufficient benchmarking to other iEEG features and iii) a lack in sufficient sampling across epilepsy etiologies.

Since there is no ground truth to drive algorithmic development, one can instead look for features that correlate with clinicians when stratified by outcome measures. Our approach sees if the feature values of the clinically annotated SOZ are "high" in success patients and "low" in failed patients. More rigorously we use feature values of SOZ and SOZ^C to predict surgical outcomes, which is a good approach since we lack ground truth labels of our desired variable. Note that developing algorithms to directly predict the SOZ will at best replicate what the current standard practice is, and achieve a rate of 30-70% surgical success rate [27, 26, 28, 62, 63]. In addition, electrodes within the SOZ may not be a part of the true EZ, but are annotated because of their "appearance" to be the onset of seizures. At best, it can be assumed that in successful surgical outcomes, the EZ is an unknown subset of the SOZ and resected zone. Hence one desires a feature that has relatively high confidence in the clinically annotated SOZ in success outcomes compared to failed outcomes.

Even with a seemingly successful feature derived from data, it is important to benchmark against existing approaches to provide a holistic view of the value of the said feature. Without benchmarking, it is easy to become overly optimistic in terms of the performance of a feature, whereas it may very well be that other iEEG features perform just as well. In this study, we benchmark neural fragility against 20 other proposed features. While other traditional features such as the power in the beta band seem to be informative in SOZ localization [64], neural fragility outperforms in effect size, p-value and interpretability.

Although predicting surgical outcomes in our experimental setup is promising, it will be important to understand why certain localizations are successful and why certain are insufficient. If neural fragility is a good marker, then we expect successful outcomes to have high fragility in their clinically annotated SOZ, and lower fragility in the SOZ^C, which is shown in this study. Understanding failed localizations and why they failed becomes more difficult. For example, in patients with lesions on MRI scans that correlate with the patient's EEG and seizure semiology, surgical resection can lead to seizure freedom in approximately 70% of patients [47]. Even in these relatively straightforward cases, localization is not perfect, possibly due to chronic effects of epilepsy such as kindling, which can cause neighboring tissue to become abnormal and epileptogenic [65, 66]. This is why there is a need to sample a heterogeneous and large patient population and derive a feature is invariant on average to epilepsy type and clinical covariates. In this study, we spent over four years to successfully collect and annotate this heterogeneous dataset of 91 patients.

Limitations of the most popular iEEG features A google scholar search using the keywords: "localization of seizure onset zone epilepsy intracranial EEG" produces over 24,000 results. This is striking considering no computational tools to assist in SOZ localization are in the clinical workflow today. The majority of proposed features lack consistency in finding objective iEEG quantities that correlate to clinically annotated SOZ because they fail to capture internal properties of the iEEG network which are critical to understand when looking for the SOZ. Proposed algorithms either (i) compute EEG features from individual channels (e.g. spectral power in a given frequency band), thus ignoring dependencies between channels [23, 67, 68, 20, 19, 22, 21] to name a few, or they (ii) apply network-based measures to capture pairwise dependencies in the EEG window of interest. Specifically, correlation or coherence between each pair of EEG channels is computed and organized into an adjacency matrix, on which summary statistics are derived including degree distribution and variants of centrality [26, 27, 28, 29, 30, 31, 32, 35, 34, 25, 33]. Such network-based measures are not based on well formulated hypotheses of the role of the epileptic tissue in the iEEG network, and many different networks (adjacency matrices) can have identical summary statistics [69] resulting in ambiguous interpretations of such measures.

A popular EEG feature that has been proposed as an iEEG marker of the SOZ and reported in over 1000 published studies is High-frequency oscillations (HFOs) ([70, 23, 71, 72, 73, 74, 75, 76, 71, 77, 74, 78, 79, 72, 80, 81, 82, 83, 62, 63, 84] to name a few). HFOs are spontaneous events occurring on individual EEG channels that distinctively stand out from the background signal and are divided into three categories: ripples (80–250 Hz), fast ripples (250–500 Hz), and very-fast ripples (>500 Hz) [85, 74, 86]. Regarding epilepsy, retrospective studies suggested that resected brain regions that generate high rates of HFOs may lead to good post-surgical outcome (e.g., [87, 88, 89, 90, 91, 92, 49, 93, 94, 95, 96, 97]). A 2015 reported meta-analysis investigated whether patients with high HFO-generating areas that had been resected presented a better post-surgical seizure outcome in comparison to patients in whom those areas had not been resected [98]. Although they found significant effects for resected areas that either presented a high number of ripples or fast ripples, effect sizes were small and only a few studies fulfilled their selection criteria [98]. Furthermore, several studies have also questioned the reproducibility and reliability of HFOs as a marker [99, 97, 100, 101, 72, 102] and note that there are also physiologic, non-epileptic HFOs. Their existence poses a challenge, as disentangling them from clinically relevant pathological HFOs still is an unsolved issue [103, 104, 105, 106, 107].

Similar inconclusive results hold in completed prospective studies of HFOs. In 2017, an updated Cochrane review by Gloss et al. [108] investigated the clinical value of HFOs regarding decision making in epilepsy surgery. They identified only two prospective studies at the time and concluded that there is not enough evidence so far to allow for any reliable conclusions regarding the clinical value of HFOs as a marker for the SOZ. Today, five clinical trials are listed as using HFOs for surgical planning on clinicaltrials.gov as either recruiting, enrolling by invitation, or active and not enrolling and none have reported results. The fundamental limitation of the aforementioned studies lies in the fact that they approach the SOZ EEG marker discovery process as a signal processing and pattern recognition problem, concentrating on processing EEG observations to find events of interest (e.g., HFOs) as opposed to understanding how the observations were generated in the first place and how specific internal network properties can trigger seizures.

Why Neural Fragility Performs Well Rather than analyzing iEEG data at the channel and signal processing level, we seek to model the underlying network dynamics that might give rise to seizures in the form of neural fragility in a dynamical network model. A notion of fragility in networks is commonly seen in analysis of structural [109], economic [110] and even social networks [111]. Although we are not directly analyzing the structural nature of neuronal network,

there are studies that have characterized epilepsy in terms of structural fragility and network organization [112, 38]. Specifically, in cellular studies [113, 114], epilepsy is caused by changes, or "perturbations" in the structural network (i.e. chandelier cell loss, or abnormal axonal sprouting from layer V pyramidal cells), which causes loss of inhibition or excessive excitation respectively; these biological changes cause downstream aberrant electrical firing (i.e. seizures). In this study, we analyze a functional network, characterized by a dynamical system derived from the iEEG recordings. Each electrode's effect on the rest of the network is captured by a time-varying linear model that we proposed in [37]. Each node is an electrode, which is recording aggregate neuronal activity within a brain region. By quantifying the fragility of each node, we determine how much of a change in that region's functional connections is necessary to cause seizure-like phenomena (e.g. instability). As a result, high neural fragility is hypothesized to coincide with a region that is sensitive to minute perturbations, causing unstable phenomena in the entire network (i.e. a seizure).

Presenting neural fragility as spatiotemporal heatmaps allows clinicians to qualitatively assess which electrodes and time points are most fragile within an iEEG network, aggregating any existing data sources (e.g. MRI, neuropsych evaluations, etc.) to formulate a localization hypothessis. By analyzing the fragility heatmaps of patients retrospectively, we conclude that i) fragility is high in electrodes present in the SOZ when the patient's surgery resulted in seizure freedom (i.e. Engel class I) and ii) high fragility is present in electrodes present outside the SOZ when the patient's surgery resulted in seizure recurrence (i.e. Engel class II-IV). In the context of fragility theory of a network, seizure recurrence can be due to perturbations of highly fragile regions in the epileptic network that were left untreated. Importantly, fragility of an electrode within a certain window does not correlate directly with gamma or high-gamma power, which are traditional frequency bands of interest for localizing the SOZ (Supplementary Figure 20) [28, 42, 21, 115, 49, 116, 80, 27]. This implies that neural fragility presents independent information on top of what clinicians look for in iEEG data. If translated into the clinic, neural fragility can serve as an additional source of information that clinicians can utilize for localization.

Scientific and Technological Advances Emerging from Neural Fragility Neural fragility has the potential to re-define how epilepsy surgery is performed, departing from the classical "localization paradigms" and "en-bloc resections" to a personalized "network-based" user-friendly visualization and surgical strategy. By developing a novel 3D (brain region, time, fragility) network-based method for anatomical representation of the epileptiform activity, including the seizure onset areas and the early propagation zone, this study will have high impact with the potential to offer a safer, more efficient, and cost-effective treatment option for a highly challenging group of patients with disabling DRE. More precise SOZ localization using neural fragility would also guide of chronic implantation of neurostimulation devices aimed to suppress seizures with bursts of current when detected [117, 118, 119, 120, 121, 122, 123, 124].

Neural fragility may also be relevant in detecting epileptogenic regions in the brain when applied to interictal (between seizures) iEEG recordings. Ictal or seizure iEEG data are currently the gold standard in clinical practice for localizing the SOZ [47, 46]. However, having patients with electrodes implanted for long periods of time, and requiring the monitoring of multiple seizure events over many weeks carries the risk of infection, sudden death, trauma and cognitive deficits from having repeated seizures. This contributes to the large cost of epilepsy monitoring [1, 2, 3, 4, 5, 15]. If a candidate iEEG marker could be found that is able to provide strong localizing evidence using only interictal data, then it would significantly reduce invasive monitoring time [112].

Neural fragility is an EEG marker that can also further advance our knowledge of neural mechanisms of seizure generation that will then drive more effective interventions. For example,

fragility can be used to identify pathological tissue that can be removed and tested in vitro for abnormal histopathological structure [113, 114]. Knowledge of structural abnormalities may inform new targeted drug treatments.

Finally, neural fragility may have broader implications in understanding how underlying brain network dynamics change during intervention (e.g. drugs or electrical stimulation). This could have large implications in assessing efficacy of drugs on all patients with epilepsy (i.e., not just patients with MRE), as well as multiple other neurological disorders including Alzheimher's disease or the spectrum of dementias.

4 Materials and Methods

4.1 Study Design

The aim of the presented study is to test the hypothesis that neural fragility is a robust feature representation of the SOZ in functional iEEG networks for DRE patients. To explore this aim we collected retrospective iEEG data from n=91 DRE patients who underwent intracranial EEG monitoring, which included either electrocorticography (ECoG), or depth electrodes with stereo-EEG (SEEG). Patients were selected from University of Maryland Medical Center (UMMC), University of Miami Jackson Memorial Hospital (UMH), National Institute of Health (NIH), Johns Hopkins Hospital (JHH), and the Cleveland Clinic (CClinic). Patients exhibiting the following criteria were excluded: no seizures recorded, pregnant patients, EEG sampling rate less than 250 Hz, previous surgeries more then 6 months before the current implantation, and no surgery was performed. Subjects were grouped on the basis of their epilepsy clinical complexity (1-4), surgical outcome (S/F), Engel score (I-IV) and ILAE score (1-6). The non-clinical investigators were blinded to data collection. Every patient's SOZ was labeled by 1-3 trained epileptologists (depending on the center). The corresponding SOZ complement that we define, or SOZ^C, are the electrodes that are not part of the SOZ. Surgery was performed on the basis of the clinically annotated SOZ and subjects were followed for 12+ months to determine treatment outcome. For full details of patients, see the Supplementary section.

4.2 Neural Fragility

All iEEG data is preprocessed in the same manner as in [36]. For full details on preprocessing, see the Supplementary details.

To compute fragility heatmaps from iEEG recordings, we constructed a simple linear model for each 250 ms iEEG window. We used a sparse least-squares with a 10e-5 l2-norm regularization to ensure that the model identified was stable as in [37, 36]. Then, we slid the window 125 ms and repeated the process, generating a sequence of linear network models in time. We systematically computed the minimum perturbation required for each electrode's connections (for details see Supplemental Figure 8b) to produce instability for the entire network as described in [36]. The electrodes that were the most fragile were hypothesized to be related to the SOZ in these epilepsy networks (seen as the dark-red color in the turbo colormap in Figure 3).

We also compute time-varying features that can be computed from small windows of iEEG data to serve as a benchmark against neural fragility. This includes spectral power in frequency bands, eigenvector centrality of correlation and coherence matrices and in-degree of correlation and coherence matrices, resulting in 20 benchmark features. For full details on computing the benchmark features, see Supplementary section.

4.3 Statistical Analyses

Results and graphs are shown as means +/- std. Statistical analysis was performed using Python and a suite of open-source libraries.

To determine the efficacy of neural fragility over other benchmark features, we propose a 10-fold cross-validation (CV) experiment on a fixed set of subjects with varying epilepsies. A structured RF statistical model was built using Python. We select a window of 10 seconds before marked seizure onset until the first 5% of the seizure. Each feature heatmap was structured into a set of quantile features of the SOZ and SOZ^C set over time resulting in a 20-dimensional vector over the considered window. Our driving hypothesis is that SOZ feature values should be different from the feature values of the SOZ^C in successfully treated patients. For each fold of

the CV experiment, patients are split into a training, validation and held-out test set. Parameters of the RF are trained over the training patients, and hyperparameters are optimized on the validation set. All performance metrics are reported on the held-out test set. The RF model outputs a predicted probability of success per patient. Area under the curve (AUC), and average precision are measured along with the predicted probability of success scores per patient.

Effect sizes are computed using Cohen's D. PValues are computed using the Mann-Whitney U tests for unpaired data, and paired Wilcoxon rank-sum tests for paired data [125]. We corrected for any multiple comparisons using the Holm-Bonferroni step-down method [126]. In some cases, we also present Welch's t-test and the Likelihood Q-ratio test (LQRT) results to show that the results are robust to statistical test chosen. For full details and motivations of each step in our statistical analyses, see Supplementary section.

5 Supplementary Material

Raw data is present in the link in Section 6.2. Any supplementary Information and Source Data files are available in the online version of the paper.

5.1 Dataset and definition elaborations

Clinical Center Patients						
Clinical Covariate	NIH	JHH	UMMC	UMH	CClinic	Total
Number Patients	14	4	7	5	61	91
Number Success	9	1	7	3	24	44
Number Failure	5	3	0	2	37	47
Number Male	6	n/a	7	1	30	44
Number Female	8	n/a	0	4	31	43

Table 1: A table of number of patients per clinical center (NIH = National Institute of Health; JHH = Johns Hopkins Hospital; UMMC = University of Maryland Medical Center; UMH = University of Miami Jackson Memorial Hospital; CClinic = Cleveland Clinic). Includes number of success, failures, and gender per group. Note that JHH did not retain gender information for these group of patients.

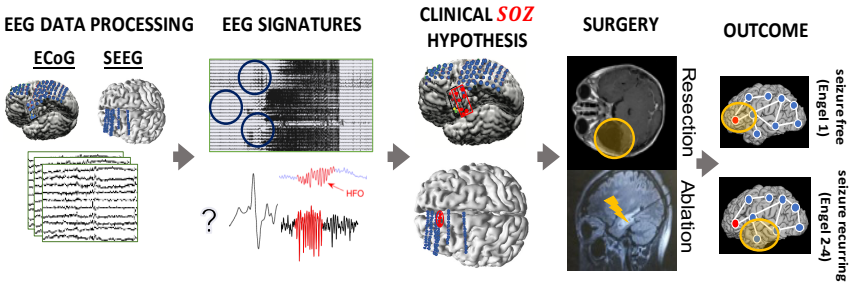


Figure 7: After non-invasive analysis is inconclusive, a complex and relatively lengthy iEEG workflow for localization takes place. During invasive monitoring, clinicians identify the clinically hypothesized SOZ as the electrodes with the earliest visually identifiable EEG signatures at seizure onset (see [47] for a review of this process). Note it is impossible to observe the EZ directly, so clinicians are required to estimate where to place iEEG electrodes and then successfully interpret the network patterns and raw iEEG patterns (e.g. HFOs, or spiking activity) to formulate the clinical SOZ hypothesis. An estimate of the true EZ is then created using a combination of the clinical SOZ hypothesis, as well non-invasive findings. When possible, subsequent surgical resection or laser ablation, generally including the SOZ along with a variable extent of additional tissue, is performed. Post-operatively, patients are followed for 12+ months and categorized as either success, or failure as defined in Section 6.2. This translates to clinical outcomes measured by Engel score, or ILAE score. As a result of seizure freedom, one can assume that the clinical SOZ included the EZ in some manner. In contrast, if seizures continue, then the clinical SOZ was not sufficient to cover the EZ.

5.2 Benchmark analysis

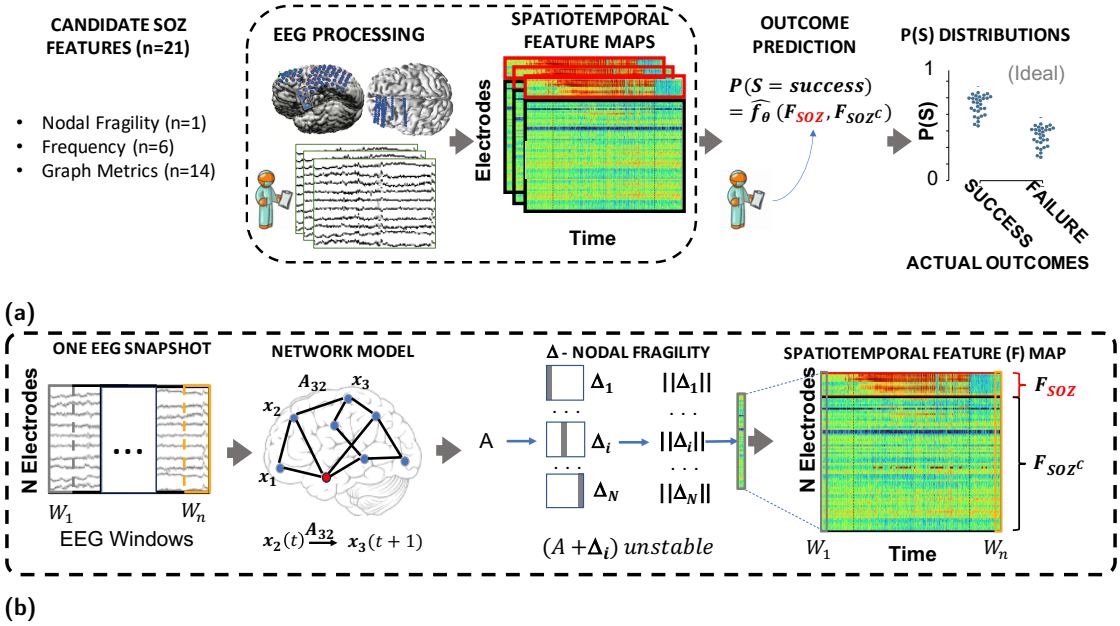


Figure 8: Computational experiment setup for all candidate SOZ features and statistical analysis. **(a)** Any candidate feature that can produce a spatiotemporal heatmap was computed from EEG data and then partitioned by the clinically annotated SOZ set and the complement, SOZ^C (i.e. non-SOZ electrodes) to compute a confidence statistic measuring the feature's belief of the clinician's hypothesis. Here F_{SOZ} and F_{SOZ^C} were the feature values within their respective sets. f_θ is the function depending on the Random Forest model parameters, θ that maps the statistics of the F_{SOZ} and F_{SOZ^C} to a confidence statistic. An ideal feature would have high and low confidence for success and failed outcomes respectively. Each point on the final CS distribution comparisons represent one patient. **(b)** A more detailed schematic of how our proposed fragility and baseline features were computed from EEG data for a single snapshot of EEG data. See fragility methods section for description of x, A and Δ. For a similar schematic of how the baseline features were computed, see Supplemental figure 10.

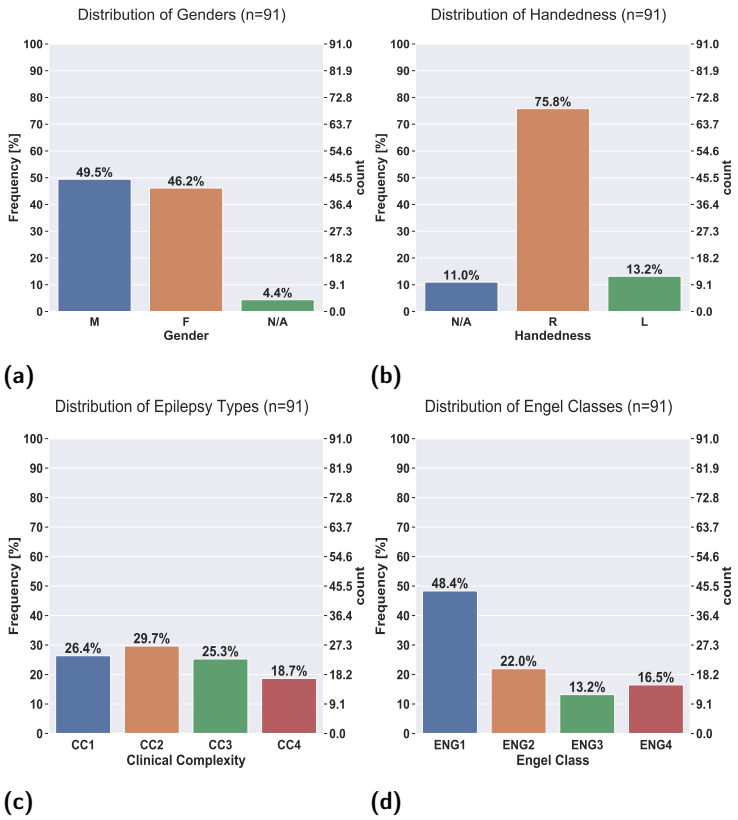


Figure 9: Distributions of the 91 patient dataset based on a variety of clinical factors, such as gender **(a)**, handedness **(b)**, clinical complexity **(c)**, and Engel class **(d)**. The plots show distributions over the 91 patients used in analysis.

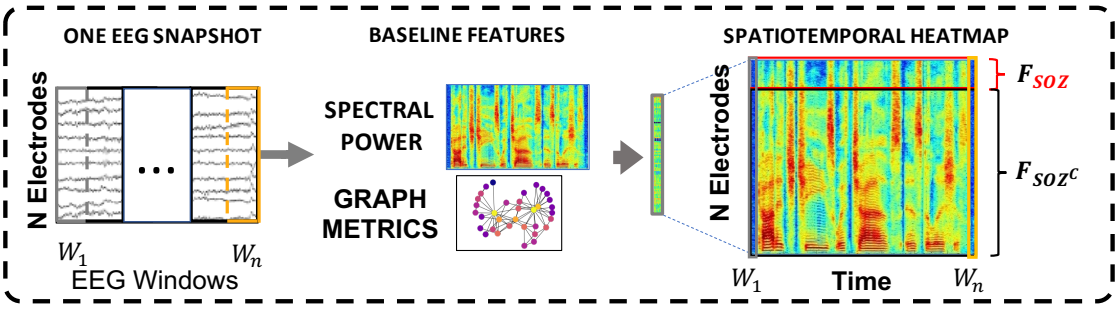


Figure 10: An accompanying schematic for Figure 8, describing how we processed baseline features, such as spectral power, and graph metrics. The feature heatmap processing is exactly the same as fragility, allowing us to compare the feature representations of neural fragility, spectral power and graph metrics of correlation and coherence derived graphs for the purposes of SOZ localization.

5.3 Neural fragility of a network

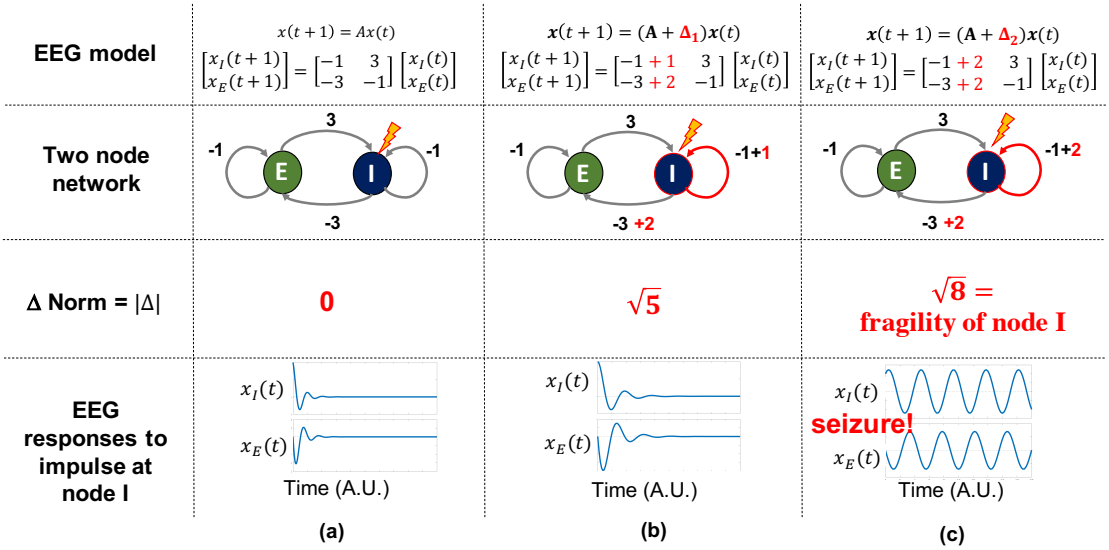


Figure 11: To build quantitative intuition on what neural fragility means in the context of a dynamical iEEG system, we construct a 2-node EEG network example with an excitatory (E) and inhibitory (I) population of neurons. For a qualitative description, see Figure 1. $x_I(t)$ and $x_E(t)$ are the EEG activity of the I and E neuronal population respectively. 'A' is a linear network model quantifying how each population affects the rest over time. Δ (i.e. the fragility), is the amount of change added to a node's connections. The fragility of the node is quantified as the minimal amount of change necessary to cause seizure-like phenomena. **(a)** shows a stable network without a perturbation added, such that the network responses due to an impulse at I result in a transient that reverts to baseline. **(b)** shows a perturbation added, but the network is still stable with a slightly larger transient when an impulse is applied to node I. Then **(c)** shows enough of a perturbation is added, such that the network becomes unstable; an impulse applied at node I results in oscillatory activity that does not quickly return to baseline. The magnitude of the Δ added in **(c)** is the fragility of node I.

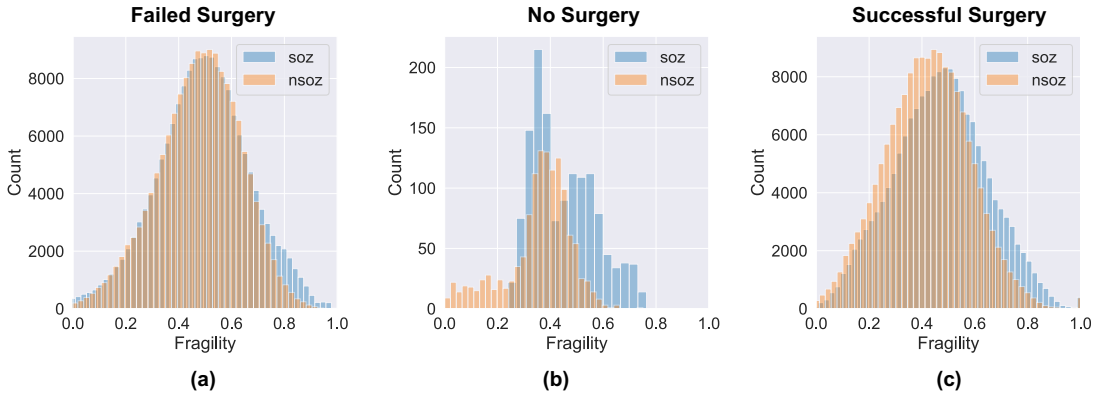


Figure 12: Pooled fragility distribution analysis for all 91 patients: failed (a), no surgery (b) and successful surgery (c) datasets. Each SOZ (soz in blue bars) and SOZ^C ('nsoz' in orange bars) distribution per patient was bootstrap sampled (see Methods section for more information on sampling) and then compared using the one-sided Mann-Whitney U test. The corresponding test yielded a statistic of 2776334 (PValue = 0.355) for the failed patient outcomes and a statistic of 36836739 (PValue = 3.326e-70) for the successful patient outcomes. The patients without resection were not included in the analysis comparing to outcome, but these patients can present as interesting case studies where the SOZ was hypothetically localizable, but perhaps was too close to eloquent areas.

5.4 Pooled subject and SOZ vs SOZ^C analysis

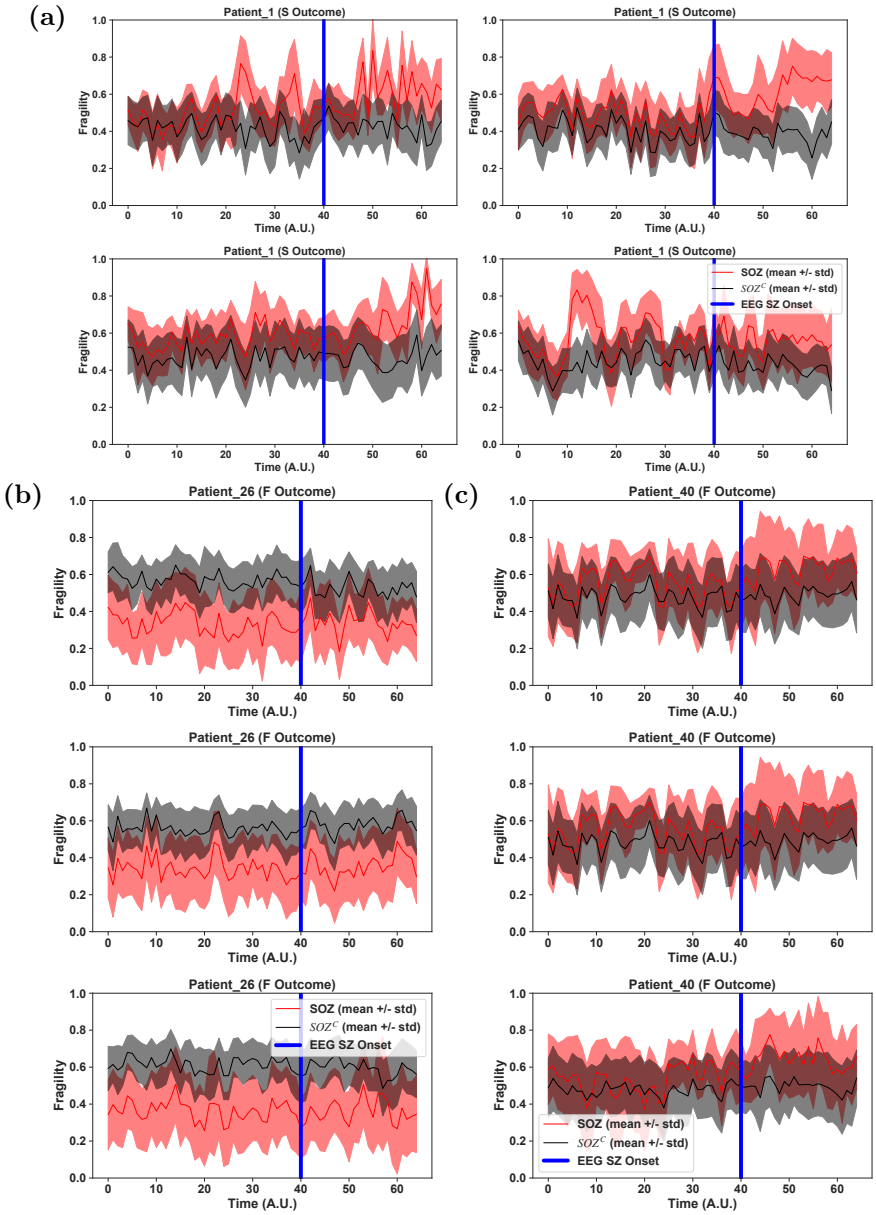
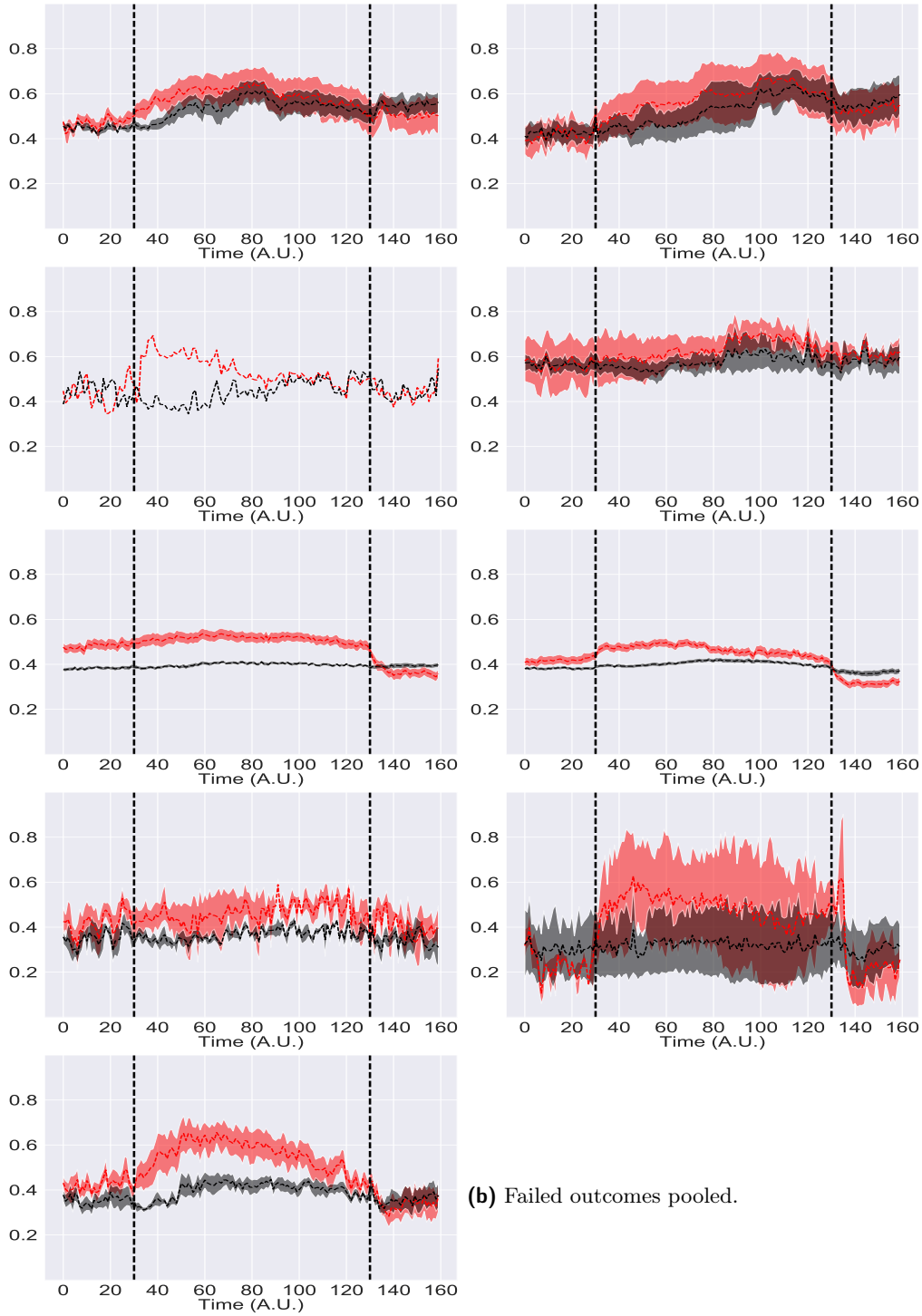


Figure 13: Red SOZ vs black SOZ^C signals for patients presented in Figure 3: Patient_01 (a), Patient_26 (b), Patient_40 (c). For each patient, the ictal snapshots available are visualized around seizure onset with 5 seconds before onset until the first 20% of the seizure. Note that not necessarily all electrodes in the clinically annotated SOZ are part of the EZ when the patient had a successful outcome. Therefore, if neural fragility had value in contrasting true EZ electrodes from non-EZ electrodes, then any extra electrodes clinically annotated in the SOZ should have relative lower fragility.



(b) Failed outcomes pooled.

(a) Successful outcomes pooled.

Figure 14: Red SOZ vs black SOZ^C fragility signals for pooled patients within each of the five centers with successful **(a)** and failed outcomes **(b)** for NIH (n=14), JHH (n=4), CC (n=61), UMH (n=5), and UMMC (n=7) (top to bottom respectively). Note UMMC only had successful outcomes, so there was no curve for the failures. Seizure periods were resampled and normalized to 100 samples for averaging and viewing purposes. In JHH and UMH, there were only one and two patients in successful outcomes respectively.

5.5 Results and interpretations

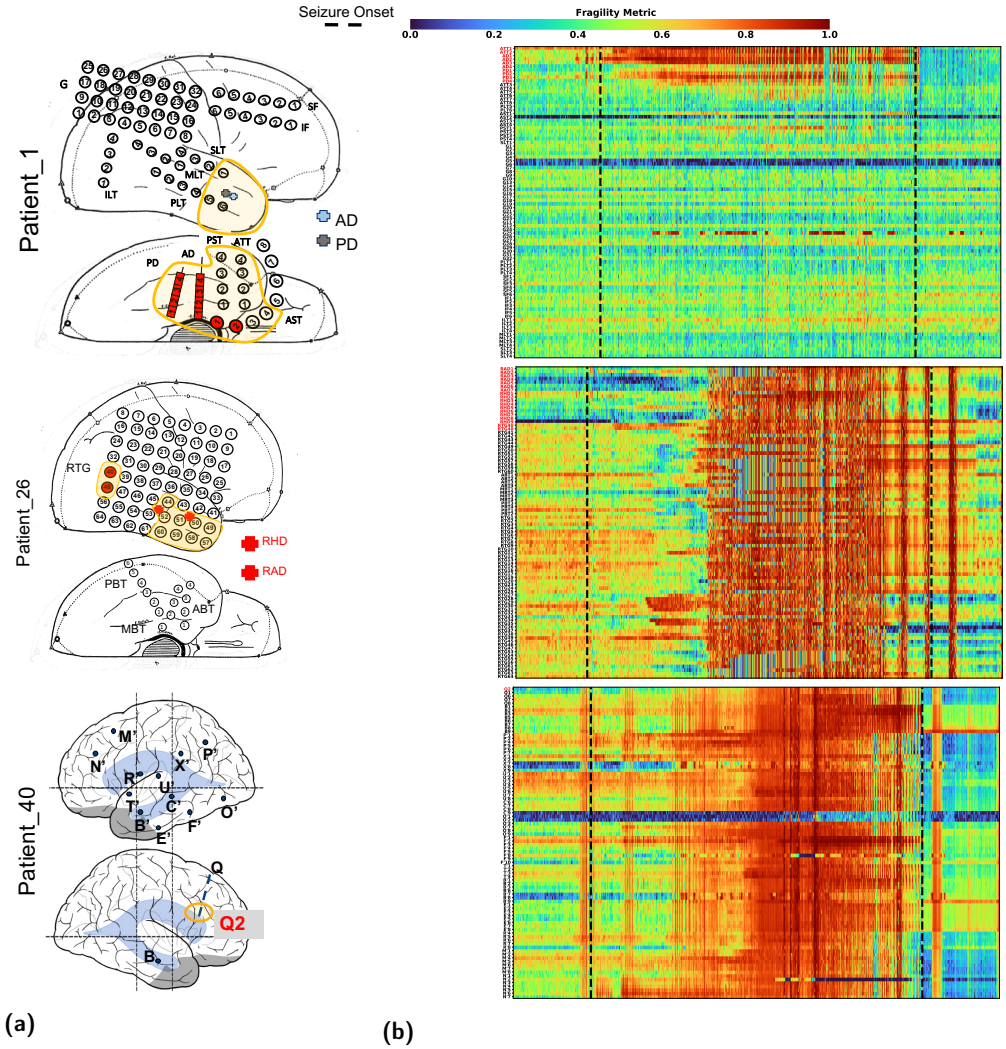


Figure 15: Fragility heatmaps with electrodes on y-axis and time on x-axis with the dashed white-lines denoting seizure onset and offset. Shows a period of 30 seconds before seizure onset and 30 seconds after seizure offset. **(a)** Shows clinically annotated maps of the implanted ECoG/SEEG electrodes with red denoting SOZ contacts. **(b)** shows spatiotemporal fragility heatmaps for example of successful outcome (Patient_01), and failed outcome (Patient_26 and Patient_40). The color scale represents the amplitude of the normalized fragility metric, with closer to 1 denoting fragile regions and closer to 0 denoting relatively stable regions. The contacts in red and orange are part of the SOZ and RZ, respectively as defined in Methods section. Note that the red contacts are also part of the RZ. Visualized with Turbo continuous colormap. Best seen if viewed in color.

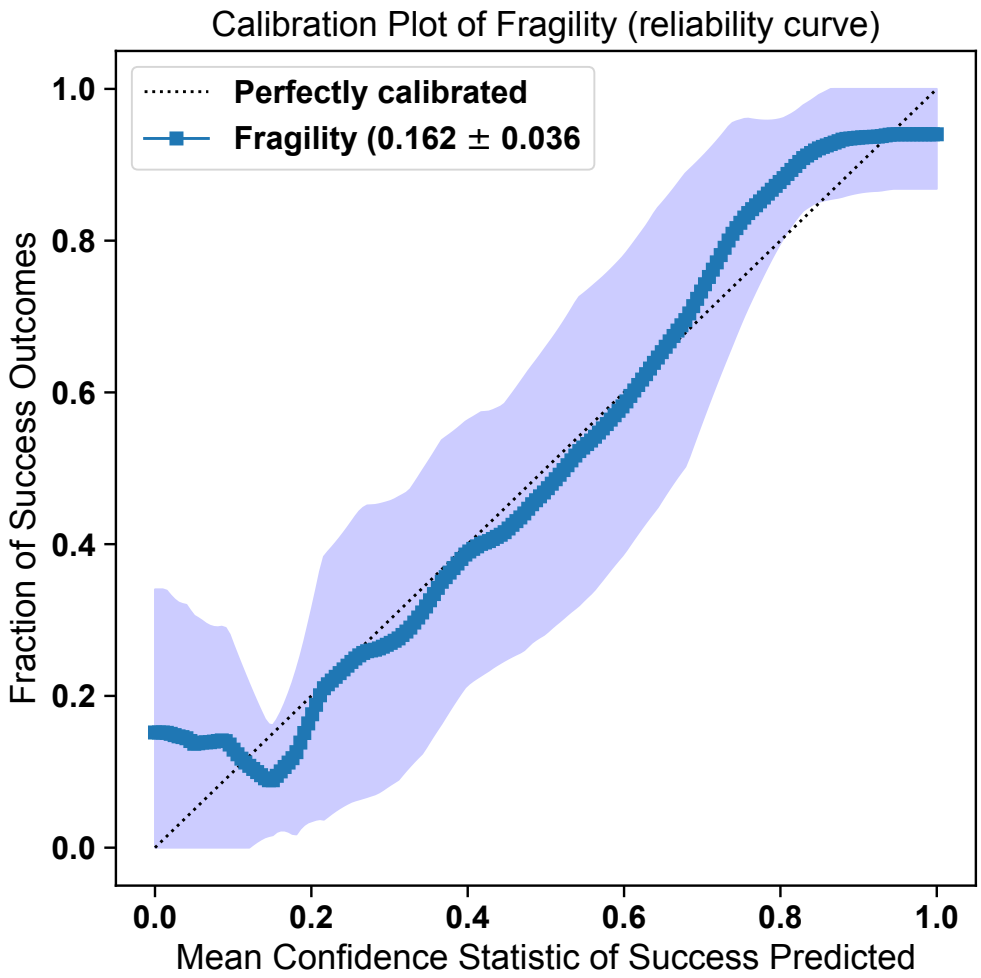


Figure 16: Calibration curve showing the fraction of actual successful surgical outcomes on the y-axis vs the average CS output on the x-axis. This curve measures how calibrated the predicted success probability values are to the true risk stratification of the patient population. The closer a curve is to the $y = x$ line, then the more calibrated a model is. It is quantified by the Brier-loss (closer to 0 is better), which is shown in the legend, and is significantly lower then the next best feature (an improvement of 15%).

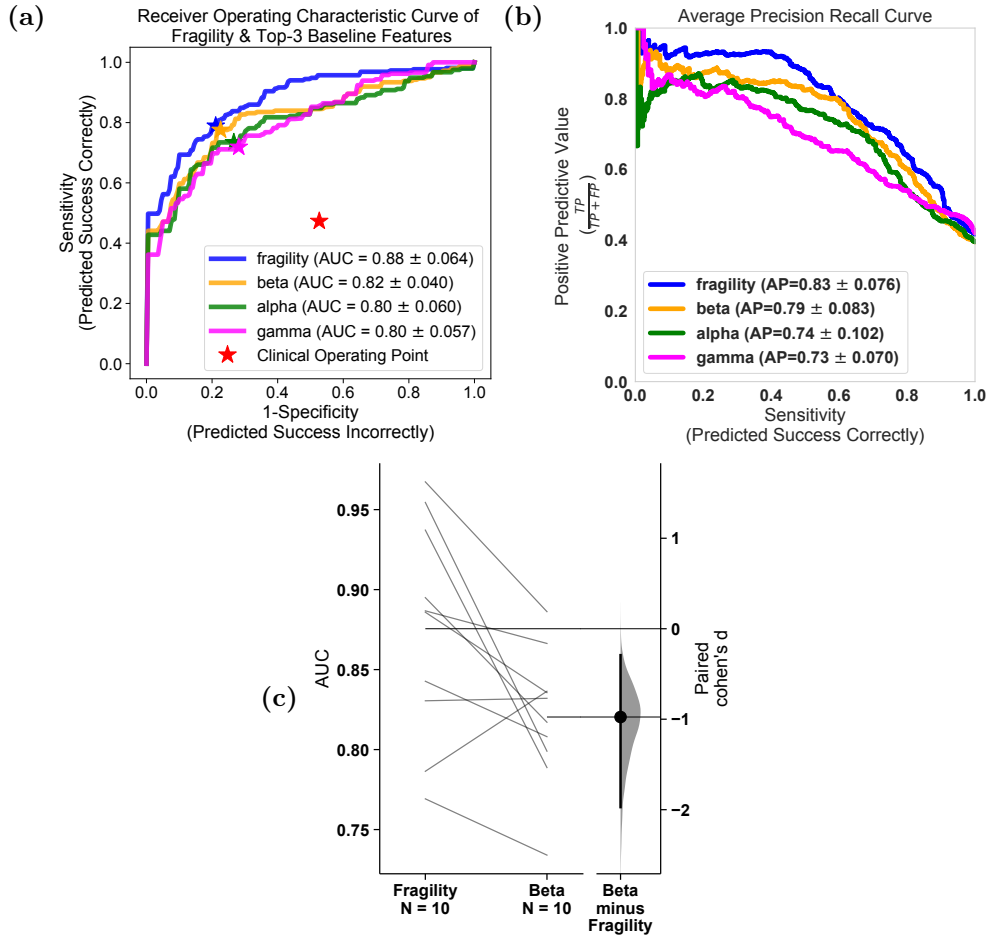


Figure 17: **(a)** The ROC curve over 10 folds of cross-validation of the held-out test set obtained by applying a Random Forest model onto the spatiotemporal heatmaps to predict surgical outcome (See 4 Section). Fragility and the top-3 baseline features in terms of AUC are visualized. The shaded area represents the standard deviation of the curve obtained by linear interpolation for visualization purposes. The AUC of fragility obtained a 0.88 +/- 0.064 over the 10 standard deviation with a relative improvement of 7.2% improvement in AUC compared to the next best feature representation (i.e. the beta frequency band). At the Youden point (stars), neural fragility obtains a balanced accuracy score of 0.76 +/- 0.06, and an improvement of 0.32 in TPR and 0.32 in FPR compared to the clinical operating point (red star). **(c)** A paired estimation plot showing how the same test set of patients differed in AUC depending on whether it was using the fragility, or beta feature heatmap representation. The paired Cohen's D effect size was computed at -0.975 (-1.97 to -0.29; 95% CI). The p-values associated with the difference between Neural Fragility and the Beta frequency band were 0.0204, 0.0273, and 0.0225 using the Wilcoxon rank-sum test, permutation test, and the paired student t-test respectively. Comparing Fragility to all other feature representations were more different in terms of larger effect sizes and smaller p-values.

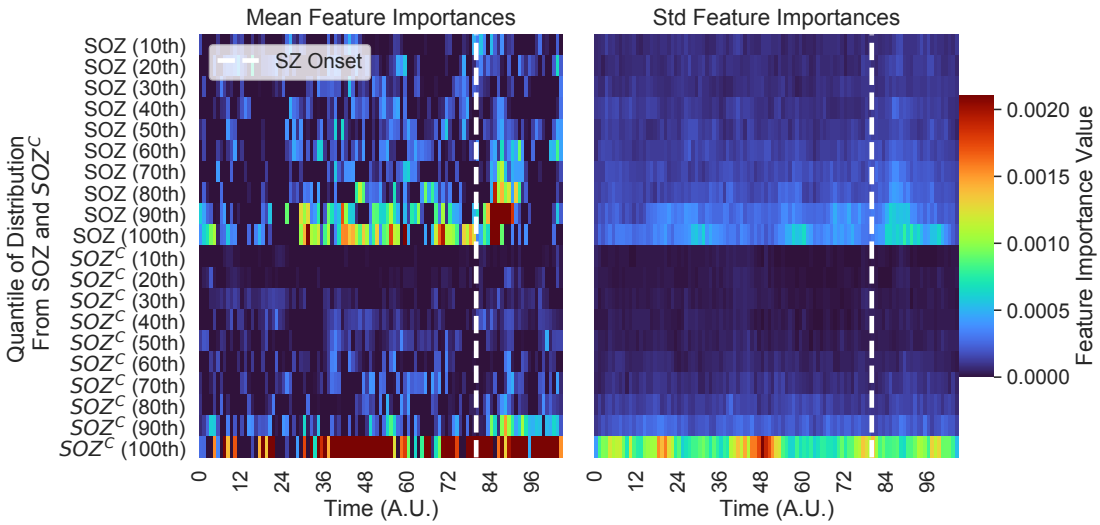


Figure 18: Estimated feature importance (mean and stdev) of the associated fragility heatmap used estimated using permutation. The metric of interest was the concordance statistic (i.e. AUC) of the ROC curve. The original feature map is transformed into a 20-dimensional set of time-varying statistics of its SOZ and SOZ^C electrodes describing the quantiles of the spatiotemporal heatmap (10% - 100% quantiles). This time-varying summary allows these heatmaps to be pooled together across subjects when training a Random Forest classifier as described in Section 4.

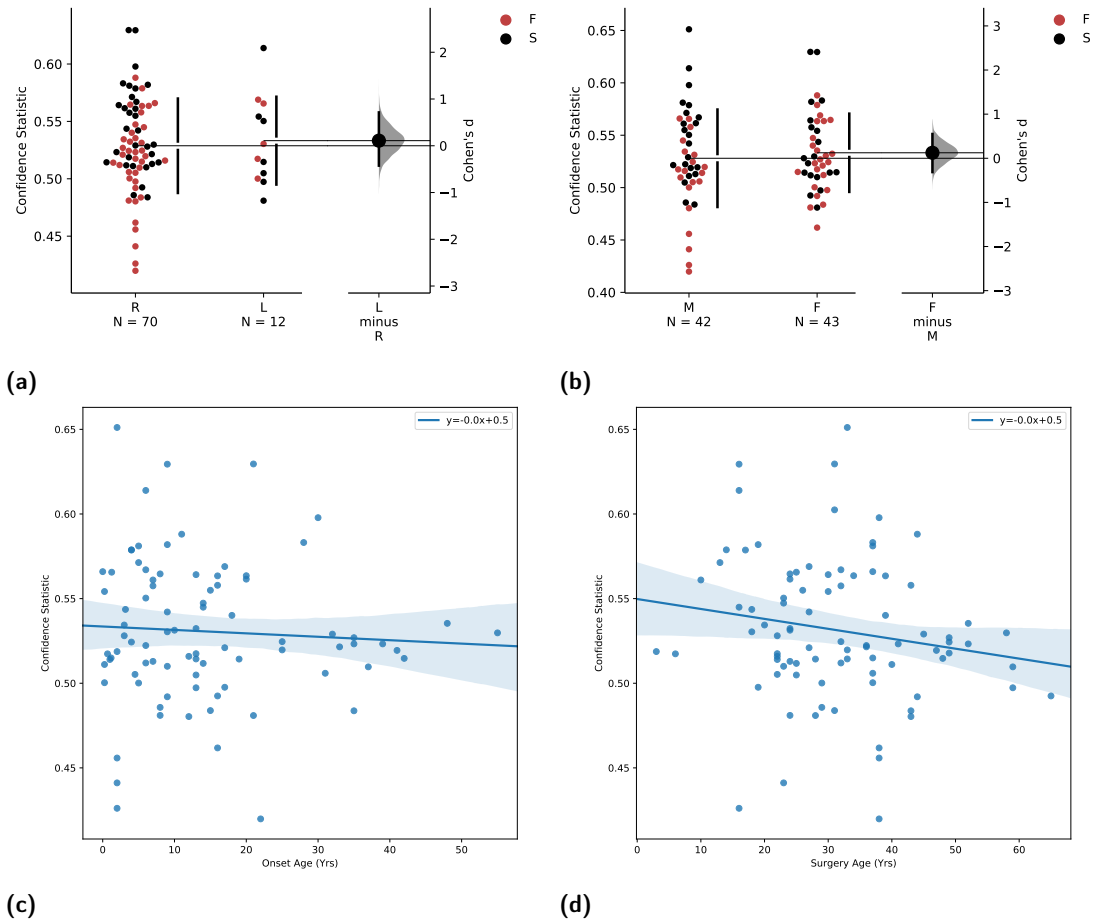


Figure 19: Fragility success probabilities split by clinical factors, such as handedness **(a)**, gender **(b)**, onset age **(c)** and age at surgery **(d)**. Effect sizes were estimated using the permutation test and Mann Whitney U test described in section 4. The corresponding effect sizes and p-values were (0.1/0.99) for handedness, and (0.12/0.7) for gender. The slopes were all negligibly close to 0 for onset age and surgery age linear fit. Note that not all patients had data for each of these categories, so the subset of available data was used.

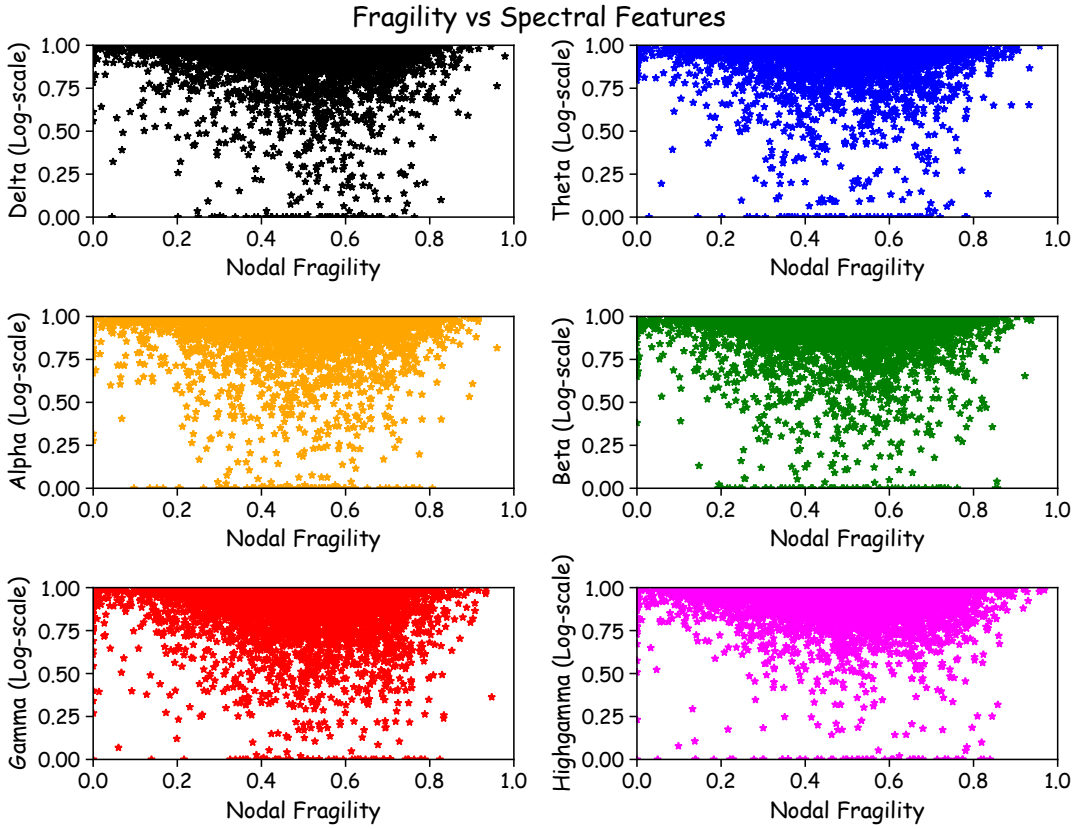


Figure 20: Fragility versus frequency power in the delta, theta, alpha, beta, gamma and highgamma band for Patient_01, Patient_26, and Patient_40. For band definitions, refer to Section 6.6. Every point represents the spectral power and neural fragility value from a randomly chosen window and electrode from one of the patients. No significant correlation is seen or computed from the data. Each spectral feature and fragility are normalized as described in Section 4.

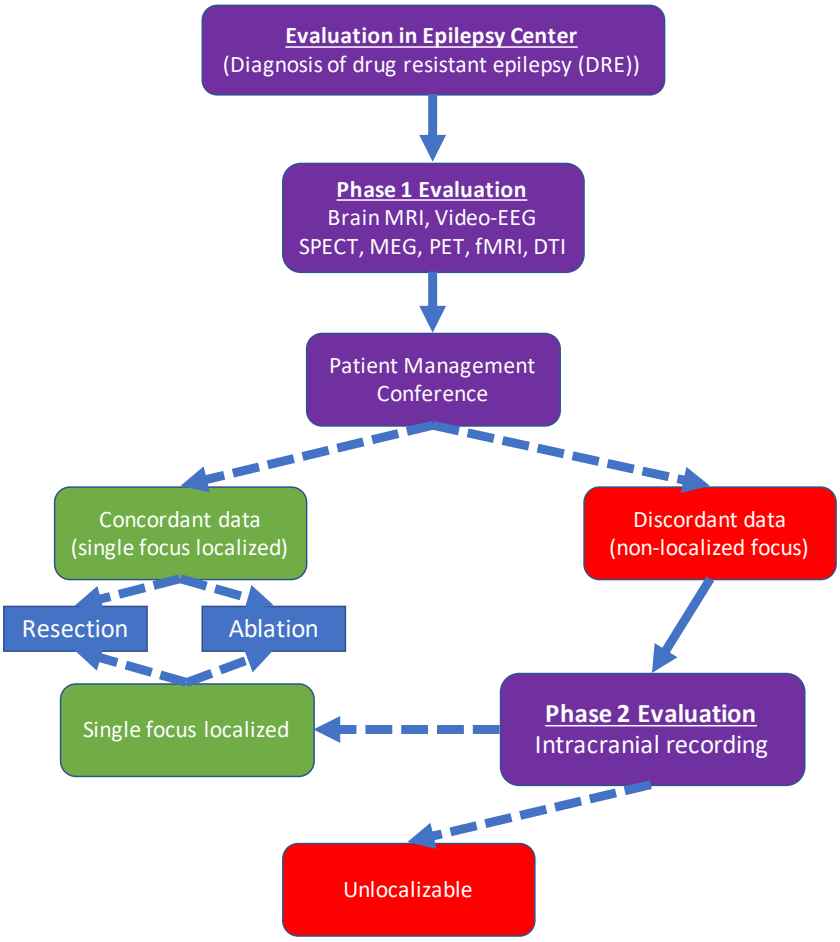


Figure 21: An overview of the clinical pathways that lead to epilepsy surgery. After a neurosurgical consultation, in few cases, concordant non-invasive data allows a patient to proceed to surgery immediately. In most other cases, invasive monitoring takes place, where localization is attempted based on stereotypical electrographic signatures, described in Section 6.2 and in [127]. If the SOZ was localized confidently, and eloquent regions are not affected, then a resection, or laser ablation is performed. The dashed lines indicate where a decision split was made depending on the clinical analyses of the data.

6 Supplementary Materials and Methods

6.1 Code and data availability

All code related to generate the figures are at `https://github.com/adam2392/fragility_in_ieeg` (will be made public once published). We include a jupyter notebook written in Python to help reproduce figures. We also released the raw iEEG data for patients from NIH, UMH, UMMC, and JHH in the OpenNeuro repository in the form of BIDS-iEEG (`https://openneuro.org/datasets/ds003029`). Due to restrictions on data sharing from CClinic, we were unable to release the raw iEEG data that we received from this site.

6.2 Dataset collection

iEEG data from 91 DRE patients who underwent intracranial EEG monitoring, which included either electrocorticography (ECoG), or depth electrodes with stereo-EEG (SEEG) were selected from University of Maryland Medical Center (UMMC), University of Miami Jackson Memorial Hospital (UMH), National Institute of Health (NIH), Johns Hopkins Hospital (JHH), and the Cleveland Clinic (CClinic). Patients exhibiting the following criteria were excluded: no seizures recorded, pregnant patients, EEG sampling rate less than 250 Hz, previous surgeries more then 6 months before the current implantation, and no surgery was performed (possibly from SOZ localizations in eloquent areas). We define successful outcomes as seizure free (Engel class I and ILAE scores of 1 and 2) at 12+ months post-op and failure outcomes with seizure recurrence (Engel classes 2-4). Of these 91 patients, 44 experienced successful outcomes and 47 had failed outcomes (age at surgery = 31.52 +/- 12.32 years) with a total of 462 seizures (seizure length = 97.82 +/- 91.32 seconds) and 14703 total number of recording electrodes (159.82 +/- 45.42 per subject). [128, 129, 46].

Our clinician team categorized patients by surgical outcome, Engel class and ILAE score [130]. In addition, we categorized patients by their clinical complexity (CC) as follows: (1) lesional, (2) focal temporal, (3) focal extratemporal, and (4) multi-focal (Figure 2) [46, 47]. Each of these were categorized based on previous outcome studies that support this increasing level of localization difficulty. Lesional patients have success rates of $\tilde{7}0\%$, experiencing the highest rate of surgical success because the lesions identified through MRI are likely to be part of the SOZ [131, 132, 42, 59, 41, 133]. Localization and surgical success are more challenging patients with non-lesional MRI, with average surgical success rates in temporal, extratemporal and multi-focal epilepsy of $\tilde{6}0\%$, $\tilde{4}5\%$ and $\tilde{3}0\%$, respectively [55, 56, 57, 58]. Patients that fit into multiple categories were placed into the more complex category. Next, clinicians identified electrodes that they hypothesized as SOZ. In general, this was a subset of the resected region for all patients, unless otherwise noted. The epileptologists define the clinically annotated SOZ as the electrodes that participated the earliest in seizures. Every patient's SOZ was labeled by 1-3 trained epileptologists (depending on the center). The corresponding SOZ complement that we define, or SOZ^C, are the electrodes that are not part of the SOZ. For more detailed information regarding each patient, see Supplemental Figure 9 and Supplemental clinical data summary Excel file.

At all centers, data were recorded using either a Nihon Kohden (Tokyo, Japan) or Natus (Pleasanton, USA) acquisition system with a typical sampling rate of 1000 Hz (for details regarding sampling rate per patient, see Supplementary file table). Signals were referenced to a common electrode placed subcutaneously on the scalp, on the mastoid process, or on the subdural grid. At all centers, as part of routine clinical care, up to three board-certified epileptologists marked the electrographic onset and the termination of each seizure by consensus. The time of seizure onset was indicated by a variety of stereotypical electrographic features which included, but were

not limited to, the onset of fast rhythmic activity, an isolated spike or spike-and-wave complex followed by rhythmic activity, or an electrodecremental response. Each snapshot available for a patient was clipped at least 30 seconds before and after the ictal event. We discarded electrodes from further analysis if they were deemed excessively noisy by clinicians, recording from white matter, or were not EEG related (e.g. reference, EKG, or not attached to the brain) which resulted in 97.23 +/- 34.87 (mean +/- std) electrodes used per patient in our analysis. Data was initially stored in the form of the European Data Format (EDF) files [134]. We preprocessed data into the BIDS-iEEG format and performed processing using Python3.6, Numpy, Scipy, MNE-Python and MNE-BIDS [135, 136, 137, 138, 139, 140, 141]. Figures were generated using Matplotlib and Seaborn [142, 143]. Statistical analyses were performed with mlxtend, pingouin, dabest, scipy and scikit-learn [125, 144, 145, 141, 146]. Decisions regarding the need for invasive monitoring and the placement of electrode arrays were made independently of this work and part of routine clinical care. All data were acquired with approval of local Institutional Review Board (IRB) at each clinical institution. The acquisition of data for research purposes was completed with no impact on the clinical objectives of the patient stay. Digitized data were stored in an IRB-approved database compliant with Health Insurance Portability and Accountability Act (HIPAA) regulations.

6.3 Surgical Workup

Epilepsy surgery is a complicated workflow that depends on many clinical factors, such as the history of the patient, and non-invasive testing [48]. We give a brief overview here to provide the reader with a sense of the complexities inherent in evaluating efficacy of localization algorithms, and even the surgical outcomes (see a schematic in Supplementary Figure 21). Every DRE patient generally begins with what is known as a Phase 1 evaluation, where a plethora of non-invasive procedures are done, such as: MRI to look for structural abnormalities, SPECT and PET to look at blood flow and metabolism, fMRI/MEG to map eloquent brain regions and/or hyperactive regions, and neuropsycological testing to determine if cognitive deficits are related to seizure onsets and provide a picture of likely post-operative cognitive deficits. These are just samples of tests that can be done, and active research is being done to explore their utilities in the context of SOZ localization.

When a patient has undergone all the necessary tests, then Phase 2 evaluation begins, where invasive monitoring is considered. iEEG electrodes are implanted into the patient's brain, and are monitored for several weeks in order to observe many seizures. While they are being monitored, patients are taken off anti-epileptic drugs (AEDs), which present serious dangers to the patient because they may have recurring seizures that could lead to sudden death. Once clinicians have sufficiently formed an SOZ hypothesis based on the iEEG data they have collected, and deem that surgery will not affect eloquent regions, then they proceed with surgery. The types of procedures can be broadly classified as a resection, or laser ablation [147, 128, 148, 149].

Assuming that surgery includes the true epileptic tissue, then the patient will be seizure free. On the other hand, there can be many reasons for a failed surgery, such as: neoepileptogenesis [150, 151], kindling [152, 65], a mismatch between the SOZ and resected region, or insufficient resection. Due to the complex nature of epilepsy and the brain, it is difficult to determine the cause of failure. We simply require that a useful feature representation of the electrodes separate the SOZ from the SOZ^C in successful surgeries, and vice-versa in failed surgeries.

6.4 Preprocessing of data

Every dataset was notch filtered at 60 Hz (with a cutoff window of 2 Hz) and bandpass filtered between 0.5 and the Nyquist frequency with a fourth order Butterworth filter. A common average reference was applied to remove any correlated noise [153]. EEG sequences were broken down into sequential windows and the features were computed within each window (see Methods 6.5, 6.6 and 6.7 for details). Each proposed feature representation produces a value for each electrode for each separate window, and results in a full spatiotemporal heatmap when computed over sequential windows with electrodes on the y-axis, time on the x-axis and feature value on the color axis. In total, we computed 20 different baseline feature representations from the iEEG data: 6 frequency power bands, 7 eigenvector centralities (one for each frequency band coherence connectivity matrix and one for a correlation connectivity matrix), 7 in-degrees (one for each frequency band coherence connectivity matrix and one for a correlation connectivity matrix). Values at each window of time were normalized across electrodes to values that could range from 0 up to at most 1, to allow for comparison of relative feature value differences across electrodes over time; the higher a normalized feature, the more we hypothesized that electrode was part of the SOZ [36].

6.5 Neural fragility of a network

To demonstrate how fragility is computed from a dynamical model, we consider a 2-node network as shown in Supplemental Figure 11. In panel a), a stable network is shown where excitation and inhibition are balanced. The network model is provided in the top row and takes a linear form of $\mathbf{x}(t+1) = \mathbf{A}\mathbf{x}(t)$, where t is a time index (typically one millisecond). When the inhibitory node is stimulated by an impulse, both nodes transiently respond and the EEG returns to baseline (bottom row in Supplemental Figure 11). In panel b), the inhibitory node's connections are slightly perturbed in a direction that makes the inhibitory node *less* inhibitory (see red changes to its connectivity to the excitatory node). These changes are reflected in the model and diagram. Now, when the inhibitory node is stimulated by an impulse, the EEG responses from each node have a larger transient response but still return to baseline. Finally, in panel c), the inhibitory node's connections are further perturbed in a direction that makes the inhibitory node *less* inhibitory. Now, when the inhibitory node is stimulated by an impulse, the EEG responses oscillate demonstrating that the network has gone unstable. The fragility of the inhibitory node is thus quantified as square root of 8 which is the norm of the perturbation vector applied to the first column in the network model.

To compute fragility heatmaps from iEEG recordings, we constructed simple linear models as described above, but one for each 250 ms iEEG window. We used a sparse least-squares with a 10e-5 l2-norm regularization to ensure that the model identified was stable as in [37, 36]. Then, we slid the window 125 ms and repeated the process, generating a sequence of linear network models in time as in Supplemental Figure 8b). We systematically computed the minimum perturbation required for each electrode's connections (Supplemental Figure 8b) to produce instability for the entire network as described in [36]. The electrodes that were the most fragile were hypothesized to be related to the SOZ in these epilepsy networks (seen as the dark-red color in the turbo colormap in Figure 3).

6.6 Baseline features - spectral features

We constructed spectral-based features from frequency bands of interest by applying a multi-taper Fourier transform over sliding windows of data with a window/step size of 2.5/0.5 seconds [154, 28]. We required relatively longer time windows in order to accurately estimate some of the

lower frequency bands. Each EEG time series was first transformed into a 3-dimensional array (electrodes X frequency X time), and then averaged within each frequency band to form six different spectral feature representations of the data. We break down frequency bands as follows:

1. Delta Frequency Band [0.5 - 4 Hz]
2. Theta Frequency Band [4 - 8 Hz]
3. Alpha Frequency Band [8 - 13 Hz]
4. Beta Frequency Band [13 - 30 Hz]
5. Gamma Frequency Band [30 - 90 Hz]
6. High-Gamma Frequency Band [90 - 300 Hz]

Note that the High-Gamma frequency band includes frequencies of "ripples" - a category of HFOs.

6.7 Baseline features - graph analysis of networks

We computed a time domain model using Pearson correlation (equation 1) and a frequency domain model using coherence (equation 2). We computed the connectivity matrix using MNE-Python and used the default values [138, 137]. In the equations, (i, j) are the electrode indices, Cov is the covariance, σ is the standard deviation, f is the frequency band, and G is cross-spectral density. Note that these connectivity models attempt to capture linear correlations either in time, or in a specific frequency band, but are not dynamical system representations of the data (i.e. $x(t+1) = Ax(t)$). For each network-based feature, a sliding window/step size of 2.5/0.5 seconds were used, resulting in a sequence of network matrices over time resulting in 3-dimensional arrays (electrodes X electrodes X time) [27, 28, 30].

$$Corr_{ij} = \frac{Cov(x_i, x_j)}{\sigma_{x_i}\sigma_{x_j}} \tag{1}$$

$$Coh_{ij}(f) = \frac{|G_{ij}(f)|^2}{G_{ii}(f)G_{jj}(f)} \tag{2}$$

From each network matrix, we computed the eigenvector centrality [28, 27], and the in-degree [26] features of the network for each electrode across time. Centrality describes how influential a node is within a graph network. In-degree is the weighted sum of the connections that connect to a specific node. Both features are potential measures that attempt to capture the influence of a specific electrode within an iEEG network, through the lens of graph theory.

6.8 Experimental design

We tested if the neural fragility representation of iEEG data localized the SOZ better compared to other proposed features, compared to clinicians and compared to chance. Electrodes with extreme feature values deviating from the average were hypothesized as part of the SOZ.

In order to compute a probability of successful surgical outcome for each patient, we trained a non-parametric machine learning classifier, a Random Forest classifier to output a probability value that the surgery was a success. We input the distribution of feature values in the clinically SOZ and the SOZ^C. We considered a set of hyperparameters that we then performed evaluation over a 10-fold nested cross-validation scheme. We then performed statistical analysis on the final classification performance to determine the most robust feature representation.

Pooled Patient Analysis First, we analyzed the difference in the distributions of neural fragility between SOZ and the SOZ^C. We pooled all patients together, stratified by surgical outcome and compared the neural fragility distributions using a one-sided Mann-Whitney U test (Success pvalue = 3.326e-70, Fail pvalue = 0.355; Supplementary Figure 12). This suggested that there is some effect on average where fragility is higher in the SOZ for successful outcomes, so we next looked at the distributions per patient's seizure snapshot around seizure onset. In Supplementary Figure 13, success outcome patients have a higher neural fragility in the SOZ. This effect is seen when pooling patients across all centers as well, where neural fragility is either i) higher before the seizure onset, or ii) has a marked difference starting at seizure onset (Supplementary Figure 14). Next, we performed a classification experiment (Figure 2) that would determine the robustness of the neural fragility representation at the patient level benchmarked against 20 other features.

Non-parametric Decision-Tree Classifier In order to determine the value of a feature representation of the iEEG data, we posed a binary classification problem where the goal would be to determine the surgical outcome (success or failure) for a particular patient's spatiotemporal heatmap. Each spatiotemporal heatmap was split into its SOZ and SOZ^C set (F_{SOZ} and F_{SOZ^C} in Figure 8). Then each set of electrodes' feature distributions were summarized with its quantile statistics over time, resulting in twenty signals: ten quantiles from 10-100 of the SOZ and ten quantiles of the SOZ^C over time. We used a Random Forest (RF) classifier [52]. Specifically, it was a variant known as the Structured Random Forest, or Manifold Random Forests [51, 50]. The manifold RF approach allows one to encode structural assumptions on the dataset, such as correlations in time of the fed in data matrix. The input data matrix for each RF is a multivariate time series that summarizes the quantile statistics of the SOZ and SOZ^C over time. As a result, we obtained better results using the manifold RF because it took advantage of local correlations in time, rather than treating all values in the data matrix input as independent as done in a traditional RF model. This approach allowed our classifiers to learn faster with less data, compared to treating all inputs as independent in the traditional RF model. For more information on how manifold RF improves on traditional RF, we refer readers to [50, 51]. For every model, we used the default parameters from scikit-learn and the rerf package [146, 51]. The output of the decision tree classifier is:

$$\hat{f}_\theta(F_{SOZ}, F_{SOZ^C}) = P(success)$$

where θ are the trained RF parameters and F_{SOZ} are the heatmap values for the SOZ and F_{SOZ^C} are the heatmap values for the SOZ^C for a specific feature representation F (shown in Supplemental Figure 8). $\hat{f}$ is the function we are trying to estimate, which predicts the probability of successful surgical outcome for a given feature heatmap.

Hyperparameters When looking at iEEG data, clinicians inherently select windows of interest due to external and prior information about when the seizure clinically manifests. Similarly, we consider a window of 10 seconds before seizure onset to the first 5% of the seizure event. This window was chosen apriori to analysis, and we repeated the analyses with slightly varying windows, but the results were consistent. In order to provide further contrast to the spatiotemporal heatmaps, we considered also a threshold between 0.3 and 0.7 (spaced by 0.1) that would be applied to the heatmap such that values below were set to 0. This is analagous to clinicians being able to look at a iEEG data and hone in only on areas that are extreme relative to the rest. We selected a fixed threshold through nested cross-validation, where thresholds are selected on the train/validation set, and then performance is measured on the held out test set.

Structured Heatmap Input In order to train a RF classifier on patients with varying number of channels, we structured the spatiotemporal heatmap inputs to summarize the distributions of the feature values in the SOZ and SOZ^C sets. We converted these into a vector of 20 quantile values (10 quantiles for SOZ and 10 for SOZ^C), taken evenly from the 10th to the 100th quantiles. This forms a data matrix per heatmap, which summarizes the SOZ and SOZ^C distributions over time, fixed around the seizure onset time.

Nested Cross-Validation Feature Evaluation It is common practice when building a supervised machine learning model to incorporate cross-validation (CV) where the dataset is split into a training and testing set. However, when one has hyperparameters (in addition to the machine learning model parameters) it is necessary to control against over-fitting to the test dataset [155]. Due to our hyperparameter selection of the optimal heatmap thresholds, we used a nested CV scheme, where we split the dataset into a training, validation and a "lock-box" dataset, where the hyperparameters were tuned on an inner CV (70% of the dataset) with the training and validation data, and then performance was evaluated on the lock-box dataset. All features were optimized separately, so that their hyperparameters were optimal for that feature. We computed the discrimination statistic known as the area under the curve (AUC). We repeated the nested-cross validation 10 times, resulting a 10-fold CV. In addition, we performed patient-level CV, ensuring that no patient was in multiple splits of the dataset (i.e. any one patient is only in the train, test, or lock-box dataset). We performed sampling with respect to clinical complexity to ensure that there was an approximately some lesional, temporal, extratemporal and multi-focal patients in the training dataset over each CV fold. We found that this performed slightly better for all features compared to random sampling.

Statistical analysis - Hypothesis Testing Success probability values were computed from the RF classifiers trained on the spatiotemporal feature representations of the iEEG data (fragility, spectral features, and graph metrics from correlation and coherence graphs) resulting in a distribution of probabilities for each feature.

To compare the RF model performance across the 20 proposed features of fragility, spectral power and graph metrics, we computed Cohen's D effect size differences between the groups of interest, and then Mann-Whitney U tests for unpaired data, and paired Wilcoxon rank-sum tests for paired data [125]. We corrected for any multiple comparisons using the Holm-Bonferroni step-down method [126]. In some cases, we also present Welch's t-test and the Likelihood Q-ratio test (LQRT) results to show that the results are robust to statistical test chosen. The LQRT has been shown to be more robust compared to both the Mann-Whitney U-test and t-test in the presence of noise [156, 157]. Note that it utilizes a bootstrapping procedure, so it's resolution is 0.001 (i.e. if it produces a p-value of 0, then that means it is < 0.001).

We compared the success probabilities stratified by different clinical covariates: surgical outcome, Engel class, clinical complexity, handedness, gender, onset age, and surgery age. We then estimated the effect size differences between distributions in the form of a Cohen's d statistic [158]. Cohen's d was estimated using a non-parametric bootstrapping on the observed data with 5000 re-samples used to construct a 95% confidence interval [125]. The null hypothesis of our experimental setup was that the success probabilities came from the same population. The alternative hypothesis was that the populations were different (i.e. a feature could distinguish success from failed outcomes, or between different Engel classes).

6.9 Feature evaluation using predicted probability of successful surgery

The fragility and all baseline features proposed generated a spatiotemporal heatmap using EEG snapshots of ictal data (outlined in Figure 8). To compare spatiotemporal heatmaps across features, we computed a probability of success (i.e. a confidence score) that is hypothesized to be high for success outcomes and low for failures for "good" features. We expected that fragility would follow a trend of decreasing confidence as CC, Engel score and ILAE score increase. For each clinical covariate group, we measured the effect size difference via bootstrapped sampling, and the statistical p-value between the distributions. We **hypothesized that**: i) fragility would have an effect size difference significantly different from zero when comparing success vs failed outcomes, ii) in addition, this effect size would correlate with meaningful clinical covariates, such as CC and Engel class and iii) both the effect size and p-value would be better than the proposed baseline features.

The higher the probability of success (closer to 1), the more likely the feature indicated a successful surgery, and the lower it was (closer to 0), the more likely the feature indicated a failed surgery. To compute this value, we first partition the heatmap into a SOZ and SOZ^C as seen in Supplemental Figure 8. This forms the two sets of signals that represent the spatiotemporal feature values of the SOZ set vs the SOZ^C set of electrodes. Then we take windows of interest, where clinicians find most valuable for SOZ localization: the period before seizure and right after seizure onset, and performing a nested CV of RF models. The efficacy of each proposed feature is evaluated based on how well the trained RF model is able to predict the surgical outcomes. We tested our hypotheses stated above by computing a probability of success from each feature heatmap for each patient, and estimated the distribution differences of the CS between various clinical covariates.

6.10 Spatiotemporal feature heatmap interpretability

In order to determine how valid the output probability values are, we first computed a calibration curve, which told us how well-calibrated the probability values were [159]. We furthermore compared the calibration curves across neural fragility, spectral features and graph metrics. With relatively good discrimination measured by the AUC and good calibration, we were able to then compare the CS across clinical complexity (CC) scores, Engel scores and ILAE scores. Since epilepsy and reasons for failed surgeries are so complex, these are clinical methods for stratifying patient groups based on observed etiology. We analyzed how the success probability differs across each of these categories. In addition, we looked at how the success probabilities might differ across other clinical variables, such as sex (M vs F), handedness (R vs L), epilepsy onset age (years), and age during surgery (years).

Qualitatively reading off a spatiotemporal heatmap is highly interpretable as one can match raw iEEG segments to certain time periods in the heatmap. For every patient, the time-varying quantile signals (from 10th to the 100th) of the SOZ and SOZ^C were computed for every iEEG heatmap, as specified in Methods Section 6.8. We used permutations feature importance sampling to obtain a relative weighting of each signal over time and how important it was in allowing the RF to correctly determine surgical outcome. This was visualized as a heatmap showing the mean and std importance scores of the SOZ and SOZ^C statistics, as shown in Figure 18.

In addition, we claim that human interpretability of the spatiotemporal heatmap relies on contrast between the SOZ and SOZ^C regions in successful outcomes. Using the results of the feature importance permutation test, we computed for every heatmap, an interpretability ratio, which was defined as:

$$I = \frac{F_{SOZ}(90th)}{F_{SOZ^C}(90th)}$$

where I is the interpretability ratio for a specific subject's feature heatmap F. $F_{SOZ}(90th)$ is the 90th quantile and up feature values of the SOZ, and $F_{SOZ^C}(90th)$ is for the SOZ^C. This ratio is then stratified according to surgical outcome, which we expect higher ratios for successful outcomes and lower ratios for failed outcomes. We quantified this difference using Cohen's D effect size and Mann Whitney U-test Pvalues (see Methods 6.8).

7 Acknowledgements

AL is supported by NIH T32 EB003383, the NSF GRFP (DGE-1746891), the Arcs Scholarship, Whitaker Fellowship and the Chateaubriand Fellowship. SVS is supported by NIH R21 NS103113, the Coulter Foundation, Maryland Innovation Initiative, US NSF Career Award 1055560 and the Burroughs Wellcome Fund CASI Award 1007274. Computational resources were also provided by the Maryland Advanced Research Computing Center (MARCC). The authors would like to thank Carey Priebe for useful discussions on statistical analyses, Joshua Vogelstein for introducing AL to the concepts of Structured Random Forests, Sarah Kim and Rachel June Smith for helpful reviews of the manuscript and Macauley Breault for help in creating brain figures.

8 Author Contributions

AL, JM and SVS conceived the project. AL, IC, DB, JJ, AC, AK, JHopp, SC, JHaagensen, EJ, WA, NC, ZF, JB and JM collected and supervised human epilepsy recordings and processes related to IRB at their respective centers. IC, DB, JJ, AC, AK oversaw data from UMH. JHopp, SC, and JHaagensen oversaw data from UMMC. EJ, WA and NC oversaw data from JHH. SI and KZ oversaw data from NIH. ZF, JB and JM oversaw data from CClinic. AL organized the data from all centers, converted data to BIDS, conducted the analyses and generated the figures. AL and SVS wrote the paper with input from the other authors. CH helped develop RF model and statistical analyses.

References

[1] M. J. Brodie, S. D. Shorvon, R. Canger, P. Halasz, S. Johannessen, P. Thompson, H. G. Wieser, and P. Wolf. *Commission on European Affairs: Appropriate standards of epilepsy care across Europe.* 1997. DOI: 10.1111/j.1528-1157.1997.tb01224.x.

[2] A. T. Berg and M. M. Kelly. "Defining intractability: Comparisons among published definitions." In: *Epilepsia* 47.2 (2006), pp. 431–436. DOI: 10.1111/j.1528-1167.2006.00440.x.

[3] P. Kwan and M. J. Brodie. "Early Identification of Refractory Epilepsy." In: *N. Engl. J. Med.* 342.5 (2000), pp. 314–319. DOI: 10.1056/NEJM200002033420503.

[4] A. T. Berg. "Identification of Pharmacoresistant Epilepsy." In: *Neurol. Clin.* 27.4 (2009), pp. 1003–1013. DOI: 10.1016/j.ncl.2009.06.001.

[5] WHO and W. H. Organization. *Epilepsy.* 2019.

[6] B. C. Jobst. *Consensus over individualism: Validation of the ILAE definition for drug resistant epilepsy.* 2015. DOI: 10.5698/1535-7511-15.4.172.

[7] A. K. Ngugi, C. Bottomley, I. Kleinschmidt, J. W. Sander, and C. R. Newton. "Estimation of the burden of active and life-time epilepsy: A meta-analytic approach." In: *Epilepsia* 51.5 (2010), pp. 883–890. DOI: 10.1111/j.1528-1167.2009.02481.x.

[8] C. E. Begley and T. L. Durgin. *The direct cost of epilepsy in the United States: A systematic review of estimates.* 2015. DOI: 10.1111/epi.13084.

[9] H. O. Lüders, I. Najm, D. Nair, P. Widdess-Walsh, and W. Bingman. "The epileptogenic zone: General principles." In: *Epileptic Disord.* 8.SUPPL. 2 (2006).

[10] L. Jehi. *The epileptogenic zone: Concept and definition.* 2018. DOI: 10.5698/1535-7597.18.1.12.

[11] W. Penfield. "Epileptogenic lesions." In: *Acta Neurol. Psychiatr. Belg.* 56.2 (1956), pp. 75–88.

[12] W. Penfield and H. Jasper. "Epilepsy and the Functional Anatomy of the Human Brain." In: *J. Am. Med. Assoc.* 155.1 (1954), p. 86. DOI: 10.1001/jama.1954.03690190092039.

[13] M. Morrell. *Brain stimulation for epilepsy: Can scheduled or responsive neurostimulation stop seizures?* 2006. DOI: 10.1097/01.wco.0000218233.60217.84.

[14] B. C. Jobst, T. M. Darcey, V. M. Thadani, and D. W. Roberts. "Brain stimulation for the treatment of epilepsy: Brain Stimulation in Epilepsy." In: *Epilepsia* 51 (2010), pp. 88–92. DOI: 10.1111/j.1528-1167.2010.02618.x.

[15] S. Wiebe, W. T. Blume, J. P. Girvin, and M. Eliasziw. "A randomized, controlled trial of surgery for temporal-lobe epilepsy." In: *N. Engl. J. Med.* 345.5 (2001), pp. 311–318. DOI: 10.1056/NEJM200108023450501.

[16] H. E. Scharfman. *The neurobiology of epilepsy.* 2007. DOI: 10.1007/s11910-007-0053-z.

[17] A. Korzeniewska, M. C. Cervenka, C. C. Jouny, J. R. Perilla, J. Harezlak, G. K. Bergey, P. J. Franaszczuk, and N. E. Crone. "Ictal propagation of high frequency activity is recapitulated in interictal recordings: Effective connectivity of epileptogenic networks recorded with intracranial EEG." In: *Neuroimage* 101 (2014), pp. 96–113.

[18] J. Jeppesen, S. Beniczky, A. Fuglsang-Frederiksen, P. Sidenius, and Y. Jasemian. "Detection of epileptic-seizures by means of power spectrum analysis of heart rate variability: A pilot study." In: *Technol. Heal. Care* 18.6 (2010), pp. 417–426. DOI: 10.3233/THC-2010-0606.

[19] G. Alarcon, C. D. Binnie, R. D. C. Elwes, and C. E. Polkey. "Power spectrum and intracranial EEG patterns at seizure onset in partial epilepsy." In: *Electroencephalogr. Clin. Neurophysiol.* 94.5 (1995), pp. 326–337. DOI: 10.1016/0013-4694(94)00286-T.

[20] B. R. Tharp and W. Gersch. "Spectral analysis of seizures in humans." In: *Comput. Biomed. Res.* 8.6 (1975), pp. 503–521. DOI: 10.1016/0010-4809(75)90023-3.

[21] A. V. Medvedev, A. M. Murro, and K. J. Meador. "Abnormal interictal gamma activity may manifest a seizure onset zone in temporal lobe epilepsy." In: *Int. J. Neural Syst.* 21.2 (2011), pp. 103–114. DOI: 10.1142/S0129065711002699.

[22] S. A. Weiss, G. P. Banks, G. M. McKhann, R. R. Goodman, R. G. Emerson, A. J. Trevelyan, and C. A. Schevon. "Ictal high frequency oscillations distinguish two types of seizure territories in humans." In: *Brain* 136.Pt 12 (2013), pp. 3796–3808. DOI: 10.1093/brain/awt276.

[23] S. V. Gliske, Z. T. Irwin, K. A. Davis, K. Sahaya, C. Chestek, and W. C. Stacey. "Universal automated high frequency oscillation detector for real-time, long term EEG." In: *Clin. Neurophysiol.* 127.2 (2016), pp. 1057–1066. DOI: 10.1016/j.clinph.2015.07.016.

[24] A. Li, Z. Fitzgerald, J. Hopp, E. Johnson, N. Crone, J. Bulacio, J. Martinez-Gonzalez, S. Inati, K. Zaghloul, and S. V. Sarma. "Virtual Cortical Stimulation Mapping of Epilepsy Networks to Localize the Epileptogenic Zone." In: *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS.* Vol. 2019. Cc. Institute of Electrical and Electronics Engineers (IEEE), 2019, pp. 2328–2331. DOI: 10.1109/EMBC.2019.8856591.

[25] P. Shah, A. Ashourvan, F. Mikhail, A. Pines, L. Kini, R. T. Shinohara, D. S. Bassett, B. Litt, and K. A. Davis. "Local structural connectivity directs seizure spread in focal epilepsy." In: *bioRxiv* (2018), p. 406793. DOI: 10.1101/406793.

[26] Y. H. Li, X. L. Ye, Q. Q. Liu, J. W. Mao, P. J. Liang, J. W. Xu, and P. M. Zhang. "Localization of epileptogenic zone based on graph analysis of stereo-EEG." In: *Epilepsy Res.* 128 (2016), pp. 149–157. DOI: 10.1016/j.eplepsyres.2016.10.021.

[27] A. Li, B. Chennuri, S. Subramanian, R. Yaffe, S. Gliske, W. Stacey, R. Norton, A. Jordan, K. A. Zaghloul, S. K. Inati, S. Agrawal, J. J. Haagensen, J. Hopp, C. Atallah, E. Johnson, N. Crone, W. S. Anderson, Z. Fitzgerald, J. Bulacio, J. T. Gale, S. V. Sarma, and J. Gonzalez-Martinez. "Using network analysis to localize the epileptogenic zone from invasive EEG recordings in intractable focal epilepsy." In: *Netw. Neurosci.* 02.02 (2018), pp. 218–240. DOI: 10.1162/netn_a_00043.

[28] S. P. Burns, S. Santaniello, R. B. Yaffe, C. C. Jouny, N. E. Crone, G. K. Bergey, W. S. Anderson, and S. V. Sarma. "Network dynamics of the brain and influence of the epileptic seizure onset zone." In: *Proc. Natl. Acad. Sci.* 111.49 (2014), E5321–E5330. DOI: 10.1073/pnas.1401752111.

[29] R. Yaffe, S. Burns, J. Gale, H. J. Park, J. Bulacio, J. Gonzalez Martinez, and S. V. Sarma. "Brain state evolution during seizure and under anesthesia: A network-based analysis of stereotaxic eeg activity in drug-resistant epilepsy patients." In: *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS.* IEEE, 2012, pp. 5158–5161. DOI: 10.1109/EMBC.2012.6347155.

[30] R. B. Yaffe, P. Borger, P. Megevand, D. M. Groppe, M. A. Kramer, C. J. Chu, S. Santaniello, C. Meisel, A. D. Mehta, and S. V. Sarma. "Physiology of functional and effective networks in epilepsy." In: *Clin. Neurophysiol.* 126.2 (2015), pp. 227–236. DOI: 10.1016/j.clinph.2014.09.009.

[31] A. N. Khambhati, K. A. Davis, T. H. Lucas, B. Litt, and D. S. Bassett. "Virtual Cortical Resection Reveals Push-Pull Network Control Preceding Seizure Evolution." In: *Neuron* 91.5 (2016), pp. 1170–1182. DOI: 10.1016/j.neuron.2016.07.039.

[32] A. N. Khambhati, K. A. Davis, B. S. Oommen, S. H. Chen, T. H. Lucas, B. Litt, and D. S. Bassett. "Dynamic Network Drivers of Seizure Generation, Propagation and Termination in Human Neocortical Epilepsy." In: *PLOS Comput. Biol.* 11.12 (2015). Ed. by M. Kramer, e1004608. DOI: 10.1371/journal.pcbi.1004608.

[33] D. S. Bassett and O. Sporns. "Network neuroscience." In: *Nat. Neurosci.* 20.3 (2017), pp. 353–364. DOI: 10.1038/nn.4502.

[34] M. E. Lynall, D. S. Bassett, R. Kerwin, P. J. McKenna, M. Kitzbichler, U. Muller, and E. Bullmore. "Functional connectivity and brain networks in schizophrenia." In: *J. Neurosci.* 30.28 (2010), pp. 9477–9487. DOI: 10.1523/JNEUROSCI.0333-10.2010.

[35] R. Abreu, A. Leal, and P. Figueiredo. "Identification of epileptic brain states by dynamic functional connectivity analysis of simultaneous EEG-fMRI: a dictionary learning approach." In: *Sci. Rep.* 9.1 (2019), p. 638. DOI: 10.1038/s41598-018-36976-y.

[36] A. Li, S. Inati, K. Zaghloul, and S. Sarma. "Fragility in Epileptic Networks : the Epileptogenic Zone." In: *Am. Control Conf.* 2017, pp. 1–8. DOI: 10.23919/ACC.2017.7963378.

[37] A. Li, K. M. Gunnarsdottir, S. Inati, K. Zaghloul, J. Gale, J. Bulacio, J. Martinez-Gonzalez, and S. V. Sarma. "Linear time-varying model characterizes invasive EEG signals generated from complex epileptic networks." In: *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*. 1. IEEE, 2017, pp. 2802–2805. DOI: 10.1109/EMBC.2017.8037439.

[38] D. Sritharan and S. V. Sarma. "Fragility in Dynamic Networks: Application to Neural Networks in the Epileptic Cortex." In: *Neural Comput.* 26.10 (2014), pp. 2294–2327. DOI: 10.1162/NECO_a_00644.

[39] N. Jette, A. Y. Reid, and S. Wiebe. "What are the limitations of medical therapy for epilepsy? Surgical management of epilepsy Key points." In: *CMAJ* 186.13 (2014), p. 186. DOI: 10.1503/cmaj.121291.

[40] C. E. Begley, M. Famulari, J. F. Annegers, D. R. Lairson, T. F. Reynolds, S. Coan, S. Dubinsky, M. E. Newmark, C. Leibson, E. L. So, and W. A. Rocca. "The cost of epilepsy in the United States: An estimate from population- based clinical and survey data." In: *Epilepsia* 41.3 (2000), pp. 342–351. DOI: 10.1111/j.1528-1157.2000.tb00166.x.

[41] L. E. Jeha, I. Najm, W. Bingaman, D. Dinner, P. Widdess-Walsh, and H. Lüders. "Surgical outcome and prognostic factors of frontal lobe epilepsy surgery." In: *Brain* 130.2 (2007), pp. 574–584. DOI: 10.1093/brain/awl364.

[42] J. C. Bulacio, L. Jehi, C. Wong, J. Gonzalez-Martinez, P. Kotagal, D. Nair, I. Najm, and W. Bingaman. "Long-term seizure outcome after resective surgery in patients evaluated with intracranial electrodes." In: *Epilepsia* 53.10 (2012), pp. 1722–1730. DOI: 10.1111/j.1528-1167.2012.03633.x.

[43] J. A. González-Martínez, T. Srikijvilaikul, D. Nair, and W. E. Bingaman. "Long-term seizure outcome in reoperation after failure of epilepsy surgery." In: *Neurosurgery* 60.5 (2007), pp. 873–879. DOI: 10.1227/01.NEU.0000255438.13871.FA.

[44] W. C. Stacey and B. Litt. *Technology Insight: Neuroengineering and epilepsy - Designing devices for seizure control.* 2008. DOI: 10.1038/ncpneuro0750.

[45] H. G. Wieser. "Epilepsy surgery: Past, present and future." In: *Seizure* 7.3 (1998), pp. 173–184. DOI: 10.1016/S1059-1311(98)80032-0.

[46] J. C. Bulacio, P. Chauvel, and A. McGonigal. *Stereoelectroencephalography: Interpretation.* 2016. DOI: 10.1097/WNP.0000000000000305.

[47] S. Sheikh, D. Nair, R. Gross, and J. Gonzalez Martinez. "Tracking a changing paradigm and the modern face of epilepsy surgery: A comprehensive and critical review on the hunt for the optimal extent of resection in mesial temporal lobe epilepsy." In: *Epilepsia* 60.9 (2019), pp. 1768–1793. DOI: 10.1111/epi.16310.

[48] B. Frauscher. "Localizing the epileptogenic zone." In: *Curr. Opin. Neurol.* (2020), pp. 1–9. DOI: 10.1097/WCO.0000000000000790.

[49] O. Grinenko, J. Li, J. C. Mosher, I. Z. Wang, J. C. Bulacio, J. Gonzalez-Martinez, D. Nair, I. Najm, R. M. Leahy, and P. Chauvel. "A fingerprint of the epileptogenic zone in human epilepsies." In: *Brain* 141.1 (2018), pp. 117–131. DOI: 10.1093/brain/awx306.

[50] R. Perry, T. M. Tomita, R. Mehta, J. Arroyo, J. Patsolic, B. Falk, and J. T. Vogelstein. "Manifold Forests: Closing the Gap on Neural Networks." In: *Arxiv* (2019).

[51] T. M. Tomita, J. Browne, C. Shen, J. Chung, J. L. Patsolic, B. Falk, C. E. Priebe, J. Yim, R. Burns, M. Maggioni, and J. T. Vogelstein. "Sparse projection oblique randomer forests." In: *J. Mach. Learn. Res.* 21 (2020).

[52] L. Breiman. "Random Forests." In: *Mach. Learn.* 45 (2001), pp. 5–32.

[53] R. Couronné, P. Probst, and A. L. Boulesteix. "Random forest versus logistic regression: A large-scale benchmark experiment." In: *BMC Bioinformatics* 19.1 (2018), p. 270. DOI: 10.1186/s12859-018-2264-5.

[54] S. S. Sawilowsky. "Very large and huge effect sizes." In: *J. Mod. Appl. Stat. Methods* 8.2 (2009), pp. 597–599. DOI: 10.22237/jmasm/1257035100.

[55] S. Sheikh, N. Thompson, W. Bingaman, J. Gonzalez Martinez, I. Najm, and L. Jehi. "Redefining success in epilepsy surgery: The importance of relative seizure reduction in patient reported quality of life." In: *Epilepsia* August (2019), pp. 2078–2085. DOI: 10.1111/epi.16327.

[56] P. Ryvlin, J. H. Cross, and S. Rheims. *Epilepsy surgery in children and adults.* 2014. DOI: 10.1016/S1474-4422(14)70156-5.

[57] V. N. Vakharia, J. S. Duncan, J. A. Witt, C. E. Elger, R. Staba, and J. Engel. "Getting the best outcomes from epilepsy surgery." In: *Ann. Neurol.* (2018). DOI: 10.1002/ana.25205.

[58] H. Chen, M. Agostini, K. Ding, P. Gupta, C. Madden, B. Mickey, P. Van Ness, and P. Modur. "Predictors of seizure recurrence and longitudinal outcome after epilepsy surgery." In: *Neurology* 80.1 (2013).

[59] S. J. See, L. E. Jehi, S. Vadera, J. Bulacio, I. Najm, and W. Bingaman. "Surgical outcomes in patients with extratemporal epilepsy and subtle or normal magnetic resonance imaging findings." In: *Neurosurgery* 73.1 (2013), pp. 68–76. DOI: 10.1227/01.neu.0000429839.76460.b7.

[60] J. F. Téllez-Zenteno, R. Dhar, and S. Wiebe. *Long-term seizure outcomes following epilepsy surgery: A systematic review and meta-analysis.* 2005. DOI: 10.1093/brain/awh449.

[61] J. Tellez-Zenteno. "Reasons for reoperation after epilepsy surgery: a review based on a complex clinical case with three operations." In: *Neuropsychiatr. Dis. Treat.* 6 (2010), p. 409. DOI: 10.2147/ndt.s10384.

[62] Y. Varatharajah, M. J. Chong, K. Saboo, B. Berry, B. Brinkmann, G. Worrell, and R. Iyer. "EEG-GRAPH: A factor-graph-based model for capturing spatial, temporal, and observational relationships in electroencephalograms." In: *Adv. Neural Inf. Process. Syst.* Vol. 2017-Decem. 2017, pp. 5372–5381.

[63] Y. Varatharajah, B. Berry, J. Cimbalnik, V. Kremen, J. Van Gompel, M. Stead, B. Brinkmann, R. Iyer, and G. Worrell. "Integrating artificial intelligence with real-time intracranial EEG monitoring to automate interictal identification of seizure onset zones in focal epilepsy." In: *J. Neural Eng.* 15.4 (2018). DOI: 10.1088/1741-2552/aac960.

[64] A. Sharma. "Epileptic seizure prediction using power analysis in beta band of EEG signals." In: *Int. Conf. Soft Comput. Tech. Implementations, ICSCTI 2015.* Institute of Electrical and Electronics Engineers Inc., 2016, pp. 117–121. DOI: 10.1109/ICSCTI.2015.7489552.

[65] E. Bertram. *The relevance of kindling for human epilepsy.* 2007. DOI: 10.1111/j.1528-1167.2007.01068.x.

[66] K. D. Laxer, E. Trinka, L. J. Hirsch, F. Cendes, J. Langfitt, N. Delanty, T. Resnick, and S. R. Benbadis. *The consequences of refractory epilepsy and its treatment.* 2014. DOI: 10.1016/j.yebeh.2014.05.031.

[67] I. M. Najm, R. Naugle, R. M. Busch, W. Bingaman, and H. Lüders. "Definition of the epileptogenic zone in a patient with non-lesional temporal lobe epilepsy arising from the dominant hemisphere." In: *Epileptic Disord.* Vol. 8. SUPPL. 2. 2006, S27–35.

[68] J. Jeppesen, S. Beniczky, P. Johansen, P. Sidenius, and A. Fuglsang-Frederiksen. "Detection of epileptic seizures with a modified heart rate variability algorithm based on Lorenz plot." In: *Seizure* 24.C (2015), pp. 1–7. DOI: 10.1016/j.seizure.2014.11.004.

[69] C. E. Priebe, D. L. Sussman, M. Tang, and J. T. Vogelstein. "Statistical Inference on Errorfully Observed Graphs." In: *J. Comput. Graph. Stat.* 24.4 (2015), pp. 930–953. DOI: 10.1080/10618600.2014.951049.

[70] A. Bragin, C. L. Wilson, R. J. Staba, M. Reddick, I. Fried, and J. Engel. "Interictal high-frequency oscillations (80-500Hz) in the human epileptic brain: Entorhinal cortex." In: *Ann. Neurol.* 52.4 (2002), pp. 407–415. DOI: 10.1002/ana.10291.

[71] G. Worrell and J. Gotman. *High-frequency oscillations and other electrophysiological biomarkers of epilepsy: Clinical studies.* 2011. DOI: 10.2217/bmm.11.74.

[72] T. Fedele, S. Burnos, E. Boran, N. Krayenbühl, P. Hilfiker, T. Grunwald, and J. Sarnthein. "Resection of high frequency oscillations predicts seizure outcome in the individual patient." In: *Sci. Rep.* 7.1 (2017), p. 13836. DOI: 10.1038/s41598-017-13064-1.

[73] S. A. Weiss, Z. Waldman, F. Raimondo, D. Slezak, M. Donmez, G. Worrell, A. Bragin, J. Engel, R. Staba, and M. Sperling. "Localizing epileptogenic regions using high-frequency oscillations and machine learning." In: *Biomark. Med.* 13.5 (2019), pp. 409–418. DOI: 10.2217/bmm-2018-0335.

[74] J. Jacobs, R. Staba, E. Asano, H. Otsubo, J. Y. Wu, M. Zijlmans, I. Mohamed, P. Kahane, F. Dubeau, V. Navarro, and J. Gotman. "High-frequency oscillations (HFOs) in clinical epilepsy." In: *Prog. Neurobiol.* 98.3 (2012), pp. 302–315. DOI: 10.1016/j.pneurobio.2012.03.001.

[75] B. Frauscher, F. Bartolomei, K. Kobayashi, J. Cimbalnik, M. A. van 't Klooster, S. Rampp, H. Otsubo, Y. Höller, J. Y. Wu, E. Asano, J. Engel, P. Kahane, J. Jacobs, and J. Gotman. *High-frequency oscillations: The state of clinical research.* 2017. DOI: 10.1111/epi.13829.

[76] R. S. Fisher, W. R. S. Webber, R. P. Lesser, S. Arroyo, and S. Uematsu. "High-frequency EEG activity at the start of seizures." In: *J. Clin. Neurophysiol.* 9.3 (1992), pp. 441–448. DOI: 10.1097/00004691-199207010-00012.

[77] J. D. Jirsch, E. Urrestarazu, P. LeVan, A. Olivier, F. Dubeau, and J. Gotman. "High-frequency oscillations during human focal seizures." In: *Brain* 129.6 (2006), pp. 1593–1608. DOI: 10.1093/brain/awl085.

[78] M. Zijlmans, P. Jiruska, R. Zelmann, F. S. S. Leijten, J. G. R. Jefferys, and J. Gotman. "High-frequency oscillations as a new biomarker in epilepsy." In: *Ann. Neurol.* 71.2 (2012), pp. 169–178. DOI: 10.1002/ana.22548.

[79] G. A. Worrell, L. Parish, S. D. Cranstoun, R. Jonas, G. Baltuch, and B. Litt. "High-frequency oscillations and seizure generation in neocortical epilepsy." In: *Brain* 127.7 (2004), pp. 1496–1506. DOI: 10.1093/brain/awh149.

[80] A. Matsumoto, B. H. Brinkmann, S. M. Stead, J. Matsumoto, M. T. Kucewicz, W. R. Marsh, F. Meyer, G. Worrell, S. Matthew Stead, J. Matsumoto, M. T. Kucewicz, W. R. Marsh, F. Meyer, and G. Worrell. "Pathological and physiological high-frequency oscillations in focal human epilepsy." In: *J. Neurophysiol.* 110.8 (2013), pp. 1958–1964. DOI: 10.1152/jn.00341.2013.

[81] A. Thomschewski, A. S. Hincapié, and B. Frauscher. *Localization of the epileptogenic zone using high frequency oscillations.* 2019. DOI: 10.3389/fneur.2019.00094.

[82] V. Gnatkovsky, M. De Curtis, C. Pastori, F. Cardinale, G. Lo Russo, R. Mai, L. Nobili, I. Sartori, L. Tassi, and S. Francione. "Biomarkers of epileptogenic zone defined by quantified stereo-EEG analysis." In: *Epilepsia* 55.2 (2014), pp. 296–305. DOI: 10.1111/epi.12507.

[83] S. V. Gliske, Z. T. Irwin, C. Chestek, G. L. Hegeman, B. Brinkmann, O. Sagher, H. J. L. Garton, G. A. Worrell, and W. C. Stacey. "Variability in the location of high frequency oscillations during prolonged intracranial EEG recordings." In: *Nat. Commun.* 9.1 (2018), p. 2155. DOI: 10.1038/s41467-018-04549-2.

[84] S. J. Peng, C. C. Chou, H. Y. Yu, C. Chen, D. J. Yen, S. Y. Kwan, S. P. Hsu, C. F. Lin, H. H. Chen, and C. C. Lee. "Ictal networks of temporal lobe epilepsy: Views from high-frequency oscillations in stereoelectroencephalography." In: *J. Neurosurg.* 131.4 (2019), pp. 1086–1094. DOI: 10.3171/2018.6.JNS172844.

[85] C. Haegelen, P. Perucca, C. E. Châtillon, L. Andrade-Valença, R. Zelmann, J. Jacobs, D. L. Collins, F. Dubeau, A. Olivier, and J. Gotman. "High-frequency oscillations, extent of surgical resection, and surgical outcome in drug-resistant focal epilepsy." In: *Epilepsia* 54.5 (2013), pp. 848–857. DOI: 10.1111/epi.12075.

[86] A. Bragin, J. Engel, and R. J. Staba. *High-frequency oscillations in epileptic brain.* 2010. DOI: 10.1097/WCO.0b013e3283373ac8.

[87] J. Jacobs, M. Zijlmans, R. Zelmann, C. É. Chatillon, J. Hall, A. Olivier, F. Dubeau, and J. Gotman. "High-frequency electroencephalographic oscillations correlate with outcome of epilepsy surgery." In: *Ann. Neurol.* 67.2 (2010), pp. 209–220. DOI: 10.1002/ana.21847.

[88] N. Usui, K. Terada, K. Baba, K. Matsuda, F. Nakamura, K. Usui, M. Yamaguchi, T. Tottori, S. Umeoka, S. Fujitani, A. Kondo, T. Mihara, and Y. Inoue. "Clinical significance of ictal high frequency oscillations in medial temporal lobe epilepsy." In: *Clin. Neurophysiol.* 122.9 (2011), pp. 1693–1700. DOI: 10.1016/j.clinph.2011.02.006.

[89] H. Fujiwara, H. M. Greiner, K. H. Lee, K. D. Holland-Bouley, J. H. Seo, T. Arthur, F. T. Mangano, J. L. Leach, and D. F. Rose. "Resection of ictal high-frequency oscillations leads to favorable surgical outcome in pediatric epilepsy." In: *Epilepsia* 53.9 (2012), pp. 1607–1617. DOI: 10.1111/j.1528-1167.2012.03629.x.

[90] H. Leung, C. X. Zhu, D. T. Chan, W. S. Poon, L. Shi, V. C. Mok, and L. K. Wong. "Ictal high-frequency oscillations and hyperexcitability in refractory epilepsy." In: *Clin. Neurophysiol.* 126.11 (2015), pp. 2049–2057. DOI: 10.1016/j.clinph.2015.01.009.

[91] H. Fujiwara, J. L. Leach, H. M. Greiner, K. D. Holland-Bouley, D. F. Rose, T. Arthur, and F. T. Mangano. "Resection of ictal high frequency oscillations is associated with favorable surgical outcome in pediatric drug resistant epilepsy secondary to tuberous sclerosis complex." In: *Epilepsy Res.* 126 (2016), pp. 90–97. DOI: 10.1016/j.eplepsyres.2016.07.005.

[92] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C. Y. Fu, and A. C. Berg. "SSD: Single shot multibox detector." In: *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* 9905 LNCS (2016), pp. 21–37. DOI: 10.1007/978-3-319-46448-0_2.

[93] Y. Iimura, K. Jones, K. Hattori, Y. Okazawa, A. Noda, K. Hoashi, Y. Nonoda, E. Asano, T. Akiyama, C. Go, A. Ochi, O. C. Snead, E. J. Donner, J. T. Rutka, J. M. Drake, and H. Otsubo. "Epileptogenic high-frequency oscillations skip the motor area in children with multilobar drug-resistant epilepsy." In: *Clin. Neurophysiol.* 128.7 (2017), pp. 1197–1205. DOI: 10.1016/j.clinph.2017.03.031.

[94] Y. Iimura, K. Jones, L. Takada, I. Shimizu, M. Koyama, K. Hattori, Y. Okazawa, Y. Nonoda, E. Asano, T. Akiyama, C. Go, A. Ochi, O. C. Snead, E. J. Donner, J. T. Rutka, J. M. Drake, and H. Otsubo. "Strong coupling between slow oscillations and wide fast ripples in children with epileptic spasms: Investigation of modulation index and occurrence rate." In: *Epilepsia* 59.3 (2018), pp. 544–554. DOI: 10.1111/epi.13995.

[95] M. Pail, P. Řehulka, J. Cimbálník, I. Doležalová, J. Chrastina, and M. Brázdil. "Frequency-independent characteristics of high-frequency oscillations in epileptic and non-epileptic regions." In: *Clin. Neurophysiol.* 128.1 (2017), pp. 106–114. DOI: 10.1016/j.clinph.2016.10.011.

[96] S. Wang, N. K. So, B. Jin, I. Z. Wang, J. C. Bulacio, R. Enatsu, S. Dai, Z. Chen, J. Gonzalez-Martinez, and I. M. Najm. "Interictal ripples nested in epileptiform discharge help to identify the epileptogenic zone in neocortical epilepsy." In: *Clin. Neurophysiol.* 128.6 (2017), pp. 945–951. DOI: 10.1016/j.clinph.2017.03.033.

[97] C. Cuello-Oderiz, N. von Ellenrieder, R. Sankhe, A. Olivier, J. Hall, F. Dubeau, and J. Gotman. "Value of ictal and interictal epileptiform discharges and high frequency oscillations for delineating the epileptogenic zone in patients with focal cortical dysplasia." In: *Clin. Neurophysiol.* 129.6 (2018), pp. 1311–1319. DOI: 10.1016/j.clinph.2018.02.003.

[98] Y. Höller, R. Kutil, L. Klaffenböck, A. Thomschewski, P. M. Höller, A. C. Bathke, J. Jacobs, A. C. Taylor, R. Nardone, and E. Trinka. "High-frequency oscillations in epilepsy and surgical outcome. A meta-analysis." In: *Front. Hum. Neurosci.* 9 (2015), p. 574. DOI: 10.3389/fnhum.2015.00574.

[99] A. M. Spring, D. J. Pittman, Y. Aghakhani, J. Jirsch, N. Pillay, L. E. Bello-Espinosa, C. Josephson, and P. Federico. "Interrater reliability of visually evaluated high frequency oscillations." In: *Clin. Neurophysiol.* 128.3 (2017), pp. 433–441. DOI: 10.1016/j.clinph.2016.12.017.

[100] M. A. van 't Klooster, N. E. van Klink, F. S. Leijten, R. Zelmann, T. A. Gebbink, P. H. Gosselaar, K. P. Braun, G. J. Huiskamp, and M. Zijlmans. "Residual fast ripples in the intraoperative corticogram predict epilepsy surgery outcome." In: *Neurology* 85.2 (2015), pp. 120–128. DOI: 10.1212/WNL.0000000000001727.

[101] M. A. van 't Klooster, N. E. van Klink, W. J. Zweiphenning, F. S. Leijten, R. Zelmann, C. H. Ferrier, P. C. van Rijen, W. M. Otte, K. P. Braun, G. J. Huiskamp, and M. Zijlmans. "Tailoring epilepsy surgery with fast ripples in the intraoperative electrocorticogram." In: *Ann. Neurol.* 81.5 (2017), pp. 664–676. DOI: 10.1002/ana.24928.

[102] A. B. Gardner, G. A. Worrell, E. Marsh, D. Dlugos, and B. Litt. "Human and automated detection of high-frequency oscillations in clinical intracranial EEG recordings." In: *Clin. Neurophysiol.* 118.5 (2007), pp. 1134–1143. DOI: 10.1016/j.clinph.2006.12.019.

[103] R. J. Staba, C. L. Wilson, A. Bragin, and I. Fried. "Quantitative analysis of high-frequency oscillations (80-500 Hz) recorded in human epileptic hippocampus and entorhinal cortex." In: *J. Neurophysiol.* 88.4 (2002), pp. 1743–1752. DOI: 10.1152/jn.2002.88.4.1743.

[104] P. Höller, E. Trinka, and Y. Höller. *High-Frequency Oscillations in the Scalp Electroencephalogram: Mission Impossible without Computational Intelligence.* 2018. DOI: 10.1155/2018/1638097.

[105] M. A. van 't Klooster, N. E. van Klink, D. van Blooijs, C. H. Ferrier, K. P. Braun, F. S. Leijten, G. J. Huiskamp, and M. Zijlmans. "Evoked versus spontaneous high frequency oscillations in the chronic electrocorticogram in focal epilepsy." In: *Clin. Neurophysiol.* 128.5 (2017), pp. 858–866. DOI: 10.1016/j.clinph.2017.01.017.

[106] J. Engel and F. L. da Silva. *High-frequency oscillations - Where we are and where we need to go.* 2012. DOI: 10.1016/j.pneurobio.2012.02.001.

[107] K. K, A. T, A. T, S. T, S. T, H. Y, A. M, E. F, O. M, and D. I. "Significance of High-frequency Electrical Brain Activity." In: *Acta Med. Okayama* 71.3 (2017). DOI: 10.18926/AMO/55201.

[108] D. Gloss, S. J. Nevitt, and R. Staba. *The role of high-frequency oscillations in epilepsy surgery planning.* 2017. DOI: 10.1002/14651858.CD010235.pub3.

[109] M. Wilson. "Network Topology and the Fragility of Tetrahedral Glass-Forming Liquids Temperature dependent structure of liquids, melts and glasses View project Glassy and liquid materials under pressure View project." In: *Artic. Phys. Rev. Lett.* (2009). DOI: 10.1103/PhysRevLett.103.157801.

[110] C. Piccardi and L. Tajoli. "Complexity, centralization, and fragility in economic networks." In: *PLoS One* 13.11 (2018). Ed. by S. Suweis, e0208265. DOI: 10.1371/journal.pone.0208265.

[111] W. Wei, K. Joseph, H. Liu, and K. M. Carley. "The fragility of twitter social networks against suspended users." In: *Proc. 2015 IEEE/ACM Int. Conf. Adv. Soc. Networks Anal. Mining, ASONAM 2015.* Association for Computing Machinery, Inc, 2015, pp. 9–16. DOI: 10.1145/2808797.2809316.

[112] M. A. Kramer and S. S. Cash. *Epilepsy as a disorder of cortical network organization.* 2012. DOI: 10.1177/1073858411422754.

[113] X. Jin, D. A. Prince, and J. R. Huguenard. "Enhanced Excitatory Synaptic Connectivity in Layer V Pyramidal Neurons of Chronically Injured Epileptogenic Neocortex in Rats." In: *J. Neurosci.* 26.18 (2006), pp. 4891–4900. DOI: 10.1523/JNEUROSCI.4361-05.2006.

[114] J. DeFelipe. "Chandelier cells and epilepsy." In: *Brain* 122.10 (1999), pp. 1807–1822. DOI: 10.1093/brain/122.10.1807.

[115] J. R. Hughes. *Gamma, fast, and ultrafast waves of the brain: Their relationships with epilepsy and behavior*. 2008. DOI: 10.1016/j.yebeh.2008.01.011.

[116] B. Elahian, M. Yeasin, B. Mudigoudar, J. W. Wheless, and A. Babajani-Feremi. "Identifying seizure onset zone from electrocorticographic recordings: A machine learning approach based on phase locking value." In: *Seizure* 51 (2017), pp. 35–42. DOI: 10.1016/j.seizure.2017.07.010.

[117] C. N. Heck, D. King-Stephens, A. D. Massey, D. R. Nair, B. C. Jobst, G. L. Barkley, V. Salanova, A. J. Cole, M. C. Smith, R. P. Gwinn, C. Skidmore, P. C. Van Ness, G. K. Bergey, Y. D. Park, I. Miller, E. Geller, P. A. Rutecki, R. Zimmerman, D. C. Spencer, A. Goldman, J. C. Edwards, J. W. Leiphart, R. E. Wharen, J. Fessler, N. B. Fountain, G. A. Worrell, R. E. Gross, S. Eisenschenk, R. B. Duckrow, L. J. Hirsch, C. Bazil, C. A. O'Donovan, F. T. Sun, T. A. Courtney, C. G. Seale, and M. J. Morrell. "Two-year seizure reduction in adults with medically intractable partial onset epilepsy treated with responsive neurostimulation: Final results of the RNS System Pivotal trial." In: *Epilepsia* 55.3 (2014), pp. 432–441. DOI: 10.1111/epi.12534.

[118] M. J. Morrell. "Responsive cortical stimulation for the treatment of medically intractable partial epilepsy." In: *Neurology* 77.13 (2011), pp. 1295–1304. DOI: 10.1212/WNL.0b013e3182302056.

[119] K. J. Meador, R. Kapur, D. W. Loring, A. M. Kanner, and M. J. Morrell. "Quality of life and mood in patients with medically intractable epilepsy treated with targeted responsive neurostimulation." In: *Epilepsy Behav.* 45 (2015), pp. 242–247. DOI: 10.1016/j.yebeh.2015.01.012.

[120] D. King-Stephens, E. Mirro, P. B. Weber, K. D. Laxer, P. C. Van Ness, V. Salanova, D. C. Spencer, C. N. Heck, A. Goldman, B. Jobst, D. C. Shields, G. K. Bergey, S. Eisenschenk, G. A. Worrell, M. A. Rossi, R. E. Gross, A. J. Cole, M. R. Sperling, D. R. Nair, R. P. Gwinn, Y. D. Park, P. A. Rutecki, N. B. Fountain, R. E. Wharen, L. J. Hirsch, I. O. Miller, G. L. Barkley, J. C. Edwards, E. B. Geller, M. J. Berg, T. L. Sadler, F. T. Sun, and M. J. Morrell. "Lateralization of mesial temporal lobe epilepsy with chronic ambulatory electrocorticography." In: *Epilepsia* 56.6 (2015), pp. 959–967. DOI: 10.1111/epi.13010.

[121] D. W. Loring, R. Kapur, K. J. Meador, and M. J. Morrell. "Differential neuropsychological outcomes following targeted responsive neurostimulation for partial-onset epilepsy." In: *Epilepsia* 56.11 (2015), pp. 1836–1844. DOI: 10.1111/epi.13191.

[122] C. T. Anderson, T. K. Tcheng, F. T. Sun, and M. J. Morrell. "Day-Night Patterns of Epileptiform Activity in 65 Patients with Long-Term Ambulatory Electrocorticography." In: *J. Clin. Neurophysiol.* 32.5 (2015), pp. 406–412. DOI: 10.1097/WNP.0000000000000183.

[123] F. T. Sun and M. J. Morrell. "The RNS System: Responsive cortical stimulation for the treatment of refractory partial epilepsy." In: *Expert Rev. Med. Devices* 11.6 (2014), pp. 563–572. DOI: 10.1586/17434440.2014.947274.

[124] G. K. Bergey, M. J. Morrell, E. M. Mizrahi, A. Goldman, D. King-Stephens, D. Nair, S. Srinivasan, B. Jobst, R. E. Gross, D. C. Shields, G. Barkley, V. Salanova, P. Olejniczak, A. Cole, S. S. Cash, K. Noe, R. Wharen, G. Worrell, A. M. Murro, J. Edwards, M. Duchowny, D. Spencer, M. Smith, E. Geller, R. Gwinn, C. Skidmore, S. Eisenschenk, M. Berg, C. Heck, P. Van Ness, N. Fountain, P. Rutecki, A. Massey, C. O'Donovan, D. Labar, R. B. Duckrow, L. J. Hirsch, T. Courtney, F. T. Sun, and C. G. Seale. "Long-term treatment

with responsive brain stimulation in adults with refractory partial seizures." In: *Neurology* 84.8 (2015), pp. 810–817. DOI: 10.1212/WNL.0000000000001280.

[125] J. Ho, T. Tumkaya, S. Aryal, H. Choi, and A. Claridge-Chang. *Moving beyond P values: data analysis with estimation graphics.* 2019. DOI: 10.1038/s41592-019-0470-3.

[126] S. Holm. "A Simple Sequentially Rejective Multiple Test Procedure A Simple Sequentially Rejective Multiple Test Procedure." In: *Source Scand. J. Stat.* 6.2 (1979), pp. 65–70.

[127] P. Perucca, F. Dubeau, and J. Gotman. "Intracranial electroencephalographic seizure-onset patterns: Effect of underlying pathology." In: *Brain* 137.1 (2014), pp. 183–196. DOI: 10.1093/brain/awt299.

[128] A. H. Hawasli, S. K. Bandt, R. E. Hogan, N. Werner, and E. C. Leuthardt. "Laser ablation as treatment strategy for medically refractory dominant insular epilepsy: Therapeutic and functional considerations." In: *Stereotact. Funct. Neurosurg.* 92.6 (2014), pp. 397–404. DOI: 10.1159/000366001.

[129] J. Gonzalez-Martinez and D. Lachhwani. "Stereoelectroencephalography in children with cortical dysplasia: technique and results." In: *Child's Nerv. Syst.* 30.11 (2014), pp. 1853–1857. DOI: 10.1007/s00381-014-2499-z.

[130] J. Engel, M. F. Levesque, and W. D. Shields. *Surgical treatment of the epilepsies: presurgical evaluation.* 1992.

[131] A. M. McIntosh, R. M. Kalnins, L. A. Mitchell, G. C. A. Fabinyi, R. S. Briellmann, and S. F. Berkovic. "Temporal lobectomy: Long-term seizure outcome, late recurrence and risks for seizure recurrence." In: *Brain* 127.9 (2004), pp. 2018–2030. DOI: 10.1093/brain/awh221.

[132] L. E. Jeha, I. M. Najm, W. E. Bingaman, F. Khandwala, P. Widdess-Walsh, H. H. Morris, D. S. Dinner, D. Nair, N. Foldvary-Schaeffer, R. A. Prayson, Y. Comair, R. O'Brien, J. Bulacio, A. Gupta, and H. O. Lüders. "Predictors of outcome after temporal lobectomy for the treatment of intractable epilepsy." In: *Neurology* 66.12 (2006), pp. 1938–1940. DOI: 10.1212/01.wnl.0000219810.71010.9b.

[133] H. Urbach, J. Hattingen, J. Von Oertzen, C. Luyken, H. Clusmann, T. Kral, M. Kurthen, J. Schramm, I. Blümcke, and H. H. Schild. "MR imaging in the presurgical workup of patients with drug-resistant epilepsy." In: *Am. J. Neuroradiol.* 25.6 (2004), pp. 919–926. DOI: 25/6/919[pii]ET-2004/06/19.

[134] B. Kemp, A. Värri, A. C. Rosa, K. D. Nielsen, and J. Gade. "A simple format for exchange of digitized polygraphic recordings." In: *Electroencephalogr. Clin. Neurophysiol.* 82.5 (1992), pp. 391–393. DOI: 10.1016/0013-4694(92)90009-7.

[135] C. Holdgraf, S. Appelhoff, S. Bickel, K. Bouchard, S. D'Ambrosio, O. David, O. Devinsky, B. Dichter, A. Flinker, B. L. Foster, K. J. Gorgolewski, I. Groen, D. Groppe, A. Gunduz, L. Hamilton, C. J. Honey, M. Jas, R. Knight, J. P. Lachaux, J. C. Lau, C. Lee-Messer, B. N. Lundstrom, K. J. Miller, J. G. Ojemann, R. Oostenveld, N. Petridou, G. Piantoni, A. Pigorini, N. Pouratian, N. F. Ramsey, A. Stolk, N. C. Swann, F. Tadel, B. Voytek, B. A. Wandell, J. Winawer, K. Whitaker, L. Zehl, and D. Hermes. "iEEG-BIDS, extending the Brain Imaging Data Structure specification to human intracranial electrophysiology." In: *Sci. data* 6.1 (2019), p. 102. DOI: 10.1038/s41597-019-0105-7.

[136] K. J. Gorgolewski, T. Auer, V. D. Calhoun, R. C. Craddock, S. Das, E. P. Duff, G. Flandin, S. S. Ghosh, T. Glatard, Y. O. Halchenko, D. A. Handwerker, M. Hanke, D. Keator, X. Li, Z. Michael, C. Maumet, B. N. Nichols, T. E. Nichols, J. Pellman, J. B. Poline, A. Rokem, G. Schaefer, V. Sochat, W. Triplett, J. A. Turner, G. Varoquaux, and R. A. Poldrack. "The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments." In: *Sci. Data* 3.1 (2016), p. 160044. DOI: 10.1038/sdata.2016.44.

[137] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, R. Goj, M. Jas, T. Brooks, L. Parkkonen, and M. Hämäläinen. "MEG and EEG data analysis with MNE-Python." In: *Front. Neurosci.* 7 (2013), p. 267. DOI: 10.3389/fnins.2013.00267.

[138] A. Gramfort, M. Luessi, E. Larson, D. A. Engemann, D. Strohmeier, C. Brodbeck, L. Parkkonen, and M. S. Hämäläinen. "MNE software for processing MEG and EEG data." In: *Neuroimage* 86 (2014), pp. 446–460.

[139] S. Appelhoff, M. Sanderson, T. L. Brooks, M. van Vliet, R. Quentin, C. Holdgraf, M. Chaumon, E. Mikulan, K. Tavabi, R. Höchenberger, D. Welke, C. Brunner, A. P. Rockhill, E. Larson, S. K. Herbst, R. Luke, A. Li, A. Gramfort, and M. Jas. "MNE-BIDS: MNE-Python + BIDS = easy dataset interaction." In: *Organ. Hum. Brain Mapp.* 2020. DOI: 10.5281/ZENODO.3890935.

[140] S. Van Der Walt, S. C. Colbert, and G. Varoquaux. "The NumPy array: A structure for efficient numerical computation." In: *Comput. Sci. Eng.* 13.2 (2011), pp. 22–30. DOI: 10.1109/MCSE.2011.37.

[141] P. Virtanen et al. "SciPy 1.0: fundamental algorithms for scientific computing in Python." In: *Nat. Methods* 17.3 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[142] J. D. Hunter. "Matplotlib: A 2D graphics environment." In: *Comput. Sci. Eng.* 9.3 (2007), pp. 99–104. DOI: 10.1109/MCSE.2007.55.

[143] M. W. O. B. D. O. P. H. S. L. D. C. G. T. A. Y. H. J. B. C. J. W. J. d. R. C. P. S. H. J. V. S. V. G. K. Qalieh. *mwaskom/seaborn: v0.8.1 (September 2017).* 2017. DOI: 10.5281/zenodo.883859.

[144] R. Vallat. "Pingouin: statistics in Python." In: *J. Open Source Softw.* 3.31 (2018), p. 1026. DOI: 10.21105/joss.01026.

[145] S. Raschka. "MLxtend: Providing machine learning and data science utilities and extensions to Python's scientific computing stack." In: *J. Open Source Softw.* 3.24 (2018), p. 638. DOI: 10.21105/joss.00638.

[146] F. Pedregosa, V. Michel, M. Blondel, P. Prettenhofer, R. Weiss, J. Vanderplas, D. Cournapeau, G. Varoquaux, A. Gramfort, B. Thirion, O. Grisel, V. Dubourg, A. Passos, M. Brucher, M. Perrot, and É. Duchesnay. "Scikit-learn: Machine Learning in Python." In: *J. Mach. Learn. Res.* 12.85 (2011), pp. 2825–2830.

[147] J. Gonzalez-Martinez and I. M. Najm. "Indications and selection criteria for invasive monitoring in children with cortical dysplasia." In: *Child's Nerv. Syst.* 30.11 (2014), pp. 1823–1829. DOI: 10.1007/s00381-014-2497-1.

[148] E. C. Lewis, A. G. Weil, M. Duchowny, S. Bhatia, J. Ragheb, and I. Miller. "MR-guided laser interstitial thermal therapy for pediatric drug-resistant lesional epilepsy." In: *Epilepsia* 56.10 (2015), pp. 1590–1598. DOI: 10.1111/epi.13106.

[149] L. Ross, A. M. Naduvil, J. C. Bulacio, I. M. Najm, and J. A. Gonzalez-Martinez. "Stereoelectroencephalography-Guided Laser Ablations in Patients With Neocortical Pharmacoresistant Focal Epilepsy: Concept and Operative Technique." In: *Oper. Neurosurg.* (2018). DOI: 10.1093/ons/opy022.

[150] J. Chen, Z. Huang, L. Li, L. Ren, and Y. Wang. "Histological type of focal cortical dysplasia is associated with the risk of postsurgical seizure in children and adolescents." In: *Ther. Clin. Risk Manag.* 15 (2019), pp. 877–884. DOI: 10.2147/TCRM.S203039.

[151] P. Krsek, T. Pieper, A. Karlmeier, M. Hildebrandt, D. Kolodziejczyk, P. Winkler, E. Pauli, I. Blümcke, and H. Holthausen. "Different presurgical characteristics and seizure outcomes in children with focal cortical dysplasia type I or II." In: *Epilepsia* 50.1 (2009), pp. 125–137. DOI: 10.1111/j.1528-1167.2008.01682.x.

[152] U. Heinemann, A. Konnerth, R. Pumain, and W. J. Wadman. "Extracellular calcium and potassium concentration changes in chronic epileptic brain tissue." In: *Adv. Neurol.* 44 (1986), pp. 641–61.

[153] K. A. Ludwig, R. M. Miriani, N. B. Langhals, M. D. Joseph, D. J. Anderson, and D. R. Kipke. "Using a Common Average Reference to Improve Cortical Neuron Recordings From Microelectrode Arrays." In: *J. Neurophysiol.* 101.3 (2009), pp. 1679–1689. DOI: 10.1152/jn.90989.2008.

[154] P. D. Welch. "The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms." In: *IEEE Trans. Audio Electroacoust.* 15.2 (1967), pp. 70–73. DOI: 10.1109/TAU.1967.1161901.

[155] M. Skocik, J. Collins, C. Callahan-Flintoft, H. Bowman, and B. Wyble. "I Tried a Bunch of Things: The Dangers of Unexpected Overfitting in Classification." In: *bioRxiv* (2016), p. 078816. DOI: 10.1101/078816.

[156] Y. Qin and C. E. Priebe. "Robust hypothesis testing via Lq-likelihood." In: *Stat. Sin.* 27.4 (2017), pp. 1793–1813. DOI: 10.5705/ss.202015.0441.

[157] A. Alyakin, Y. Qin, and C. E. Priebe. "LqRT: Robust Hypothesis Testing of Location Parameters using Lq-Likelihood-Ratio-Type Test in Python." In: *Arxiv* (2019).

[158] D. Lakens. "Calculating and reporting effect sizes to facilitate cumulative science: A practical primer for t-tests and ANOVAs." In: *Front. Psychol.* 4.NOV (2013). DOI: 10.3389/fpsyg.2013.00863.

[159] A. Niculescu-Mizil and R. Caruana. "Predicting good probabilities with supervised learning." In: *ICML 2005 - Proc. 22nd Int. Conf. Mach. Learn.* New York, New York, USA: ACM Press, 2005, pp. 625–632. DOI: 10.1145/1102351.1102430.