2021

Quantitative Methods For Guiding Epilepsy Surgery From Intracranial Eeg

John Bernabei
University of Pennsylvania

Quantitative Methods For Guiding Epilepsy Surgery From Intracranial Eeg

Abstract

Despite advances in intracranial EEG (iEEG) technique, technology and neuroimaging, patients today are no more likely to achieve seizure freedom after epilepsy surgery than they were 20 years ago. These poor outcomes are in part due to the difficulty and subjectivity associated with interpreting iEEG recordings, and have led to widespread interest in developing quantitative methods to localize the epileptogenic zone. Approaches to computational iEEG analysis vary widely, spanning studies of both seizures and interictal periods, and encompassing a range of techniques including electrographic signal analysis and graph theory. However, many current methods often fail to generalize to new data and are sensitive to differences in pathology and electrode placement. Indeed, none have completed prospective clinical trials. In this dissertation, I develop and validate tools for guiding epilepsy surgery through the quantitative analysis of intracranial EEG. Specifically, I leverage methods from graph theory for mapping network synchronizability to predict surgical outcome from ictal recordings, and also investigate the effects of sampling bias on network models. Finally, I construct a normative intracranial EEG atlas as a framework for objectively identifying patterns of abnormal neural activity and connectivity. Overall, the methods and results of this dissertation support the implementation of quantitative iEEG analysis in epilepsy surgical evaluation.

Degree Type

Dissertation

Degree Name

Doctor of Philosophy (PhD)

Graduate Group

Bioengineering

First Advisor

Brian Litt

Keywords

Epilepsy, Neuroengineering, Neurosurgery

Subject Categories

Biomedical | Neuroscience and Neurobiology

QUANTITATIVE METHODS FOR GUIDING EPILEPSY SURGERY FROM INTRACRANIAL EEG

John M. Bernabei

A DISSERTATION

in

Bioengineering

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2021

Supervisor of Dissertation

_____

Brian Litt, M.D.
Professor of Bioengineering, Neurology, and Neurosurgery

Graduate Group Chairperson

_____

Yale E. Cohen, Ph.D.
Professor of Otorhinolaryngology and Bioengineering

Dissertation Committee

Danielle S. Bassett, Ph.D., Professor of Bioengineering, Neuroscience, Physics & Astronomy, Psychiatry, Electrical & Systems Engineering

Timothy H. Lucas, M.D., Ph.D., Associate Professor of Neurosurgery

Russell T. Shinohara, Ph.D., Associate Professor of Biostatistics

QUANTITATIVE METHODS FOR GUIDING EPILEPSY SURGERY FROM INTRACRANIAL EEG

ACKNOWLEDGMENT

This work contained in this thesis would not have been possible without extensive support both inside and outside of the lab during the last few years.

First and foremost, I would like to thank Dr. Brian Litt for welcoming me into his lab and mentoring me throughout my time at Penn. Brian created an unparalleled environment to learn and grow as a scientist: he gave me the guidance to understand clinically important problems, the freedom to pursue my interests, and modeled an incredible academic career. Upon meeting him shortly after starting the MD/PhD program, I quickly appreciated his genuine passion for bettering the lives of patients through clinical care and translational research, as well as his commitment to bringing out the best in his trainees. Brian has taught me innumerable lessons that will shape my future career as a physician-scientist, and for that I am immensely grateful.

Some of my closest collaborators are essential parts of the work presented here. In particular, I'd like to thank the members of my thesis committee: Dr. Timothy Lucas, Dr. Danielle Bassett, and Dr. Russell Shinohara. Their mentorship has been essential in helping me become the best scientist possible and has significantly raised the level of my work, while their unique perspectives and expertise have shaped the studies contained in this thesis. I'd also like to thank Dr. Lohith Kini, Dr. Erin Conrad, and Dr. Nishant Sinha for their extended assistance. Lohith set much of the groundwork for brain network research in the lab, and we completed the work contained in Chapter 3 of this thesis together. Erin's expertise as an epileptologist and scientist have been invaluable, and we completed work

on sampling bias in network models which is featured in Chapter 4. Nishant has a wealth of experience and ideas, and has provided critical support on the work in Chapter 6.

I'd like to thank the past and present the members of the CNT who have been part of my journey. Carolyn Wilkinson and Everett Prince have done so much to keep the lab running smoothly, and I sincerely thank them for all of their work. I'd also like to thank everyone who has worked on regulatory and data support, including Jacqueline Boccanfuso, Magda Wernovsky, Adam Gibson, Amanda Samuel, and Ryan Archer. Without their efforts we would never have the building blocks with which to do our work. I also sincerely thank Dr. Kathryn Davis for her support as one of the other PIs in the CNT, providing valuable advice, data, and mentorship across many projects. I'd like to thank the many previous and current students in the laboratory: Steven Baldassano, Preya Shah, Brittany Scheid, Andy Revell, Akash Pattnaik, Peter Galer, Nina Ghosn, Kevin Xie, Ian Ong, Arjun Shankar, Daniel Kim, and Chris Painter.

I'd also like to thank Dr. Skip Brass, Maggie Krall, and the many other people who run the MD/PhD program. The environment they have created here is second-to-none, and I am truly thankful for the opportunity to train here at The Best MSTP in the Galaxy™.

I'd like to thank my closest friends here in Philadelphia who have shared this journey with me: Daniel Xu, Naveen Jain, Karun Kiani, Daniel Park, Ivan Kuznetsov, Campbell Arnold, Jason Kim, Daniel Connolly, and Courteney Ly. Finally, none of this would be possible without the extensive support of my family. Thank you for sacrificing to provide me with the best possible education, and for teaching me to value scholarship, knowledge, and curiosity. Without you I would not have been able to take this path.

ABSTRACT

QUANTITATIVE METHODS FOR GUIDING EPILEPSY SURGERY FROM INTRACRANIAL EEG

John M. Bernabei

Brian Litt

Despite advances in intracranial EEG (iEEG) technique, technology and neuroimaging, patients today are no more likely to achieve seizure freedom after epilepsy surgery than they were 20 years ago. These poor outcomes are in part due to the difficulty and subjectivity associated with interpreting iEEG recordings, and have led to widespread interest in developing quantitative methods to localize the epileptogenic zone. Approaches to computational iEEG analysis vary widely, spanning studies of both seizures and interictal periods, and encompassing a range of techniques including electrographic signal analysis and graph theory. However, many current methods often fail to generalize to new data and are sensitive to differences in pathology and electrode placement. Indeed, none have completed prospective clinical trials. In this dissertation, I develop and validate tools for guiding epilepsy surgery through the quantitative analysis of intracranial EEG. Specifically, I leverage methods from graph theory for mapping network synchronizability to predict surgical outcome from ictal recordings, and also investigate the effects of sampling bias on network models. Finally, I construct a normative intracranial EEG atlas as a framework for objectively identifying patterns of abnormal neural activity and connectivity. Overall, the methods and results of this dissertation support the implementation of quantitative iEEG analysis in epilepsy surgical evaluation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: Introduction

Epilepsy is the most common serious neurologic disorder, affecting over 60 million individuals worldwide, and is characterized by the occurrence of seizures and broad neurologic dysfunction which contribute to disability, morbidity, and mortality[1]. Over 1/3 of all people living with epilepsy have seizures which are resistant to antiepileptic medication, and thus may be candidates for epilepsy surgery to give them a chance at seizure freedom[1]. However, epilepsy surgery is often plagued by poor outcomes both due to seizure relapse, as well as neurologic, cognitive, and psychiatric issues which arise from the resection of brain in the pursuit of seizure freedom[2]. One of the main contributors to seizure relapse is the difficulty in selecting surgical candidates and anatomic targets for surgery. Intracranial EEG (iEEG) in which direct seizure recordings are made from implanted electrodes is one of the primary ways epilepsy surgery is guided[3]. However, its interpretation is difficult and subjective. In this thesis, I aim to develop and validate methods for guiding epilepsy surgery through the quantitative analysis of iEEG. After providing a background on previous quantitative iEEG methods in **Chapter 2**, I organize my contributions to this field in the following sections:

Section 1: Map epileptic networks through analysis of pre-ictal and ictal recordings

In the first section, comprising **Chapter 3**, I validate a method known as virtual resection[4] for mapping the synchronizing nature of regions across the brain. Virtual resection leverages the use of a metric from graph theory known as synchronizability, which quantifies the ease with which oscillations can propagate through a network of coupled oscillators[5,6]. I compute functional network connectivity during pre-ictal and ictal

epochs in 28 patients with drug-resistant epilepsy implanted with subdural grid, strip, and depth electrodes. I show that time-varying synchronizability can accurately predict surgical outcome before resection is performed. We also calculate the contribution of each node to network synchronizability through simulating its removal, yielding a value known as control centrality. Finally, I illustrate the potential use of synchronizability and control centrality as a clinical tools that could help select optimal surgical candidates and target resection to the parts of the brain that are most likely to cause seizures.

Section 2: Understanding the effects of sampling bias on network models

In the second section, I assess the different ways in which spatial sampling bias affects network models. Intracranial EEG electrode implants vary widely across patients and centers, with a variety of channel numbers, electrode types, and electrode locations. However, it is unknown how these differences in sampling propagate into network models and whether some approaches are only suited for certain types of implantation strategies. In **Chapter 4**, I first determine how different network metrics from graph theory have distinct ways in which they are affected by electrode placement and number of electrodes. In **Chapter 5**, I then determine how the two main methods of intracranial EEG, ECoG in which there are subdural electrodes and SEEG in which there are depth electrodes alone, have distinct sampling biases which affect the use of network models in epilepsy.

Section 3: Mapping epileptic networks through interictal analysis

In the final section, consisting of **Chapter 6**, I construct an atlas of interictal intracranial EEG data. While approaches in neuroimaging benefit from the ability to scan healthy controls to map differences between normal and abnormal brains, exceedingly few

patients without epilepsy receive intracranial EEG. Furthermore, the confounds of normal neural activity and connectivity present a challenge for traditional forms of interictal iEEG analysis. Thus, I aggregate data across >160 patients with intracranial EEG and carefully separate channels which are normal from those which can generate spikes and seizures, and map normal activity and connectivity throughout the brain. Comparing features of activity and connectivity in held-out, test patients to those of the normative atlas identifies abnormalities in neural activity and connectivity associated with patient-level brain regions which are likely to generate spikes and seizures. I validate the clinical utility of this normative atlas mapping approach and illustrate how it can be used to guide epilepsy surgery from intracranial EEG.

Overall, this thesis represents a broad series of studies which address major challenges in quantitative iEEG analysis for epilepsy surgery, however there is still much more work to be done. In **Chapter 7**, I conclude and highlight areas for additional research which may lay the groundwork for clinical trials in the near future. Ultimately, it is my hope that this research may soon support the use of quantitative iEEG models in patient care, improving accessibility and surgical outcomes for drug resistant epilepsy.

CHAPTER 2: Background

Patients with drug-resistant epilepsy often do not achieve seizure freedom after epilepsy surgery[2,7]. When no brain lesion is evident on MRI imaging, the rate of surgical success is lower still[2]. Resections can be generous, with a significant amount of healthy tissue removed, frequently resulting in post-surgical side effects, including neuropsychological deficits and reduced quality of life[8]. Even when intracranial EEG (iEEG)[9] and brain mapping are performed with the goal of localizing seizures and avoiding eloquent cortex, it is often unclear where to intervene to optimize outcome while minimizing tissue destruction and its impact on quality of life[8,10]. The chief purpose of iEEG is to localize the 'epileptogenic zone' (EZ), which is the minimal amount of cortex that must be removed to produce seizure freedom[11]. This theoretical brain area (there could be more than one) may overlap with the 'irritative zone', the region(s) responsible for generating interictal spikes, as well as the seizure onset zone (SOZ) - the part of the brain where clinical seizures are observed to originate[12]. While there is a large body of research exploring whether mapping the irritative zone through localizing interictal spikes can guide surgery[13–18], current clinical practice primarily aims to define the epileptogenic zone, the region that must be disabled to stop seizures, by proxy, through identifying all potential seizure onset zones. However, given the spatial sparsity of iEEG sampling, a limited 2-3 week period for iEEG monitoring, and the ambiguity of manually interpreting iEEG, parts of the EZ could go undetected and left out of the surgical plan. Thus it is critically important to develop better, quantitative ways of estimating the boundaries and topology of the epileptogenic zone.

Epileptologists normally evaluate iEEG through an interpretation of temporal, spatial, and spectral aspects of the signal (Figure 2.1A). The primary method of SOZ identification is the qualitative recognition of specific seizure onset patterns that are known to indicate a well-localized onset[19,20]. Common patterns include (i) low voltage fast activity (LVFA) in which low-amplitude activity in the beta to low gamma frequency range begins in a localized area before the propagation of seizure activity elsewhere, (ii) 'DC shift' or 'diffuse electrodecremental event' which is characterized by a slow shift in baseline voltage and is often followed by LVFA[19], (iii) preictal rhythmic spiking of low frequency and high amplitude, and (iv) bursts of polyspikes or spike-and-wave activity. However, seizures in some patients might not display any of these patterns, and conversely, the presence of these patterns in some channels does not mean that other brain regions are not epileptogenic. Unfortunately these patterns are frequently appreciated by their absence in specific patients, replaced by poorly localized, lower frequency or "propagated" patterns, leaving clinicians to wonder if the implant has somehow missed the region driving seizures, or if the network is "diffuse," and better treated with broader neuromodulation than focal intervention. These limitations of qualitative iEEG analysis often make it difficult for epileptologists and neurosurgeons to create a surgical plan and can lead to poor patient outcomes. Many groups therefore hypothesize that quantitative methods of mapping the EZ that extend beyond visually identifying seizure onset patterns might be able to provide better surgical targets or offer complementary information that is currently missed.

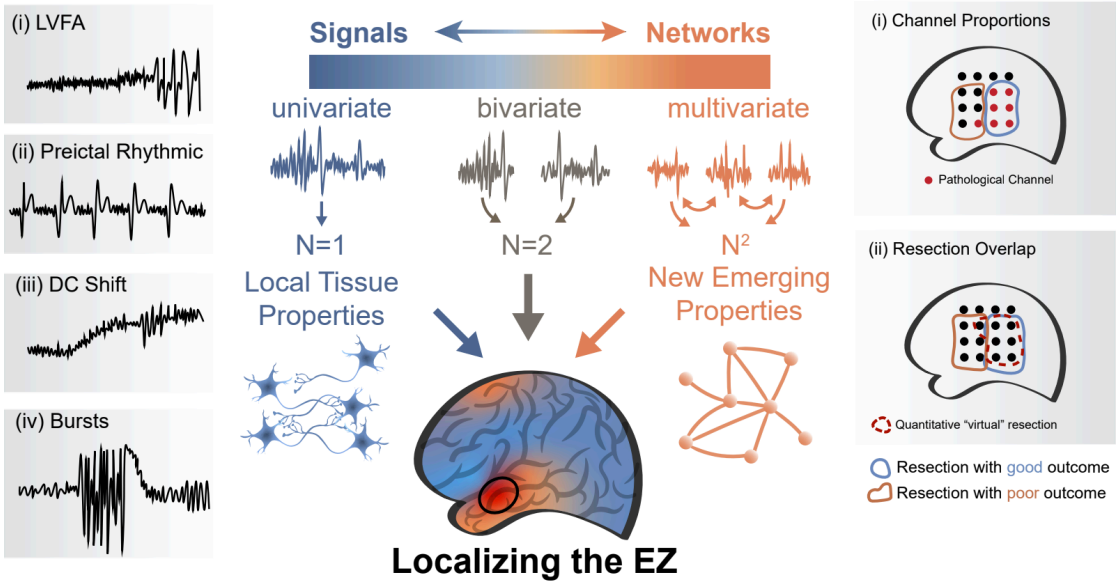


Figure 2.1. Overview of localization of the epileptogenic zone. A) Currently, the epileptogenic zone is localized from intracranial EEG by evaluation of electrographic patterns such as low voltage fast activity, DC shift, rhythmic waves, bursts, and spikes. B) Quantitative methods of ECoG analysis range can range from univariate signal processing analyses of individual channels, to bivariate analyses on pairs of channels, and to networks in which connectivity between all pairs of channels are assessed and emergent network properties are studied. C) In studies of quantitative localization methods, the value of metrics within the resection zone are compared to those values outside of the resection zone, and the result is correlated to surgical outcome.

The past two decades demonstrate a growing interest in mapping epileptogenic zones using quantitative techniques (Figure 2.1B). To date, many of these models follow a similar paradigm (Figure 2.1C). Ictal or interictal iEEG from a retrospective cohort of patients is used to calculate the quantitative metric of choice for each channel (referential or bipolar, corresponding to one or two electrode contacts, respectively) over a specified time window. When the value of this proposed localizing metric exceeds a certain threshold, the sampled brain region is deemed pathologic. Finally, one or two hypotheses are tested: (i) Good outcome patients have a greater proportion of 'pathologic' channels resected than poor outcome patients, and (ii) a potential resection zone defined by resecting

'pathologic' channels has better overlap with the true resection zone in good outcome patients compared to poor outcome patients. Beyond this general paradigm, the quantitative methods for identifying the EZ vary greatly, ranging from linear signal analysis to networked dynamical systems. It is my goal here to review the scope of available quantitative iEEG methods for localizing the EZ ranging from simple to complex, describe the retrospective literature supporting the utility iEEG methods, and discuss the path forward to clinical translation for the benefit of future patients with drug resistant epilepsy, so that surgical outcomes may be improved.

It is important to note that the above paradigm has suffered from uncertainty in electrode localization and delineation of the region of brain resected, ablated or modulated by therapy. Early studies localized electrodes using estimation by surgeons and neurologists, and resections were crudely measured during the surgery. In more recent years, robot-guided implants[21] and improved methods for coregistering electrodes using pre and postoperative imaging have dramatically improved precision in this area, though deformation of the brain during these often lengthy procedures may degrade actual electrode placement. Similarly, automated methods for segmenting post-resection images are increasing the accuracy of delineating resected or ablated brain regions, but there is a growing appreciation that both local and remote effects of these interventions[22], in part related to surgical approach and technique, may vary significantly between centers and individuals even for the same procedure. Such variability, without quantitation, makes challenges studies of the effects of different iEEG localizing techniques and their effect on outcome.

2.1 Quantitative electrographic analysis

The transition from analog to digital EEG in the 1980s and 1990s stimulated researchers to develop quantitative methods analyzing iEEG[23]. Many of the early studies focused on signal morphology and frequency content to characterize spiking activity[24] or quantify seizure onset patterns[25] in their interpretation of iEEG. A dramatic increase in computational power and digital storage rapidly emerged, followed by the steady creation of analysis software, toolboxes for EEG and digital signal processing, and more recently dynamic, open source packages on multiple platforms for signal processing, nonlinear dynamics, machine learning, and deep learning. Consequently the field has exploded, leveraging a wide variety of techniques ranging from analyzing signals from each electrode contact individually with univariate measures to bivariate metrics that relate signals from pairs of electrode contacts. While the plurality of methods cannot be covered here, I describe major divisions in the field and highlight the techniques which have generated the greatest amount of interest and which I believe carry the most potential clinical utility.

2.1.1 Univariate measures

The simplest univariate approaches to quantitative iEEG evaluation use univariate linear signal analysis, which comprise a variety of common techniques in the time and frequency domains (Figure 2.2). For example, the time domain metric of absolute value of the signal slope, also known as line length[26], can correctly identify electrode contacts which were surgically targeted in good outcome patients[27]. In the frequency domain, calculating power within physiological frequencies[28] of the delta, theta, alpha, beta, and gamma bands provides a simple and clinically interpretable manner of quantifying the activity at each

node. For example, interictal delta bandpower from iEEG lateralizes the pathologic hippocampus in mesial temporal lobe epilepsy with good accuracy[29], and localized activity within 20-80 Hz portends good clinical outcome[25]. Beyond calculating the bandpower of canonical frequencies, more complicated methods such as the wavelet transform can transform EEG from each channel into time-frequency plots in which much more granular detail about the activity at each frequency is readily available. Grinenko et al. used a wavelet transform of the SEEG signals recorded on each individual channel in patients that became seizure free to identify a 'fingerprint of the epileptogenic zone'[30,31]. They observed that the EZ is characterized by pre-ictal spikes, multiband fast activity, and suppression of low frequencies at onset, and used approaches from computer vision to extract features from the time-frequency plot which were fed into a support vector machine classifier[30]. This approach yielded a positive predictive value of >90% for determining which electrodes fell within the resection zone in good outcome patients and could be a promising technique in future surgical planning. Overall, the use of linear signal analysis can provide simple and interpretable quantifications of iEEG signal properties.

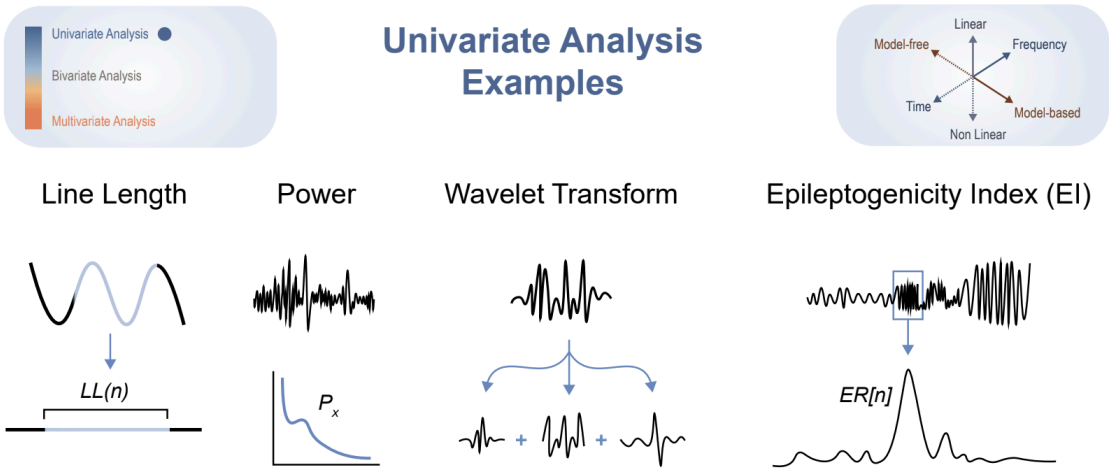


Figure 2.2. Univariate studies. A) Linear univariate metrics include those in the (i) time and (ii) frequency domains. B) Non-linear metrics, such as the Epileptogenicity Index, can identify epileptogenic tissue.

More complicated, non-linear univariate metrics that have seen substantial study include the epileptogenicity index (EI)[32] and epileptogenicity maps (EM)[33], which attempts to quantify clinically-observable patterns based on both the spectral and temporal delay patterns of iEEG. To calculate the EI, two specific metrics are calculated over a sliding window: (i) the 'ER' ratio of high frequency activity (beta and gamma bands) to low frequency activity (theta and alpha bands) and (ii) a cumulative sum algorithm used over the 'ER' signal to determine when it significantly changes, marking a shift from low frequency to high frequency activity. Thus, a channel which is involved early in the seizure will have a high EI value. EI was initially studied in mesial temporal lobe epilepsy sampled by SEEG[32] and is among the best characterized quantitative iEEG analysis methods with validation in epilepsy localized to cavernous angiomas[34], motor systems[35], and occipital networks[36]. In all cases it has yielded meaningful localization accuracy and could serve as a powerful addition to the current paradigm for surgical planning.

2.1.2 Bivariate Measures

Bivariate iEEG analyses determine the statistical relationships between pairs of channels and can be analyzed either in isolation or can be employed to construct network adjacency matrices, where the i,j^{th} matrix entry holds the metric value calculated using channels *i* and *j*. Approaches to bivariate analysis span linear techniques in the time domain and frequency domain to nonlinear methods adapted from information theory. As examples, I review correlation, coherence, and phase locking, which are among the most commonly used metrics for both network and studies and bivariate electrographic analysis (Figure 2.3).

Perhaps the most prevalent bivariate metric for iEEG analysis is Pearson cross-correlation, which quantifies the linear relationship between two signals in the time domain for a given time lag. Spatial changes in seizure activity can be quantified by mapping how the correlation coefficient between each pair of channels evolves over time in pre-ictal and ictal epochs. For example, decorrelation across channels at seizure onset is evident in epileptic fast iEEG activity[37]. The correlation coefficient can also distinguish patients who become seizure free from those who do not, as an abundance of small, homogeneous correlations between channels in temporal lobe epilepsy (TLE) portends good outcome whereas large, heterogeneous correlations do not[38]. In sum, cross-correlation is a foundational method for assessing connectivity and, independent from network analysis, may have significant explanatory value for how the activity across brain regions can be linked.

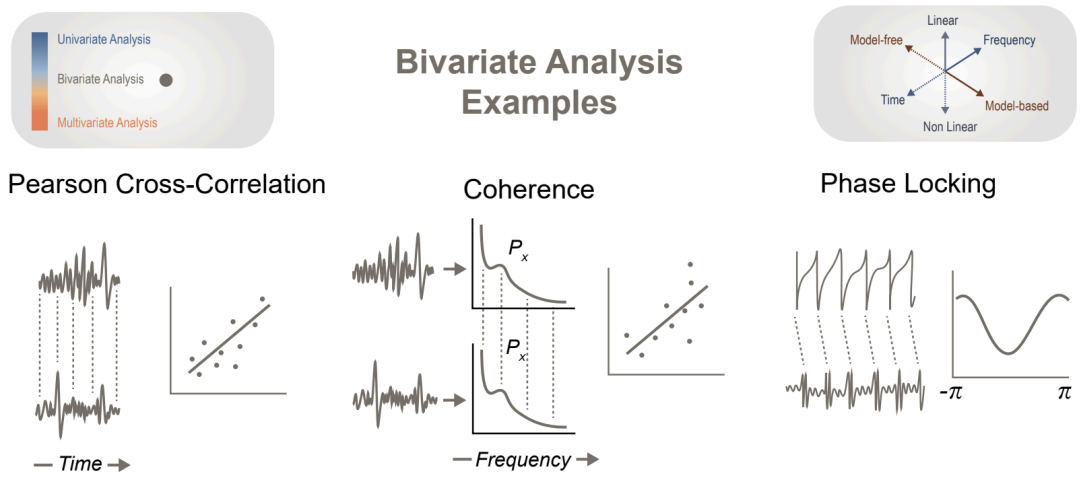


Figure 2.3. Bivariate studies. Electrographic analysis that takes data from pairs of electrode contacts to gauge functional connectivity between regions include metrics such as Pearson cross-correlation, coherence, phase-locking, and the directed transfer function. Pearson cross-correlation functions in the time domain, coherence in the frequency domain.

The frequency domain version of correlation is known as coherence, which quantifies the linear relationship of the power spectra of two signals in the frequency domain. Comparing ictal coherence between pairs of electrode contacts in the medial temporal lobe versus pairs in the lateral temporal lobe can accurately distinguish between mesial and lateral temporal lobe epilepsy[39], which can often be a difficult classification to determine. Similarly, a study of partially directed coherence in patients with focal cortical dysplasia found different connectivity patterns between cortical regions beyond the dysplasia that can generate seizures, which could serve as an important method to find less obvious EZ locations[40]. Finally, high ictal coherence in frequency bands exhibiting high frequency oscillations (HFOs) has been shown to be an EZ biomarker and tends to spatially focus in clusters of channels associated with seizure onset regions[41]. The ability for coherence to selectively assess activity in specific frequency bands may provide significant

value towards increasing our understanding of seizure dynamics as well as identifying the modes of neural synchrony which are most pathogenic.

A third commonly used bivariate metric is phase locking value (PLV), which describes the absolute value of the phase difference between two signals, and has been frequently used to search for epileptogenic zones. The driving factor behind this statistic is that areas of the brain capable of generating seizures may induce changes in long range synchronization that are visible by examining phase. High-gamma band phase locking to low frequency channels is a putative feature of the ictal core and when used as a method of EZ localization can accurately predict surgical outcome, with an area under the receiver-operating characteristic curve of 0.79, versus 0.68 for using the SOZ alone[42]. Phase-locking from high gamma to low frequencies was used in a logistic regression model to predict the EZ[43]. 96% of the channels identified using this approach were resected in good outcome patients, while over 31% of the identified channels fell outside of the resection zone in poor outcome patients. With multiple successful retrospective studies, phase-locked high gamma may be a good candidate for prospective studies in localizing the EZ.

Finally, directed connectivity metrics have also provided essential insight into the relationship of the EZ and the surrounding brain and could serve as important localizing tools. In particular, partial directed coherence (PDC) and the directed transfer function (DTF) have revealed that the EZ has unexpectedly high inward connectivity compared to other parts of the brain, however outwards strengths are not different[44]. Using these measures as inputs to a classifier for determining epileptogenic regions yielded an area under the curve of 0.88, and increased to 0.93 when restricted to the EZ of only good

outcome patients. The finding of enhanced inwards connectivity could be possibly due to an 'inhibitory surround' which could be attempting to suppress epileptiform activity[45]. In the future, it is likely that directed connectivity measures could be used in surgical evaluation for their simplicity, interpretability, and localizing value.

2.2 Network methods

There is an increasing body of evidence supporting the hypothesis that epilepsy is a disease of disordered brain networks[46–48]. Generalizing the basis of bivariate electrographic signal analysis to measure functional connectivity leads us to the essentials of network neuroscience[49]. Using statistical metrics we can define network 'nodes' as brain regions recorded by single or groups of iEEG electrodes and 'edges' as the relationship between nodes[50]. Any bivariate metric including correlation, coherence, mutual information, and phase locking value can be used to calculate edge weights and yield a functional connectivity (FC) matrix. Networks can be calculated during interictal or ictal epochs, and EEG can be windowed before network calculation to yield multilayer networks which can quantify network dynamics over time. These principles give rise to a wide variety of EZ localization methods, some which study patterns of connectivity in a model-free fashion, and others which use neural models to overlay dynamics on top of connectivity (Figure 2.4).

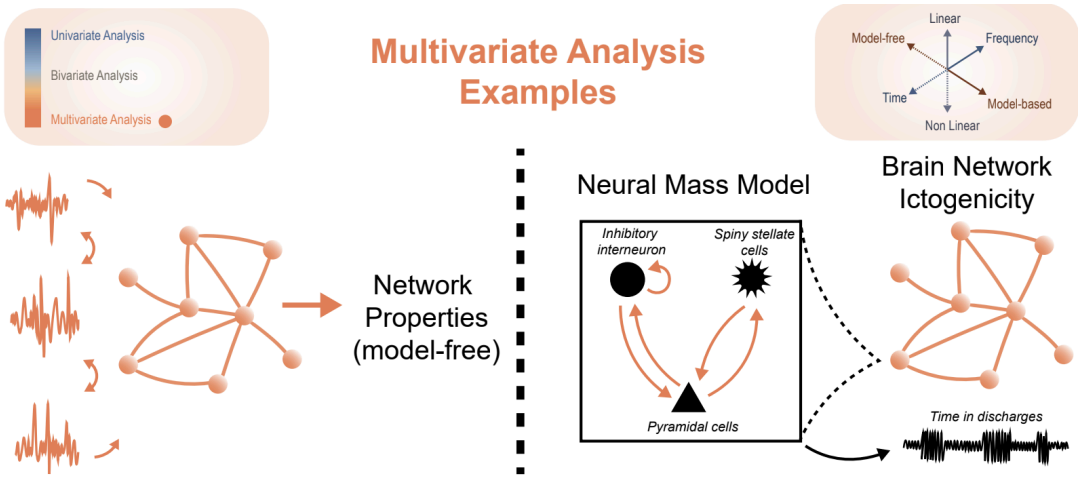


Figure 2.4. Network studies. Networks can be constructed from bivariate metrics to quantify the connectivity between all possible pairs of electrodes, and these adjacency matrices can be used towards graph studies or other models. A) Model-free network studies can use nodal metrics like node strength, betweenness centrality, and eigenvector centrality. They can also use global metrics like efficiency, path length, and synchronizability. B) Model-based methods can use connectivity to parametrize dynamical systems. In many studies, the dynamical systems generate surrogate EEG which can be assessed for its probability of being in an 'ictal-like' state, such as in Goodfellow et al and Sinha et al.

2.2.1 Model-free approaches

The simplest network neuroscience paradigms for quantitative localization of the EZ use 'model free' approaches which apply traditional statistics that describe different aspects of network structure. The bulk of these metrics were developed outside of the framework of neuroscience and have been well-studied and characterized in the fields of physical, social, and information networks. The simplest model-free network method for EZ localization is using a metric known as 'degree' for binary networks or 'node strength' for weighted networks, which is the sum of all a given node's connections. Node strength is a simple yet powerful metric for surgical targeting, as the EZ has high interictal intra-zone connectivity and connectivity to outlying brain in both SEEG[51] and ECoG[52]. These

significant findings warrant further study, as the ability to define the boundaries of the EZ using simple interictal network measures alone would obviate the need for a lengthy and costly hospital stay for the purpose of recording seizures.

Node-level centrality metrics capture the relationship between each node and local or global patterns in the network. For example, betweenness centrality quantifies the number of shortest paths that are routed through a given node, and is higher in the resection zone of good outcome versus poor outcome patients[53]. A similar metric known as eigenvector centrality estimates the influence of each node on the entire network based on the eigenstructure of the network adjacency matrix. Burns et al[54] found eigenvector centrality to be higher in resected regions in good outcome versus poor outcome patients. Another method that quantifies a type of centrality is known as 'virtual resection'[4,55], in which network synchronizability estimates the stability of the fully synchronous state, and control centrality, which is the contribution of each node or brain region to synchronizability. Our group has used these metrics to distinguish seizure types[4], predict surgical outcome[55], and identify surgical targets by mapping out synchronizing and desynchronizing regions of the network.

To quantify patterns that extend beyond the individual node or edge level, a variety of techniques exist for studying medium and large scale network structure[56]. One such technique, non-negative matrix factorization (NMF)[57,58], can deconstruct the network adjacency matrix into sub-graphs which have time-varying coefficients and together sum to the full network activity. Performing NMF is able to uncover patterns in connectivity that represent the seizure onset zone in both interictal and ictal time frames[59]. Furthermore,

ictal and interictal sub-graphs have similar topologies which may provide methods for determining good surgical targets. Another study determined that epileptic brain networks have 'rich club' structure characterized by high connectivity nodes that themselves were highly connected to each other[60], and that resecting more of the 'rich club' population leads to better surgical outcome. Assessing broad network properties may serve as a powerful method of both understanding the influence of brain networks on epileptic activity as well as guiding surgical intervention.

2.2.2 Model-based approaches

In contrast to traditional network metrics, functional connectivity derived from iEEG can be used to weight links between neural mass models (NMM) which use principles from dynamical systems to simulate epileptic activity. In a NMM, a small number of variables representing excitatory or inhibitory currents affect the state or mean field of a neural population, typically at the scale of columns or cortical areas. There exist a wide variety of neural mass models that have been thoroughly studied in epilepsy[37,61–63], but there exists a common underlying theme. Each node, representing a local population of neurons, can inhabit one of two states representing interictal and ictal activity, and stochastic input from internal or external sources drives switching between states. Dynamical modeling has the potential advantage of better clinical interpretability when compared to model-free network approaches as it provides the ability to perturb the network model. Perturbations can simulate the consequences of different resection or ablation interventions, which could be a useful in surgical planning in the case that the primary hypothesis lies in eloquent cortex[64].

Sinha et al[65] used this principle and assigned each brain region a differential equation that governed a dynamical system. They calculated average interictal correlation between pairs of channels to derive a functional connectivity matrix which they used to link neural mass models. They used the hypothesis that transition between the bistable states of non-seizure and seizure is caused by stochastic noise. The authors measured the time to seizure transition for each node, and defined the EZ by identifying nodes which quickly transitioned into seizures. They showed that these nodes are more often resected in good versus poor outcome patients, and they were able to use this model to identify potential surgical targets and predict outcome. They provided an important validation step of showing that this model performed better than just purely identifying nodes with high strength of weighted connections. It is promising that excellent performance can be achieved using interictal recordings alone, and the approach in this paper may be a good candidate for future clinical translation.

Another well-characterized phenomenological model known as brain network ictogenicity (BNI) carries significant promise for planning epilepsy surgery[66]. Rather than attempting to simulate seizures, the BNI paradigm simulates spiking activity using the Wendling model[67], a set of twelve ordinary differential equations that quantify post-synaptic potentials at each node and has been shown to replicate seizure dynamics[68]. Each neural population can range from quiet to active, in which it generates discharges. To determine connection weights between the neural mass models at each node, they calculated mutual information from patient ECoG during clinical seizures. The authors define BNI as the fraction of time the model spends in the discharge state, and showed that

high NI nodes were more likely to be removed in patients that became seizure free after surgery compared to poor outcome patients. They also showed high ΔBNI in good outcome versus poor outcome patients, which gives evidence for this model's utility in both predicting surgical outcome as well as identifying putative targets.

2.3 The path forward

While many of the methods I have discussed here show early promise in retrospective studies, the path forward requires that clinicians trust quantitative results and learn to integrate their findings into the clinical workflow. These methods will not replace traditional iEEG evaluation entirely but rather will likely serve as additional data for clinicians to consider when identifying surgical targets for epilepsy surgery. In this section I discuss a few significant points towards improving previous models and moving towards clinical translation (Figure 2.5).

2.3.1 Generalizability

A critical aspect of any model is its ability to generalize to new data on which it was not developed. In the field of EZ localization, one chief barrier to generalizability is differences in electrode recording paradigms. Historically, centers in the United States have favored electrocorticography though grid and strip electrodes (ECoG) while centers in France and Italy have favored stereotactically implanted depth electrodes (SEEG). ECoG usually samples a restricted amount of cortical surface on one hemisphere with a regular interelectrode spacing such as on a typical grid array, whereas SEEG typically records from more distributed brain regions in a sparse manner using a number of individual depths to sample deep structures and white matter which are inaccessible to ECoG. These differences

in spatial sampling might not cause significantly different results of univariate electrographic methods, but could greatly affect bivariate functional connectivity and subsequent network analyses, which depend on interelectrode distance and the underlying connectivity patterns between distant brain regions. Thus, translating EZ localization methods from ECoG to SEEG or vice-versa should be done with care and may require methodological modification in many instances.

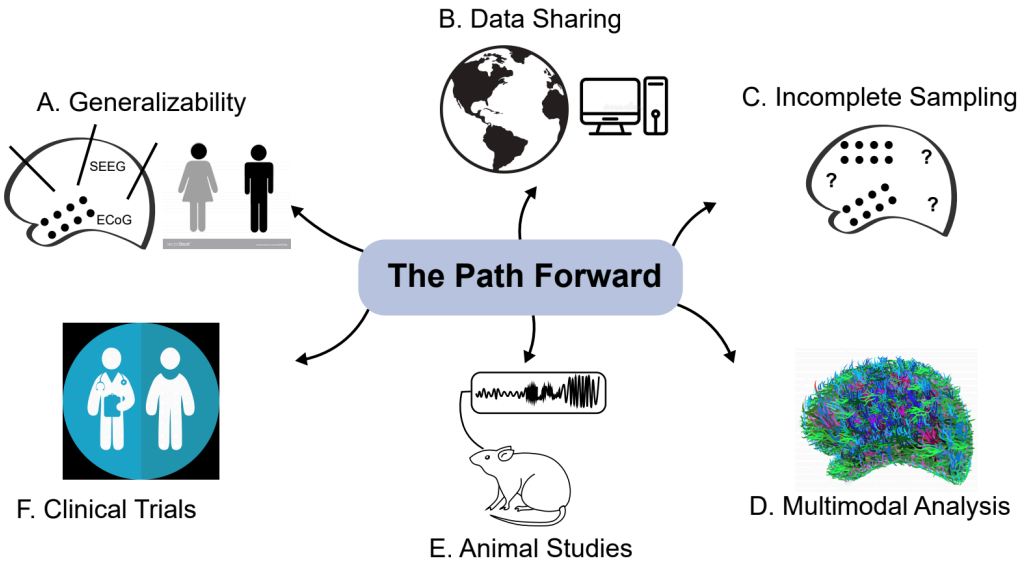


Figure 2.5. The path forward. A) Generalizing current studies of EZ localization methods will require (i) validating methods in both ECoG and SEEG and (ii) accounting for inter-patient heterogeneity by testing at different centers. B) Data-sharing will be essential for aggregating enough data to capture less common types and patterns of seizures for use in future studies. C) Methods must be developed to understand and account for sampling bias introduced by sparse implantation of electrodes. D) Multimodal studies which incorporate neuroimaging or MEG with iEEG may be able to improve models. E) Animal studies could be a bridge from retrospective human studies to prospective human studies. F) Prospective human studies must eventually be undertaken.

Another generalizability barrier will be the significant patient heterogeneity that exists among individuals and across patient populations that are treated by different epilepsy centers. For example, methods which were developed primarily for patients with

mesial temporal lobe epilepsy may not generalize well to patients with neocortical disease. Even within the same epilepsy syndrome, measures that accurately delineate the EZ boundary of patients with a lesional MRI might not do so in non-lesional patients. Many centers do not have the ability to acquire large volumes of similar patients and thus many models will need retrospective validation in cohorts of aggregated patients across centers before prospective trials can be initiated.

Another area of contention is whether epilepsy surgery aims to resect focal pathology or intervene upon a broader epileptic network. The approach of signal analysis lends itself to the hypothesis of a seizure-generating focus by determining which candidate brain region out of all regions sampled by iEEG is the most pathogenic. On the other hand, network-based methods of EZ localization seem to favor the idea that epileptic networks are distributed by considering how the entire brain influences the behavior at a single node and vice versa, as well as the complicated patterns of seizure initiation and propagation that could arise out of a disordered network. Advances in experimentation and imaging may provide evidence for one hypothesis over the other.

2.3.2 Data sharing

Ultimately, the development of a high-quality quantitative method of EZ localization cannot be completed by a single center alone. This is in part due to the wide variability across patients as to their clinical disease, intracranial implant characteristics, and ultimate therapy received. Thus, the path to clinical translation will require broad collaboration and the investment in resources for data sharing. One such platform is ieeg.org[69,70], an online electrophysiology sharing portal with over 5000 users worldwide

which houses over 100 public intracranial EEG recordings complete with accompanying imaging and clinical metadata from the University of Pennsylvania and the Mayo Clinic, in addition to and thousands of scalp, animal and other human studies. The repository on ieeg.org is sufficient to support developing quantitative EZ localization methods, however more multicenter data will be necessary for validation and to prevent overfitting of models to the most popular and available datasets. While the challenges of open-source science are significant, the benefits will be invaluable for translating the methods described in this review towards clinical practice.

2.3.3 Incomplete sampling of brain

Another significant barrier to the clinical translation of quantitative iEEG methods is the clinical constraint of sparse brain sampling. While the current approach to iEEG implantation favors testing clinical hypotheses on the location of the EZ with the minimum number of electrodes possible, quantitative analysis could yield incorrect predictions if the epileptic network isn't properly characterized. Preliminary evidence suggests network metrics have different sensitivities to subsampling on intracranial EEG[71], and the predictive power of personalized network models for epilepsy is improved with a greater number of nodes[72]. We need better methods of determining when the network is sufficiently characterized to deploy quantitative measures with confidence. One possible method of doing so could be examining concordance of results across models. For example, multiple neural mass models have concordance with nodal heterogeneity is high, which is known to be a network feature which is correlated with good outcome. Another method could be jackknife resampling of electrode contacts to simulate alternate implantations, which could

generate spatial confidence intervals for localization of EZ based on network models[71]. Thus, if resampling the network has a minor influence on the localization, it may mean that the network is well characterized. Perhaps the best method, however, would be to aggregate large numbers of recording from patients with similar syndromes and different interventions and outcomes, to better represent the range of individual expressions of common medically refractory epilepsy syndromes Further study is needed to validate methods of confirming sufficient localization and doing so will hasten the deployment of quantitative iEEG paradigms in clinical practice.

2.3.4 Integration of multimodal data

In order to move beyond the limitations of intracranial EEG analysis such as sampling bias and to potentially alleviate the need for extensive intracranial implants in the future, non-invasive measures from MEG, TMS, structural imaging, or functional imaging may be combined with iEEG analysis. One significant example of this is the virtual epileptic patient (VEP), which uses diffusion weighted imaging to estimate connection weights between neural mass models which are tuned to closely approximate the seizure activity in each individual patient[73]. Another method under study is structure-function coupling in which diffusion imaging is compared with functional imaging or iEEG to understand how activity propagates through the brain's underlying connections. However, methods which alleviate the need for intracranial EEG in patients for which it is currently indicated remain elusive.

2.3.5 Clinical trials

The only route to translating quantitative methods of localizing the EZ to patient care is through clinical trials. However, given the substantial morbidity involved with a failure of resective or ablative epilepsy surgery, it is difficult to clinically validate EZ localizing tools. One of most challenging aspects of designing a prospective trial will be defining the exact application of the candidate model. Will it be used to refine the surgical resection boundaries once a general brain region target has been established? Or would it identify candidate resection zones in difficult cases where clinicians are stumped. Currently in the EPINOV trial for the Virtual Epileptic Patient model, a list of the highest likelihood targets and their probabilities of lying within the EZ provided to clinicians. Another trial, currently underway at ICL in London, allows addition of up to 3 additional electrodes to standard implants, in a controlled, randomized trial of adding quantitative methods to standard clinical practice to assess their ability to improve patient outcome. As more iEEG analysis tools reach maturity, more such trials may be initiated and methods can be refined for the exact clinical scenarios for which they are best suited.

CHAPTER 3: Analysis of seizure networks

3.1 Abstract

Patients with drug-resistant epilepsy often require surgery to become seizure-free. While laser ablation and implantable stimulation devices have lowered the morbidity of these procedures, seizure-free rates have not dramatically improved, particularly for patients without focal lesions. This is in part because it is often unclear where to intervene in these cases. To address this clinical need, several research groups have published methods to map epileptic networks, but applying them to improve patient care remains a challenge. In this study we advance clinical translation of these methods by: (1) presenting and sharing a robust pipeline to rigorously quantify the boundaries of the resection zone and determining which intracranial EEG electrodes lie within it, (2) validating a brain network model on a retrospective cohort of 28 patients with drug-resistant epilepsy implanted with intracranial electrodes prior to surgical resection, and (3) sharing all neuroimaging, annotated electrophysiology, and clinical metadata to facilitate future collaboration. Our network methods accurately forecast whether patients are likely to benefit from surgical intervention based on synchronizability of intracranial EEG (area under the ROC curve of 0.89) and provide novel information that traditional electrographic features do not. We further report that removing synchronizing brain regions is associated with improved clinical outcome, and postulate that sparing desynchronizing regions may further be beneficial. Our findings suggest that data-driven network-based methods can identify patients likely to benefit from resective or ablative therapy, and perhaps prevent invasive interventions in those unlikely to do so.

3.2 Introduction

Epilepsy affects 65 million people, one third of whom are resistant to antiepileptic medications. In these cases, surgery is often necessary to help reduce seizures[10]. Resections can be generous, with a significant amount of healthy tissue removed, frequently resulting in post-surgical side effects, including neuropsychological deficits and reduced quality of life[8]. Unfortunately, some patients have seizure recurrence despite removal of assumed critical seizure generators, as mapped by extensive electroencephalography, multimodal neuroimaging, and neuropsychological evaluation prior to surgery[10]. Due to the limitations of traditional epilepsy surgery, clinicians are turning to other less destructive therapeutic approaches. Specifically, responsive neurostimulation, deep brain stimulation, and targeted laser ablation techniques are increasingly being used to alleviate seizure burden and improve quality of life[74–76]. These interventions are hypothesized to act by disrupting connections and pathways involved in seizure spread[76]. Identifying these important control regions is a critical step toward realizing the potential of these newer, less invasive techniques and for optimizing the use of established resective surgery.

While focal brain lesions have long been a target of epilepsy surgery with favorable success rates, patients without clear lesions may have seizures that arise from abnormal connectivity in broader networks that can be measured at the scale of electrocorticography (ECoG). Recent work supports the hypothesis that epilepsy can arise from disordered brain networks[47]. In epileptic networks, the seizure onset zone (SOZ) often not only drives seizure initiation and propagation, but also recruits regions that extend well beyond it to act as central hubs. These regions appear to strengthen in connectivity to each other, while weakening in connectivity to remaining regions[77,78]. Because the epileptic network may be

characterized by pathologic foci embedded in this web of structural and functional connections[79], it is important to understand how aberrant cortical functioning drives seizure dynamics and manifests in the diverse roles of regions such as the epileptogenic, irritative, and propagation zones. Thus, a network approach that quantifies the complex synchronization and spread of neural activity is well suited to studying epilepsy in which changes in brain connectivity manifest across a wide range of spatiotemporal scales.

Recent efforts to translate and extend methods originally developed in network science have generated novel *in silico* approaches to model epileptic networks[49,54,73] and identify important regions to target therapeutically with surgical[4,52,60,65,66] or non-surgical[80] interventions. While some of these models quantify brain dynamics through data-driven network models[4,52,54], others integrate network architecture estimated by intracranial EEG or imaging with generative mathematical models that parameterize behavior at each node[60,65]. The Virtual Epileptic Patient (VEP)[73] is one notable model that uses structural connectivity estimated from diffusion weighted MRI to parameterize coupled 'Epileptor' oscillators[81] which predict seizure propagation and spread. The VEP framework is currently being studied in a prospective clinical trial to augment clinician decision making in epilepsy surgery (US National Library of Medicine, 2018). Previous studies have established the potential of modeling to enhance our understanding of epileptic networks, but their translation to clinical care has been challenging. There are multiple potential reasons for this. Some approaches such as the VEP use models that generate synthetic seizures that look remarkably similar to clinical events, but because they are synthetic there is concern that they may not capture the complex interplay between brain regions with

inherently different control properties. Many studies do not use expert clinical annotations from interictal, pre-ictal, and seizure epochs to evaluate the full spectrum of epileptic activity in each patient. Finally, validation has been challenging, particularly for studies that use clinical data. For a variety of reasons, most groups have not openly shared their methods and data so that other centers can reproduce and extend their studies. We aim to specifically overcome each of these shortcomings in this study using the virtual resection framework developed by our group.

Virtual resection[4] implements a brain network model in which the regions measured by individual intracranial electrode contacts are defined as nodes, and the statistical relationships between pairs of nodes known as functional connectivity are defined as edges[49]. This approach focuses on calculation of synchronizability, which describes the ease with which neural activity can propagate through the network. Virtual resection uses a push-pull framework in which synchronizing nodes dynamically oppose desynchronizers and thus the properties of the network are modulated as a function of time. By virtually removing nodes and recalculating synchronizability, we quantify each node or region's control centrality[4] and thus estimate how well dynamic brain activity such as seizures would spread throughout the network if a given region were removed. While this framework has been shown to robustly characterize the spatiotemporal regulators of seizure dynamics[4], in our prior work we did not engage in any effort to predict surgical outcomes nor did we investigate whether control centrality within the resection zone played a key role in accurately predicting that outcome. In the work presented here, we hypothesize that removing synchronizing versus desynchronizing brain regions as determined by the virtual

resection method is predictive of good versus poor outcomes. In addition, we hypothesize that control centrality within the resection zone is predictive of epilepsy surgical outcomes. In this multi-center retrospective cohort study, we seek to determine whether virtual resection network features can predict surgical outcome in 28 patients with drug resistant epilepsy who underwent surgical therapy. Specifically, we ask: What is the spatiotemporal evolution of synchronizability during a seizure? Is there a relationship between network synchronizability and important clinical variables used to guide therapy? What is the accuracy of predicting surgical outcome from the control centrality of combined group nodes overlying resection cavities? We hypothesize that virtual resection will uncover important seizure dynamics as measured by network metrics that can separate patients based on post-surgical outcome and that network metrics will be sensitive to clinical variables such as seizure type.

3.3 Materials and Methods

3.3.1 Patient dataset

All patients included in this study gave written informed consent in accordance with the Institutional Review Board of the University of Pennsylvania. All patients from both the Hospital of the University of Pennsylvania and the Mayo Clinic gave consent to have anonymized full-length electrophysiology recordings and brain MRI and CT scans available to the public on the open online portal IEEG.org[69].

Twenty-eight patients undergoing surgical treatment for medically refractory epilepsy (Table 3.1) underwent implantation of subdural and depth electrodes to localize the seizure-onset zone after presurgical evaluation with scalp electroencephalogram (EEG)

recording of ictal epochs, MRI, fluorodeoxyglucose - positron emission tomography (FDG-PET), and neuropsychological testing. The results of the presurgical studies suggested that focal cortical resection might be a therapeutic option. Patients were then implanted with intracranial electrodes including grid, strip and sparse depth electrodes, to better define the epileptic network. De-identified patient data were retrieved from the online International Epilepsy Electrophysiology Portal (IEEG Portal). Patients were excluded from the study if they (1) did not undergo resection, (2) did not have complete post-resection imaging, or (3) did not have complete electrophysiology data. Surgical outcome was measured at a minimum of one year after surgery and determined based on medical records from the last available follow up with a clinician. Patients who had surgical outcome of Engel I or ILAE 1-2 were marked as having favorable outcome and patients who had Engel II-IV or ILAE 3-6 were marked as having poor outcome (Wieser *et al.*, 2001).

ECoG signals for patients from the Hospital of the University of Pennsylvania were recorded and digitized at 500 Hz sampling rate and pre-processed to eliminate line noise. Cortical surface electrode configurations, determined by a multidisciplinary team of neurologists and neurosurgeons, consisted of linear and two-dimensional arrays (2.3 mm diameter with 10 mm inter-contact spacing) and sampled the neocortex and mesial cortex for epileptic foci. Signals were recorded using a referential montage with the reference electrode, chosen by the clinical team, distant to the site of seizure onset and spanned the duration of a patient's stay in the epilepsy-monitoring unit. All EEG recording systems and intracranial electrodes used were FDA approved and commercially available.

	Good surgical outcome (Engel I or ILAE 1-2)	Poor surgical outcome (Engel II-IV or ILAE 3-6)	*p* - value
Total number of subjects	17	11	
Age at surgery*			$p = 0.99^a$
Mean ± std. dev.	36.2 ± 11.1	33.1 ± 18.6	
Sex			$p = 0.93^b$
Male	9	6	
Female	8	5	
Resected / Ablated region			$p = 0.11^b$
LTL	3	6	
RTL	8	1	
LFL/LPL/LFPL	4	2	
RFL/RFTL/RFPL	2	2	
MRI*			$p = 0.97^b$
Lesional	7	4	
Non-Lesional	9	5	
Pathology*			$p = 0.02^b$
HS/MTS	6	2	
Gliosis	3	8	
Malformations of cortical development	5	0	
Tumor/vascular/infection	2	0	
Seizure Type*			$p = 0.20^b$
Aura/Focal aware	14	3	
Focal impaired awareness	51	30	
Focal w/generalization	31	11	
Type of resection			$p = 0.68^b$
Anterior temporal lobectomy and/or amygdalohippocampectomy	3 (left), 7 (right)	6 (left), 2 (right)	
Anterior temporal lobectomy+	1 (left)	None	
Partial resection/lesionectomy	3 (left), 2 (right)	2 (left), 1 (right)	
RF ablation	1 (right)	None	
Volume of resection			
Volume of tissue (cc)			$p = 0.83^a$
Mean ± std. dev.	19.3 +/- 12.8	23.5 +/- 18.9	
Nodes removed (%)			$p = 0.72^a$
Mean ± std. dev.	19.3 +/- 14.4	21.5 +/- 10.8	

Table 3.1. Virtual resection dataset. Patients were grouped by surgical outcome. Second column shows patients that had a favourable surgical outcome. Third column shows patients that had a poor surgical outcome. FCD = focal cortical dysplasia; HS = hippocampal sclerosis; LFL = left frontal lobe; LFPL = left frontoparietal lobe; LPL = left parietal lobe; LTL = left temporal lobe; MTS = mesial temporal sclerosis; PNH = periventricular nodular heterotopia; RF = right frontal; RFL = right frontal lobe; RFPL = right frontoparietal lobe; RPL = right parietal lobe; RTL = right temporal lobe; TSC = tuberous sclerosis complex.

[a]*T*-test; [b]Pearson chi-square test. *Data for these fields were unknown in a minority of patients and was not thus included in the table.

3.3.2 Clinical marking of seizure events

Seizure-onset zone was marked on ECoG according to the standard clinical protocol in the Penn Epilepsy Center. Initial clinical markings are made on IEEG by board-certified staff epileptologist attendings. These IEEG markings were made by clinicians blinded to surgical outcome but provided with available surgical conference notes that contained patient clinical data related to other multimodality testing, such as brain MRI, PET scan, neuropsychological testing, and ictal single-photon emission computed tomography (SPECT) scanning, used to finalize surgical approach and planning. The following seizure times were annotated for each seizure: (1) earliest electrographic change (EEC)[82], (2) unequivocal electrographic onset (UEO), and (3) termination of seizure (END). A pre-seizure state that spanned a period equal in duration from EEC to END was used to mark the baseline pre-ictal period for any given seizure. An epileptologist involved in the study examined the patient's primary seizure type for the presence of localizing factors such as low voltage fast activity (LVFA), DC shift, and a clearly focal seizure onset zone which are known to predict outcome[83–86]. If a disagreement regarding annotation arose, at least two epileptologists discussed the seizure in question until reaching consensus. All seizures were identified according to the current International League Against Epilepsy (ILAE) classification system, as focal aware seizures, focal impaired awareness seizures, or focal to bilateral tonic-clonic seizures. When clinical notes were unclear, the reviewing epileptologist made a decision on seizure type based on all available clinical data. In order to support the potential generalizability of our methods to any new patient, network measures were computed on all seizures, interpolated to fit into ten

sequential time bins spanning the pre-seizure and seizure epochs, and averaged within a patient's group of seizures.

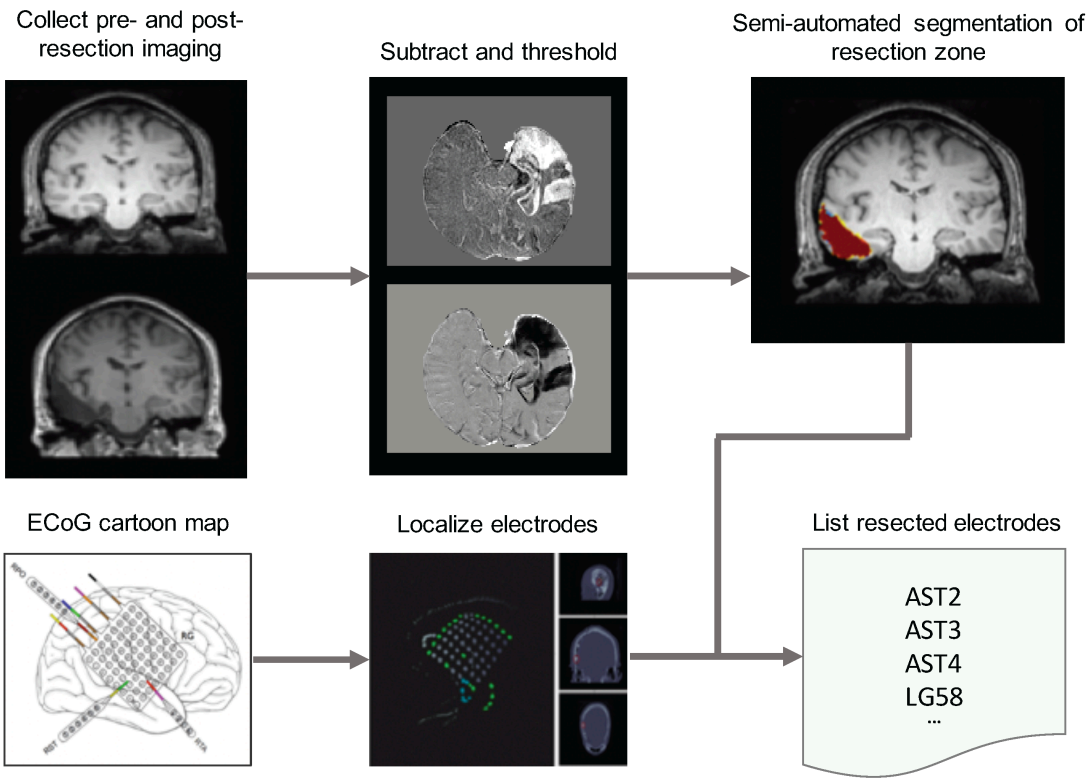


Figure 3.1. **Imaging pipeline for resection zone estimation.** The imaging pipeline registers pre-operative MRI, post-implant CT, and post-resection MRI together in order to allow for resection zone mapping and electrode localization. In-house software in addition to ITK-SNAP were used to map resection zones and localize electrodes for all patients using cartoon maps presented in surgical conference notes. Nodes that overlapped with the resection zone were virtually resected from the network to measure effect on synchronizability.

3.3.3 Image Processing

All patients, as part of their clinical neuroradiological workup, underwent a clinical epilepsy neuroimaging protocol. Pre-implant T1-weighted MPRAGE MRI, post-implant T1-weighted MPRAGE MRI, and post-implant CT images were acquired in order to localize electrodes. In addition, patients underwent a post-resection imaging protocol

acquired on average 6-8 months after implant and resection, which consisted of T1-weighted MRI and axial FLAIR MRI sequences. All images were stripped of headers, anonymized and registered to patient's native pre-implant T1 MRI space for localization and segmentation (Figure 3.1). In-house software[87] was used to assist in localizing electrodes after registration of pre-implant and post-implant neuroimaging data. All electrode coordinates and labels were saved and matched with ECoG electrode names on IEEG.org. All electrode localizations were verified by a board-certified neuroradiologist.

Pre-implant MRI imaging was registered diffeomorphically using the Advanced Normalization Toolkit (ANTs)[88] to post-resection imaging in order to accurately segment the resection zone. Resection zones were estimated semi-automatically with the use of a random forest classifier and region-growing algorithm as part of the ITK-SNAP toolkit[89]. All resection estimates were confirmed by a board-certified neuroradiologist.

3.3.4 Network methods

The virtual resection method (Figure 3.2) is described in greater detail in Khambhati *et al.* Electrodes in which the ECoG signal was obscured by artifact, as noted by an attending epileptologist, were removed from analysis in order to avoid biasing our results. A common average reference was applied to all neural signals by first computing a time-varying signal averaged across all electrodes and then by subtracting this signal from each electrode. All ECoG signals were notch filtered at 60 Hz to remove power line noise. We constructed functional brain networks in each time window using multitaper coherence estimation, which defines an edge between electrode pairs in terms of the correlation of the power spectra of signal activity over a specific frequency band. This

procedure was done across different physiological frequency bands, namely: α/θ (5 – 15 Hz), β (15 – 25 Hz), low-γ (30 – 40 Hz), high-γ (95 – 105 Hz) and very high frequencies (>105 Hz). In addition, broadband cross-correlation was used to generate functional dynamic networks without regard to frequency specific information. In this study we compute functional networks directly from clinical recordings of intracranial EEG and not on modeled or simulated data. These networks were generated as N x N symmetric adjacency matrices, describing the network for all T time windows.

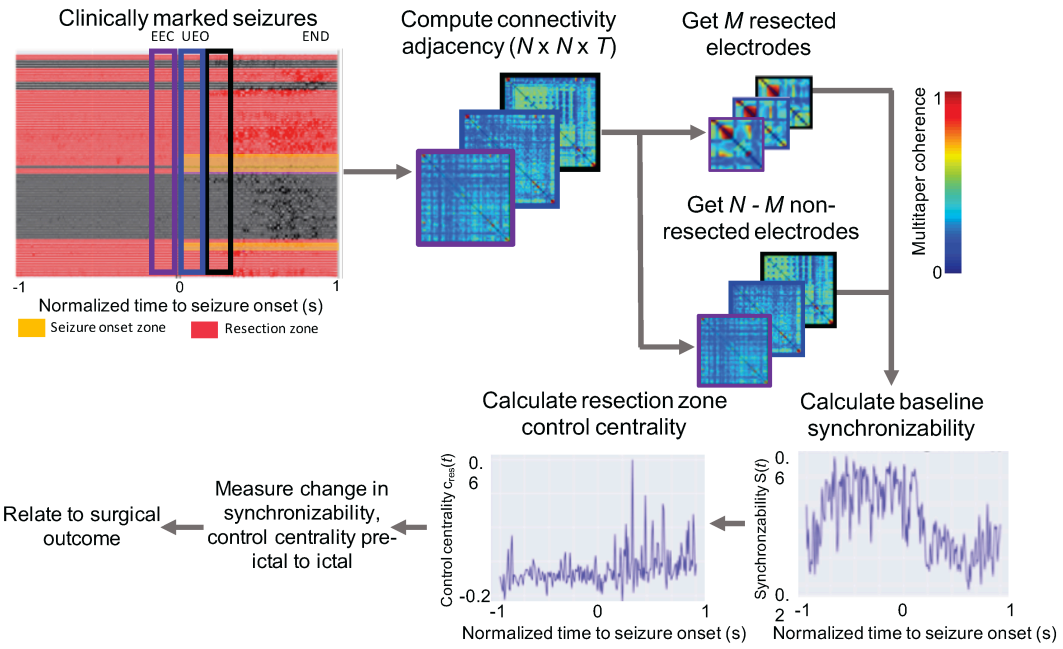


Figure 3.2. Electrophysiology pipeline for epileptic network analysis. The electrophysiology pipeline uses coherence to construct adjacency matrices for each event, separately for seizure and pre-seizure data. Baseline synchronizability of the network and control centrality of each individual node is calculated. From the imaging pipeline, resected electrodes are determined, and control centrality of the resection zone is calculated. Metrics are used to generate predictions of surgical outcome.

Because of its importance to seizure spread, we measured network synchronizability by first computing the Laplacian matrix of each adjacency matrix at time

one-second time windows. The Laplacian matrix can be interpreted as measuring the ease with which information diffuses between nodes in a network[5]. Next, at each time epoch t across all T epochs, we calculated the synchronizability measure, s(t) as the ratio of the second smallest eigenvalue to the largest eigenvalue of the Laplacian matrix which quantifies the stability of the synchronous state[5]. In order to model the effects of resective surgery, we used the approach of virtual cortical resection which quantifies control centrality as the contribution to synchronizability. Control centrality can be calculated either at each node or for the entirety of a region of interest by removing the node or nodes in question from the network and recalculating synchronizability. In this study, we remove the resection zone *en bloc* to calculate control centrality when comparing across patients, while for whole brain visualizations we calculate control centrality at the node level. This measure of change in synchronizability is referred to as the control centrality, or $c_{res}(t)$, and can be used to identify a region as (1) desynchronizing (removal of which increases post-resection network synchronizability) characterized by positive control centrality, or (2) synchronizing (removal of which decreases post-resection network synchronizability) characterized by negative control centrality.

3.3.5 Statistical methods

All averages computed in this study use the median because it represents a better measure of centrality in skewed, non-normal distributions such as seen in the distribution of network measures[90]. We compared median time-varying network metrics between seizures in patients who experienced a favorable surgical outcome and patients who experienced a poor surgical outcome. We performed this comparison by normalizing each

seizure event into ten sequential time bins spanning the pre-seizure and seizure epochs, and employing functional data analysis (FDA) to statistically test differences in temporal dynamics between seizure types independently in each state[91]. FDA allowed us to test whether the area under the good outcome curves and poor outcome curves were significantly different by comparing the true area to the area expected in an appropriate permutation-based null model. The null model was created by re-assigning surgical outcome to adjacency matrices uniformly at random up to 10,000 times and computing the median area under the resulting curves of functional network metrics.

The median network measures during pre-ictal epochs are compared to those of ictal epochs using Wilcoxon rank-sum test. The $\Delta s(t)$ change from pre-ictal to ictal time periods was used as the feature for quantifying the ability to predict surgical outcome. We varied the threshold of $\Delta s(t)$ to predict patients as having either good or poor surgical outcome which generated a receiver operating characteristic (ROC) curve. We measured area under the ROC curve (AUC) as a marker for accuracy in predicting good versus poor surgical outcome; the ROC curve quantifies the tradeoff between the true positive rate and the false positive rate for a binary classifier. We used the non-parametric DeLong test to compare ROC curves and determine whether any single predictive model derived from a specific frequency band performed significantly better than the others[92]. In the analysis of control centrality of the resection zone we present p values noting that the significance threshold is 0.0083 to correct for six comparisons introduced by the separate frequency bands according to the Bonferroni method.

3.3.6 Data availability

Our codebase comprises the imaging and electrophysiology pipelines (https://github.com/ieeg-portal/EpiVR), and allows researchers to easily fetch data situated on the IEEG.org portal as well as to perform virtual cortical resection. The electrophysiology pipeline is dependent on Echobase, which can be found at https://github.com/akhambhati/Echobase. We have additionally made pre- and post-resection imaging as well as annotations of seizures along with their ECoG recordings from the entirety of their epilepsy monitoring unit (EMU) stay available to the public through IEEG.org. From this unique and powerful dataset we hope that other investigators may validate our methods or compare their performance to other virtual resection tools.

3.4 Results

Twenty-eight patients with drug resistant epilepsy underwent implantation with ECoG electrodes to localize the seizure onset zone preceding surgical resection. Patient specific network models were constructed from clinically annotated ECoG recordings stored on the cloud platform IEEG.org[69,70], and a quantitative pipeline using pre- and post-surgical imaging was used to determine which electrodes were resected.

3.4.1 Network synchronizability predicts surgical outcome

We began our virtual resection approach by examining the dynamic network changes in synchronizability s(t) that occur in the transition state from pre-ictal to ictal periods. Seizures were interpolated to be equal length across patients and onset times were aligned such that s(t) curves could be compared. Figure 3.3A shows the time course of

median broadband synchronizability over pre-ictal and ictal periods between good and poor outcome patients. These curves were different in the ictal period between outcome groups (FDA curve area test, 10,000 permutations, $p < 0.001$) suggesting that there is predictive information in the magnitude and temporal evolution of s(t). Patients with favorable surgical outcomes had decreased synchronizability at the time of seizure onset across frequency bands while s(t) remained high in those with poor surgical outcomes. Synchronizability curves for other frequency bands can be found in Supplementary Figure 3.1.

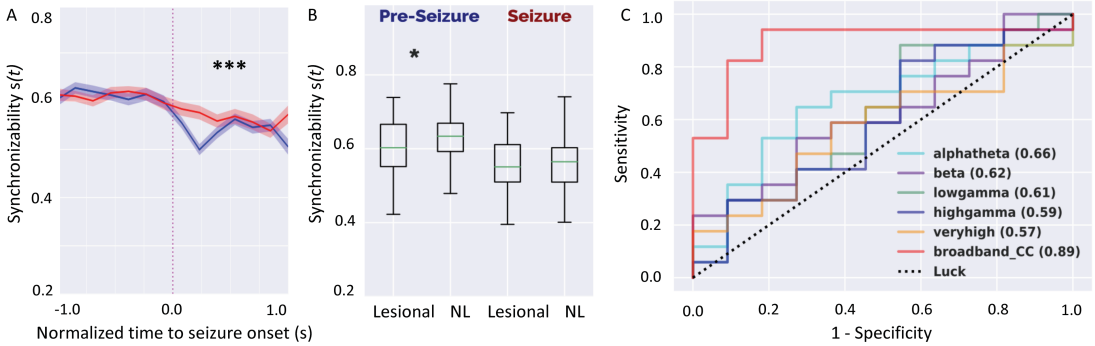


Figure 3.3. Time-varying network synchronizability is predictive of surgical outcome. (A) Median base network synchronizability in good outcome patients (blue) and poor outcome patients (red) for broadband intracranial EEG (IEEG), *** = $p < 0.001$. Shaded areas show 95% CIs. (B) Patients with lesional MRI have higher pre-ictal synchronizability than non-lesional (NL) patients, * = $p < 0.05$. (lesional, pre-ictal: min = 0.42, 25% ile = 0.55, median = 0.60, 75% ile = 0.67, max = 0.74. NL, pre-ictal: min = 0.33, 25% ile = 0.59, median = 0.63, 75% ile = 0.67, max = 0.85. lesional, ictal: min = 0.40, 25% ile = 0.51, median = 0.55, 75% ile = 0.61, max = 0.70. NL, ictal: min = 0.40, 25% ile = 0.51, median = 0.56, 75% ile = 0.60, max = 0.74) (C) ROC curves were constructed by calculating difference in s(t) from pre-ictal to ictal periods and sweeping the threshold for classification. Broadband IEEG predicts surgical outcome significantly better than other bands as assessed by the DeLong test which statistically compares ROC curves generated from correlated data.

We also asked whether there is an association between synchronizability and clinical variables used to guide therapy such as lesion status on MRI. We found that patients

who were non-lesional (n = 14) had higher pre-ictal synchronizability in broadband compared to patients with lesions (n = 11) (rank-sum statistic -1.984, $p = 0.047$) (Figure 3.3B), even though there was no correlation with lesion status and outcome (chi-square test = 0.178, $p = 0.67$) or electrode number (rank-sum statistic 267, $p = 0.79$) in our study.

In order to quantify the ability of synchronizability to predict surgical outcome, we calculated median change in $s(t)$ from pre-ictal to ictal periods, and performed a sweep of that feature to generate a receiver operating characteristic (ROC) curve (Figure 3.3C). Despite the small sample size, the $\Delta s(t)$ derived from broadband ECoG data (AUC=0.89, 95% CI 0.76-1.00) is predictive of surgical outcome. This ROC curve showed greater statistical significance than that observed in each of the other frequency bands. Selecting the point on the ROC which gave the greatest number of correct classifications, we choose a threshold of $\Delta s(t) = 0.0279$ to determine the performance of our predictive model based on broadband networks across all 28 patients in our cohort. We find an accuracy of 0.86 with a true positive rate of 0.88 and a true negative rate of 0.82 (Supplementary Table 3.1).

3.4.2 Virtual resection provides novel clinical insight

We sought to determine whether the virtual resection model provides additional information on patient prognosis that is not merely correlated with traditional electrographic findings such as DC shift, clear seizure focus, and low voltage fast activity. While a significant decrease in broadband synchronizability at seizure onset carries an odds ratio of 35 for good surgical outcome, having a focal seizure onset zone, low voltage fast activity, or a DC shift do not perform as well with odds ratios of 4.3, 1.5, and 7 respectively

(Supplementary Table 3.2). We present the EEG and synchronizability of two patients that our model predicted correctly (Figure 3.4). In these cases, following electrographic features alone did not correlate with outcome and we find that virtual resection provides novel clinical information not captured by traditional clinical analysis. We then asked the following questions: How does control centrality $c_{res}(t)$ of the resection zone change before and during a seizure? How does it differ between patients who fare favorably and patients who fare poorly after surgery? We observed that median $c_{res}(t)$ was lower in good outcome patients compared to poor outcome patients during the ictal period. In focal impaired awareness seizures the β frequency analysis was strongest (rank-sum statistic –2.10, $p = 0.036$), while in focal to bilateral tonic-clonic seizures this effect was strongest in high-γ (rank-sum statistic –2.43, $p = 0.015$) (Figure 3.5). We note that after adjusting the alpha level to 0.0083 for multiple comparisons, these results no longer reached statistical significance. The effect was not significant during pre-ictal periods or in other frequency bands (Supplementary Figure 3.2). Additionally, calculated $c_{res}(t)$ is robust to segmentation error at the resection zone margin (Supplementary Figure 3.3). These findings align with our initial hypotheses as well as with the theoretical understanding of the virtual resection network analysis.

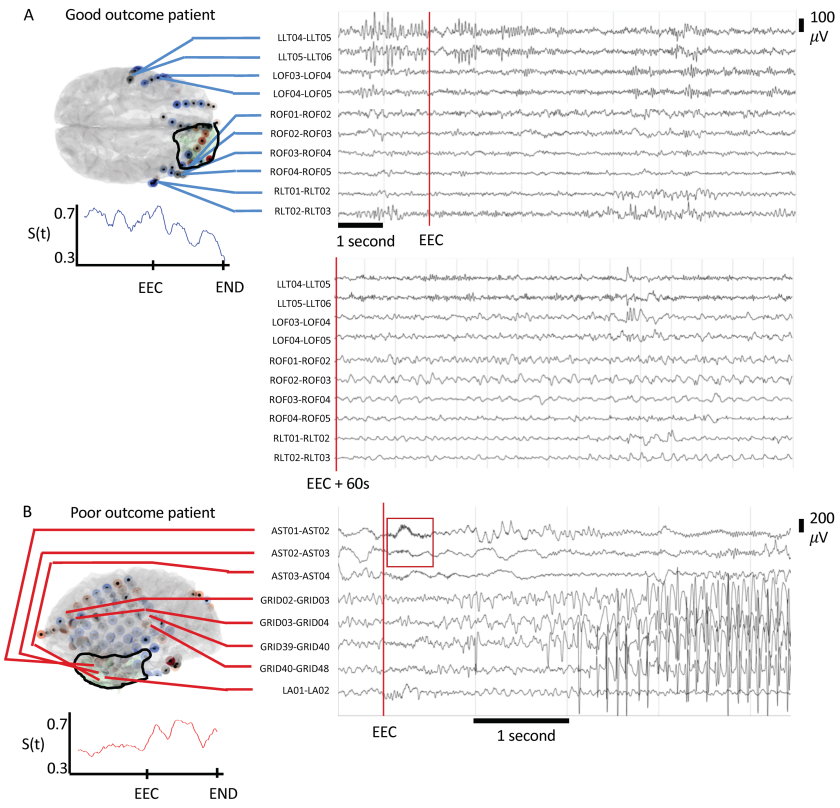


Figure 3.4. Synchronizability provides novel clinical information. (A) Our model correctly predicted HUP073 to be seizure free after surgery. A board certified epileptologist determined there is no DC shift, low voltage fast activity, or clearly focal seizure onset. This stereotyped clinical seizure results in arousal from sleep without evident EEG change. While the clinically marked seizure onset began with an arousal pattern in LLT04-06, here rhythmic activity begins in ROF1-3 60 seconds into seizure, later progressing to RLT1-3. .Resection was performed in the right frontal region. Synchronizability decreases throughout the seizure in this patient, predicting seizure freedom after surgery. (B) Our model correctly predicted the poor outcome of patient HUP080. In the displayed seizure, a board certified epileptologist determined seizure onset electrodes of AST1-3 with the presence of low voltage fast activity (red box), and a clearly focal seizure onset but no DC shift. Synchronizability increases after EEC in this patient, correctly predicting the surgical outcome of this patient to be poor. EEC = earliest electrographic change.

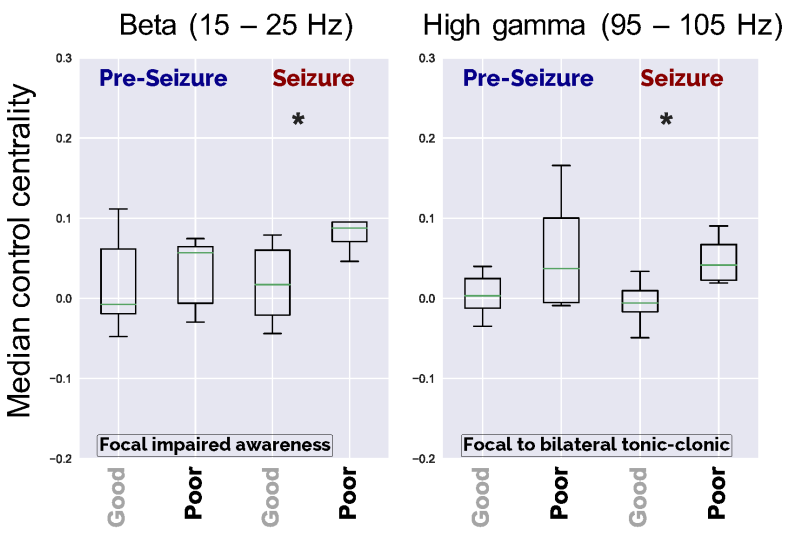


Figure 3.5. Control centrality of resection zone. Median node-level control centrality -- calculated for the entire group of nodes lying within the resection zone -- are shown for different seizure types. Good outcome patients have regions that play a greater synchronizing role resected compared to poor outcome patients. This effect is greatest in the β band for focal impaired awareness seizures (good: min = -0.044, 25% ile = -0.021, median = 0.017, 75% ile = 0.060, max = 0.220, poor: min = 0.046, 25% ile = 0.071, median = 0.088, 75% ile = 0.095, max = 0.183) and in the high-γ band for focal to bilateral tonic-clonic seizures (good: min = -0.050, 25%ile = -0.017, median = -0.006, 75% ile = 0.010, max = 0.034, poor: min = 0.019, 25% ile = 0.023, median = 0.042, 75% ile = 0.067, max = 0.090).

3.4.3 Virtual resection maps spatial anomalies in seizure networks

To examine the implications of virtual resection results on clinical management, we sought to elucidate the role that various clinical features can play on virtual resection features. We sought to study the spatial distribution of $c_{res}(t)$ in patients who had undergone focal resections with malformations of cortical development (MCD), such as focal cortical dysplasia. There were six patients with MCD of which two were read as MRI-normal. Five of these patients had favorable outcome. Figure 3.6 shows patients with MCD on pathology and their respective mean control centrality across pre-ictal and ictal epochs derived using

the broadband cross-correlation metric for functional connectivity. In high frequency bands, during pre-ictal periods, we observed that MCDs exhibit stronger synchronizers perilesionally (pre-seizure: rank-sum statistic –2.08, p = 0.04). We also found strongly desynchronizing nodes in the β band during ictal periods perilesionally within the resection zone compared to non-lesional patients (seizure period: rank-sum statistic 1.71, p = 0.09). However, neither of these results achieved statistical significance after correcting for multiple comparisons. These closely localized desynchronizing nodes near lesions may be important controllers that act on nearby seizure generating regions next to dysmorphic dysplastic tissue. Since this finding was present even in the two non-lesional patients with MCD, spatial maps of control centrality may act as a biomarker to uncover hidden epileptogenic lesions that may not be easy to identify on standard clinical neuroimaging.

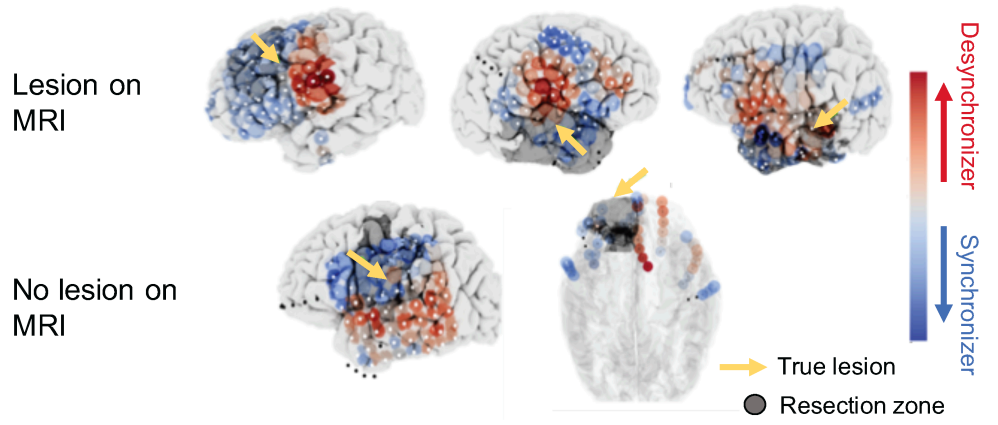


Figure 6. Control centrality uncovers malformations of cortical development. Spatial maps of mean control centrality for each node are shown for all five patients: three lesional and two non-lesional on MRI, seen above. All patients had pathology-confirmed malformations of cortical development (yellow arrow). Two patients had non-lesional findings on MRI despite presenting a spatial pattern of node-level control centrality similar to those in other patients with MRI-positive MCDs. Specifically, strong desynchronizing regions (red) are seen in all resected regions (darkened zone with white outline) which contain MCDs (synchronizers, blue). Artifactual electrodes not included in analyses are denoted with black dots.

3.5 Discussion

In this study, we investigate the ability of virtual resection to accurately predict surgical outcome using functional networks derived from ECoG data, expertly annotated seizure markings, and rigorous coregistration of intracranial electrodes and resected brain regions. We determine that decreased synchronizability at the time of seizure onset is predictive of good outcome, performing better than traditional electrographic features. We further suggest that good outcome patients have brain regions removed with a relatively greater synchronizing effect on broader networks than those of poor outcome patients. Finally, we propose that our robust pipeline incorporating rigorous clinical marking and validation of ECoG, quantified resection zones on standardized MRI after surgery, and sharing of all code and data are novel contributions that make this study important.

In our validation of the virtual resection method, we uncover relationships between synchronizability and important clinical variables. Broadband synchronizability has a significant decrease at seizure onset in good but not poor outcome patients, and pre-ictal synchronizability is higher in non-lesional patients than those with a clear lesion on MRI. The notion that good outcome patients have a decrease in synchronizability at seizure onset is intriguing because it suggests a greater resistance to the propagation of oscillatory epileptic activity throughout the network early in seizures. For patients without this decrease in synchronizability, traditional resections may not provide a cure as the existing epileptic network may poorly constrain abnormal activity. Furthermore, heightened pre-ictal synchronizability in non-lesional patients could underlie differences in the pathophysiology of seizure generation. Furthermore, the robustness of our methods to

predict both non-lesional and lesional patient outcomes equally well is exciting and a clear strength of this study as there is substantial literature supporting improved outcomes in lesional epilepsy[83,86]. While our retrospective analysis is not designed to conclusively prove the mechanism by which either of these observations occur, future experimental studies may explore these concepts *in vivo*. As we continue to extend virtual resection and provide support of its utility we further aim to initiate a prospective clinical trial in the future.

We used differences in synchronizability between good and poor outcome patients to predict surgical outcome and show that it provides novel information not present in traditional electrographic features. Using the change in median s(*t*) from the pre-ictal to the ictal period, we predict surgical outcome with accuracies that compare favorably to other recently-published *in silico* models of resective epilepsy surgery[65,66,80,93]. We also show that examining synchronizability curves can uncover novel information about the epileptic network that is often not present in traditional electrographic localizing features of focal SOZ, DC shift, and LVFA. These findings support the use of our model as an adjunct to traditional EEG interpretation and may identify patients for whom surgery could eliminate their seizures as well as those unlikely to benefit from intervention. We envision clinicians interacting with our model in a similar approach to Figure 3.4, where synchronizability, control centrality, and EEG are viewed alongside each other. A holistic assessment of the full clinical information such as electrographic features is always warranted, however our findings support the use of synchronizability to support surgical decision-making.

Examining the contribution of the resection zone to overall network dynamics may be a powerful tool for assessing surgical intervention. While the results of this analysis are not statistically significant after adjusting for multiple comparisons, median control centrality of the resection zone is lower in good outcome patients (Supplementary Figure 3.2) across all frequency bands which are correlated. We observe that good outcome patients may be more likely to have lower control centrality of their resection zones, meaning that the resected tissue plays more of a synchronizing role in the overall network. It follows that such patients would have good outcome as the topology of the resulting functional brain network would have decreased synchronizability. Performing a high resolution spatial mapping of individual synchronizing and desynchronizing nodes such as in Figure 3.6 may also provide insights into these concepts in the context of specific pathologies. However, larger collaborative datasets are needed to robustly uncover disease specific patterns as our results do not reach statistical significance after adjusting for multiple comparisons. It is tempting to infer that our finding suggests that these regions function as a macro-scale "inhibitory surround," analogous to that seen in more controlled studies of seizure propagation in animal models and humans[45], but this hypothesis would require further investigation.

Broadband cross-correlation may have higher predictive power compared to individual frequency bands because it provides the most general assessment of network connectivity without frequency-specific information. Given that broadband has high correlation with all other frequency bands (Supplementary Figure 3.4), broadband may be the most generalizable as it would not be as affected by differences in individual seizure

pathophysiology. Furthermore, previous network models of epilepsy have found broadband to be highly predictive[93]. On the other hand, when examining control centrality of the resection zone parsed by seizure types we find frequency differences that are not apparent in the broadband analysis. The finding that high-γ synchronization was most associated with good outcome in focal seizures that generalized whereas β synchronization was associated with good outcome in focal seizures that did not spread may be rooted in different mechanisms underlying these events.

Recent work has attempted to study the spatiotemporal dynamics of seizures from initiation to termination[4,94]. Onset patterns are not determined by initiation of aberrant activity in the seizure onset zone core alone, but additionally by how changes in excitability in surrounding healthy tissue cause the onset to become evident[95]. These network state changes may serve as control mechanisms enabling desynchronous activity to disrupt seizures or to coalesce tightly bound and functionally cohesive network components[47]. As a result, some nodes may be seizure desynchronizers that should potentially be left intact in any resection plan (Figure 3.7). In Figure 3.7, the proposed resection zone is the group of synchronizing nodes that fell within the original resection zone, while nodes marked to avoid are desynchronizing nodes that were resected. The predictions of our model need to be validated in a prospective trial before translating them into patient care. Additionally, other nodes may act as strong synchronizers during seizure evolution and perhaps could be especially targeted for ablation, resection or stimulation. Studying these findings broadly across connected brain regions could identify potential targets for focal therapy outside of the seizure onset.

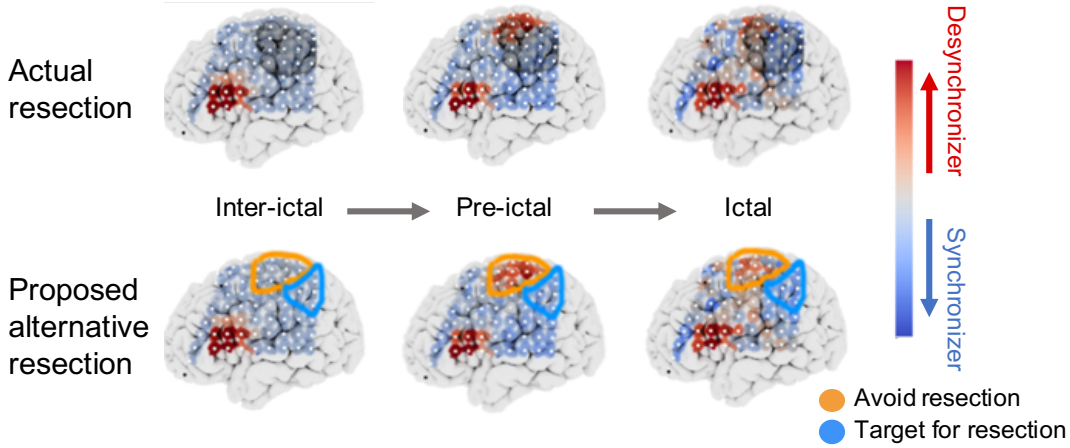


Figure 3.7. Proposed framework for optimal resection targets using virtual resection. Median node-level control centralities during interictal, pre-ictal and ictal periods are shown for a sample patient who had poor outcome. The top row shows spatial maps with the resection zone (darkened region with white outline) and the bottom row shows regions that perhaps should be targeted for resection due to the presence of strong synchronizing nodes (blue), or should be avoided during resection due to the presence of strong desynchronizing nodes (red). Future experiments assessing the effects of stimulation of these nodes and subsequent changes in synchronizability will allow clinicians to better predict the effect of targeted resection of these nodes.

The virtual resection method allows for mapping of control mechanisms of epileptic networks outside of seizures as well. The model can be extended to interictal periods without the need for ictal markings because repetitive, stable topographical patterns in functional connectivity emerge across long interictal time periods. Additionally, subgraphs identified in interictal periods are similar topologically to those identified during seizures[59], allowing for generalization of the virtual resection method to interictal epochs. Interictal epileptiform discharges may also be incorporated into the virtual resection framework by studying the regions that generate spikes in terms of their interictal, pre-ictal and ictal control centrality. The flexibility of our methods to various modes of analysis is a clear strength of our study.

The natural next step in testing the idea that resection should target synchronizers and preserve desynchronizers is to relate the effects of local stimulation to regional control centrality. Brain stimulation performed either intraoperatively or in the epilepsy monitoring unit may provide a safe and effective way to experimentally test network hypotheses by determining whether activation of certain nodes results in the generation or interruption of epileptic activity. Relating stimulation to virtual resection measures may be useful to fully describe spatiotemporal dynamics and would provide an avenue to formulate an algorithm to target resection. As recent technology allows intracranial EEG streaming to online cloud platforms[96] our virtual resection method and calculation of network metrics could also be implemented in real time during stimulus-based mapping. If successful, these methods could be extended to guide therapy with closed-loop neurostimulation devices such as the Neuropace RNS.

The methods and results of our model are derived from patient-specific ECoG and imaging data and compare favorably with previous studies of *in silico* models of resective epilepsy surgery. In particular, studies have often performed the identification of ictogenic nodes whose resection influences outcome by using neural mass models parameterized by functional connectivity. First, Sinha *et al.* (2017)[65] identified nodes that caused the network to transition the fastest into seizure dynamics via a subcritical Hopf bifurcation[97]. Second, Goodfellow *et al.* (2016)[66] used a more mechanistic model that identified nodes where removal would reduce epileptiform dynamics via saddle-nodes on a limit cycle bifurcation. In contrast, our current virtual resection study uses a network framework to directly describe the node level and global dynamics of each patient's seizures as they occur. Our

approach is thus not constrained by simulated seizures whose dynamics may be at odds with those observed in clinical data, and instead allows clinicians to assess intracranial EEG network properties derived from the very same data on which they currently base clinical decisions. Furthermore, our quantitative imaging pipeline and high quality dataset is larger than either of the previous studies, which bolsters the generalizability of our results. We also use data from each seizure and a corresponding preictal period for each patient in our study rather than just the first seizure or purely interictal data. Each of these advantages of our study brings us closer to a clinical tool and demonstrates the significant novelty of our work in the field of personalized network models of epilepsy.

Our study has several important limitations. The small number of patients tested decreases the statistical power of this study, and before the technique can be applied clinically, it must be tested on a broader range of epilepsy types. It also must be validated on stereotactic EEG recordings which are rapidly becoming the standard for many epilepsy centers worldwide as they use depth electrodes to sample broader brain networks yet require a less invasive implantation procedure. Another limitation of our network neuroscience approach is the sensor-level brain regions that we call 'nodes' are neither spatially discrete nor fixed in their locations across patients as intracranial electrodes are not implanted with connectivity studies in mind. One limitation of all studies using functional connectivity derived from ECoG is the sampling bias introduced by electrode placement, which may impact the reliability of summary statistics derived from network models. We have found that even upon 20-40% resampling of electrodes, network metrics including control centrality and synchronizability are reliable and perform in line with

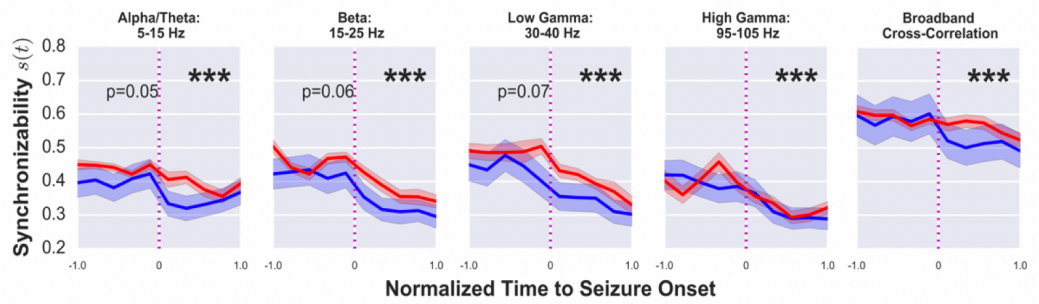
other metrics used in the field[71]. Furthermore, statistical methods such as jackknife resampling can be used to determine relative spatial confidence in network model results[71]. Unfortunately, we cannot validate whether the predictions regarding changes in synchronizability after resection result in an altered epileptic network that is less likely to manifest seizures. Direct brain recordings post-surgery would be an ideal method of model validation but these are not performed as part of clinical care at our institution.

A further nuance of our study is that all seizures were considered and analyzed for every patient. Patients undergoing electrode implantation and monitoring are frequently observed to have aberrant seizures and discharges attributed to electrode trauma as well as events occurring in regions that do not give rise to the patient's stereotyped clinical events[3,98,99]. The true causes of these seizures and their significance is unknown. Additionally, in our synchronizability analysis we average all seizures within a patient, even for mixtures of focal impaired awareness and bilateral tonic-clonic seizures. This approach diminishes sensitivity in detecting dynamic changes associated with different seizure subtypes. However, we feel that "human filtering" of data by selecting subtypes would add bias to our results and add a level of subjectivity to our methods that would make them quite difficult to translate to clinical practice, particularly at different medical centers using slightly different ECoG interpretation criteria. While this attests to the generalizability of the method, we need to further refine spatiotemporal mapping.

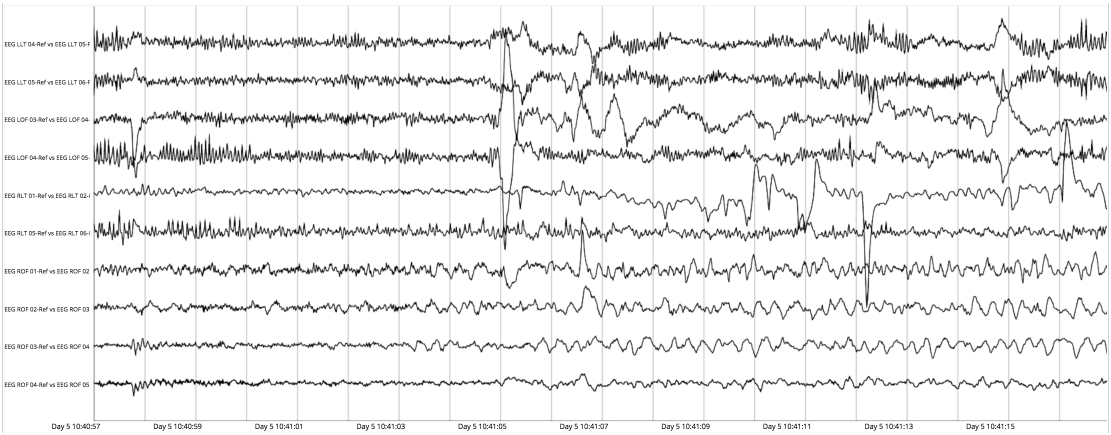
We present a method for rigorously mapping epileptic networks and predicting outcome based upon rigorously validated resection of network nodes. In spite of study limitations, our results suggest that these tools may have value in planning epilepsy

surgeries, identifying patients in advance who are likely to have good outcomes, and also those who are less likely to benefit from surgical resection. We hope that this work may be a step in further standardizing invasive epilepsy procedures and treatment, and initiating multi-center clinical trials that reduce individual variation from center to center in this vital part of patient care.

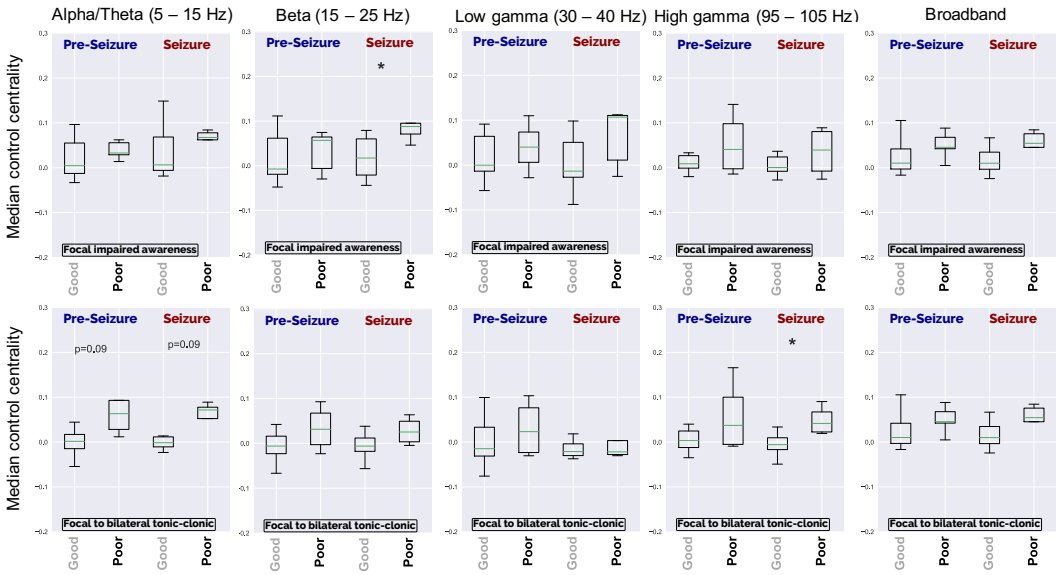
3.6 Supplemental materials



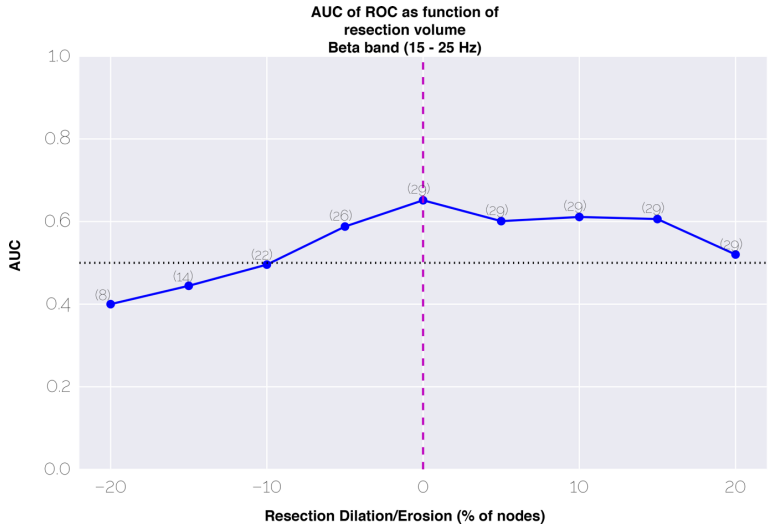
Supplemental Figure 3.1. Pre-resection synchronizability curves for all frequency bands. The above figure shows the time-varying distribution of s(t) during pre-ictal and ictal periods across different outcome groups and frequency bands after time-normalization to 10 time bins. Immediately following seizure onset, mean network synchronizability decreased in both outcome groups, with greater decrease in good outcome patients. With the use of functional curve analysis (10,000 permutations), all ictal events were significantly different ($p < 0.001$) in shape and time evolution.



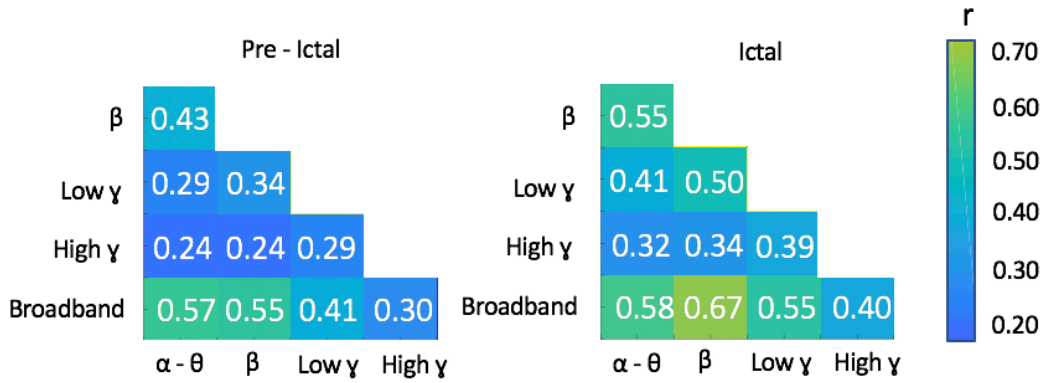
Supplementary Figure 3.2. HUP073 ictal pattern 60 seconds after EEC.



Supplementary Figure 3.3. Resection zone control centrality for all frequency bands. Patients with good surgical outcome have resection zones with lower control centrality, implying that tissue with greater synchronizing properties was resected. * = $p < 0.05$.



Supplementary Figure 3.4. Outcome prediction is robust to error in resection zone mapping. AUC of ROC curves are shown as a function of resection volume measured by percentage of network nodes removed. Segmentations were both dilated and eroded up to 20% of the network and resulting network measures were used to predict outcome. The numbers in parenthesis indicate the number of patients that still had electrodes left after erosion. AUC values remained high despite variance up to 10% in resected nodes. These results suggest fewer electrodes can be targeted for resection and potentially augur favorable outcome.



Supplementary Figure 3.5. Correlation of functional connectivity between frequency bands. Similarity values were generated by measuring Pearson correlation between edges in the functional connectivity matrix. The Pearson correlation coefficient values were averaged across time, seizures and patients. We considered the pre-ictal and ictal periods separately. We observe moderate correlations across frequency bands ($r = 0.24 – 0.67$) overall, with broadband networks more correlated to frequency specific bands than frequency specific bands to each other. Frequency specific bands were more highly correlated to adjacent frequency bands than bands much different in frequency. Average correlation across frequency bands in the ictal period was higher than in the pre-ictal period.

Feature	OR (95% CI)	*p* value
Δs(t)	32 (3.2-271)	0.00005
LVFA	1.5 (0.3-7.2)	0.71
Focal SOZ	4.3 (0.6-29)	0.17
DC shift	7 (0.7-68)	0.098

Supplementary Table 3.1. Odds ratios for associations with good surgical outcome. A board certified epileptologist determined the presence of low voltage fast activity, clearly focal SOZ and DC shift in all patients in our study, and we compared the odds ratio of the presence of each of these localizing EEG features with that of having a decrease in broadband synchronizability at seizure onset. P values determined by two-tailed Fisher exact test.

		SURGICAL OUTCOME	
	N = 28	**Good**	**Poor**
PREDICTED OUTCOME	**Good**	TP = 15	FP = 2
	Poor	FN = 2	TN = 9
	Accuracy = 0.86	*TPR = 0.88*	*TNR = 0.82*

Supplementary Table 3.2. Confusion matrix for virtual resection predictions. TP = true positives, FP = false positives, FN = false negatives, TN = true negatives, TPR = true positive rate, FPR = false positive rate.

CHAPTER 4: The effects of sampling bias on network models

4.1 Abstract

Focal epilepsy is a clinical condition arising from disordered brain networks. Network models hold promise to map these networks, localize seizure generators, and inform targeted interventions to control seizures. However, incomplete sampling of epileptic brain due to sparse placement of intracranial electrodes may profoundly affect model results. In this study, we evaluate the robustness of several published network measures applied to intracranial electrode recordings and propose an algorithm, using network resampling, to determine confidence in model results. We retrospectively evaluated intracranial EEG data from 28 patients who were implanted with grid, strip, and depth electrodes during evaluation for epilepsy surgery. We recalculated global and local network metrics after both randomly and systematically resampling subsets of intracranial EEG electrode contacts. We found that sensitivity to incomplete sampling varied significantly across network metrics, and that this sensitivity was independent of the distance of removed contacts from the seizure onset zone. We present an algorithm, using random resampling, to compute patient-specific confidence intervals for network localizations on both global and nodal network statistics. Our findings highlight the difference in robustness between commonly used network metrics and provide tools to assess confidence in intracranial network localization. We present these techniques as an important step toward assessing the accuracy of intracranial electrode implants and translating personalized network models of seizures into rigorous, quantitative approaches to invasive therapy.

4.2 Introduction

Epilepsy is a significant cause of disability worldwide, particularly among the one third of patients whose seizures cannot be controlled by medications[10]. While these patients may benefit from surgery or implanted devices, many continue to experience seizures after invasive therapies[1,86,100]. One reason for this persistence of seizure activity may be the difficulty in localizing seizure-generating brain regions, the drivers of complex epileptic brain dynamics.

Clinicians and scientists now agree that epilepsy is in part a disease of brain networks[47]. Driven by clinical observations, scientists have applied formal models from network theory to better understand seizure dynamics and target therapy[49]. In these models, the brain is discretized into regions represented by network nodes, while network edges are used to represent their structural or functional connectivity. Network theory applied to epilepsy employs a wide variety of metrics to understand seizure generation and control, including node strength[101], eigenvector centrality[54], betweenness centrality[53], clustering coefficient[102], and control centrality[4,55], as well as global metrics including global efficiency[103], synchronizability[4], and transitivity[104]. Collectively, these network measures have been used to predict neuronal firing as seizures begin and spread, track seizure progression, identify the seizure onset zone, and predict surgical outcome[53,54,65,105–107].

When using invasive sensors such as intracranial EEG (iEEG) to estimate functional connectivity, sampling from the full brain is impossible, and the network measures available for modeling depend on the location and number of electrodes implanted. In fields outside of epilepsy, missing data are known to affect the results of

network analyses[108–111]. The effect of missing data on network models and clinical care in epilepsy has not been rigorously explored. While network models have potential to add rigor to clinical decision-making, their application may be limited by uncertainty in estimated network metrics and the unknown interaction between that uncertainty and sparse brain sampling. In this study we seek to rigorously assess the extent to which different network metrics are sensitive to intracranial electrode sampling. Our goal is not to determine which, if any, network statistic correctly localizes the seizure onset zone or predicts surgical outcome, as this important work is currently under way by several groups[55,65,112,113]. Rather, our goal is to determine (a) whether and how incomplete spatial sampling affects the practical utility of network statistics, and (b) how sensitivity to spatial sampling can estimate patient-specific uncertainty in network model predictions. This computational work is a vital first step to deploying network models as an adjunct to clinical decision-making.

4.3 Materials and Methods

We use a high-quality dataset that has been included in multiple network studies in epilepsy[4,55] and is publicly available at www.IEEG.org. We randomly eliminate nodes from functional networks to simulate the uncertainty consequent to variable sampling of brain regions by iEEG and to determine the reliability of different network metrics within and across patients. Based upon the assumption that the main drivers of epilepsy network behavior might localize to an epileptogenic region, we ask to what extent electrode contacts far away from the seizure onset zone impact the estimated values of various network metrics, and whether subsampling that targets the seizure onset zone disproportionately

affects network statistics compared with subsampling that spares the seizure onset zone. We then randomly remove nodes by jackknife subsampling in order to derive patient-specific estimates of confidence in network statistics.

4.3.1 Patient Selection, Intracranial EEG Recording, and Electrode Localization

All patients gave written informed consent in accordance with the Institutional Review Board of the Hospital of the University of Pennsylvania (HUP) and the Mayo Clinic in Rochester. Furthermore, all patients consented to publishing their full-length iEEG recordings on the public web portal IEEG.org[70]. This study was performed in accordance with the Declaration of Helsinki.

A total of 28 patients with drug-resistant epilepsy underwent iEEG recording during presurgical evaluation at HUP or the Mayo Clinic. Electrode configurations (Ad Tech Medical Instruments, Racine, WI) consisted of linear cortical strips and two-dimensional cortical grid arrays (2.3-mm diameter with 10-mm intercontact spacing), and linear depth electrodes (1.1-mm diameter with 10-mm intercontact spacing). EEG signals were recorded at a sampling frequency of 512 Hz at HUP and 500 Hz at Mayo Clinic. All electrode and EEG recording systems were FDA approved and are commercially available.

Each patient underwent MPRAGE T1-weighted magnetic resonance imaging (MRI) on a 3T Siemens Magnetom Trio scanner (Siemens, Erlangen, Germany) prior to electrode implantation, and they also underwent spiral CT imaging (Siemens, Erlangen, Germany) after electrode implantation. We cross-referenced the CT images with a surgical cartoon map to localize electrode coordinates. To segment the resection zone, we registered the preimplant MRI to postresection imaging and the postimplant CT using the Advanced

Normalization Toolkit[114] (ANTs). We utilized a random forest classifier with ITK-SNAP to semiautomatically estimate the resection zone and identify electrodes overlying resected cortex[89].

Seizures were identified clinically and confirmed in a clinical case conference discussion. Board-certified epileptologists then reviewed the seizures and identified the earliest electrographic change[82] (EEC) and the electrode contacts of seizure onset (identified using the clinical standard for recognizing the electrode contact with the EEC) for each seizure. We performed our primary analysis on the first seizure identified for each patient. For patients with more than one seizure ($N = 26$), we also performed the analysis on the second seizure to assess the sensitivity of our results to the choice of seizure. For patients with three or more seizures ($N = 23$), we also performed the analysis on the patient's last seizure in order to evaluate more temporally distant seizures, given evidence that temporally clustered seizures may have similar dynamics, and given the possibility that earlier seizures may be atypical because of postimplantation effect[115,116]. One patient (HUP111) had two separate electrode implantations, and we analyzed both implantations separately.

4.3.2 Calculating Functional Networks

We examined 1-s time windows (sampled at 512 Hz at HUP and 500 Hz at Mayo Clinic) at each of the following time periods: 10 s prior to the EEC, 5 s prior to the EEC, at the EEC, 5 s after the EEC, and 10 s after the EEC. We chose 1-s time windows so as to have sufficient data to perform coherence calculations and because of the validation of this time window in prior publications[4,55,117]. To determine the sensitivity of our results to this

choice, we repeated this analysis with time windows of 2 s. We performed our primary analysis on the time period at the EEC given evidence for changes in network parameters that occur at the EEC[4,77]. We then repeated the analysis for each other time window in order to assess the sensitivity of our results to the choice of time period, and given the evidence that both interictal networks and post-EEC networks localize the seizure onset zone[54,113].

A common average reference was applied to iEEG signals to remove common sources of noise. Data were filtered using an elliptic bandpass filter with cutoff frequencies of 5 Hz and 115 Hz, as well as a 60-Hz notch filter to remove power line noise. Signals were pre-whitened using a continuous autoregressive model to account for slow dynamics and to accentuate higher frequencies known to be involved in seizure dynamics. This also enhanced local neural population dynamics in order to minimize the effect of signal mixing[77,118,119]. For each 1-s window, we constructed functional networks in which iEEG electrode contacts represented network nodes. Edges were weighted by multitaper coherence, which estimates the correlation between two electrode contact signals in the frequency domain and is frequently used to calculate functional networks in neuroscience publications[4,120]. We calculated coherence in the high gamma frequency band (95–105 Hz), which we chose because of its importance in seizure propagation and spread[4]. We also repeated the analysis in beta (15–25 Hz) to assess the sensitivity of our results to the choice of frequency band, and in acknowledgment of the fact that the beta frequency is also thought to be important in epileptic networks[121]. This separation of the data resulted in an adjacency matrix for each frequency band representing a network with undirected,

weighted edges for each patient, where each row and each column represented an electrode contact, and each matrix element represented the signal coherence between the two contacts.

To determine the sensitivity of our results to the choice of network density, we also performed weight-based thresholding in which we set matrix elements below a weight *w* to 0, where *w* was tuned for each patient to achieve a network density of 0.5 (in addition to the unthresholded network).

4.3.3 Network Metrics

For each functional network, we calculated several global and nodal network metrics, chosen because of their importance in graph theory and their use in recent epilepsy publications as described above. The global metrics were synchronizability, global efficiency, and transitivity. The nodal metrics were node strength, control centrality, clustering coefficient, eigenvector centrality, and betweenness centrality. The methods for calculating these metrics have been previously described, and we briefly summarize each below. We specifically describe their calculations for an undirected, weighted network. We calculated each using the Brain Connectivity Toolbox[122], or using custom code for synchronizability and control centrality[4].

Global efficiency is a global measure that is thought to represent how easily information travels throughout the network[123]. It is defined as

$$E = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{\sigma_{ij}},$$

where E is global efficiency, N is the number of nodes, and σ_{ij} is the shortest weighted path length between node i and node j, for example estimated using Dijkstra's algorithm[124]. A

high global efficiency is thought to reflect a greater ease of information transfer throughout the network. Path lengths were weighted by the inverse of the values of the adjacency matrix, to reflect the fact that information is thought to be transferred more readily along stronger edges[125].

Synchronizability is a global metric that quantifies the stability of the fully synchronous network state[6,126] and has been shown to predict seizure generalization[4]. It is calculated by first computing the weighted Laplacian $L = D - A$ as the difference between the node strength matrix D and the adjacency matrix A. Synchronizability is then obtained by the equation $Sync = \lambda_2/\lambda_{max}$, where $Sync$ is synchronizability, λ_2 is the second smallest eigenvalue of the Laplacian, and λ_{max} is the largest eigenvalue. Greater synchronizability reflects a smaller spread between eigenvalues, which suggests greater ease for a network to synchronize its dynamics.

Transitivity is another global measure that represents the degree to which nodes in a graph tend to cluster together[127–129]. It is defined as $T = \sum\tau_\Delta/\sum\tau$, where T is transitivity, $\sum\tau_\Delta$ is the sum of the weights of closed triplets, and $\sum\tau$ is the sum of the weights of all triplets. A triplet is defined as a set of three nodes connected by either two or three edges. A closed triplet, more specifically, is a set of three nodes connected by three edges. Higher transitivity implies that nodes tend to cluster together into exclusive groups.

Node strength represents the total strength of connections involving a particular node[130], and is defined as

$$s_i = \sum_{j=1}^{N} A_{ij},$$

in which s_i is the strength of node i, A_{ij} is the adjacency matrix element containing the edge weight between node j and node i, and N is the number of nodes. A high node strength implies that the total weight of its connected edges is large. Eigenvector centrality is a similar nodal measure that weights individual node influence by the relative influence of each of its connected nodes[106]. It is specifically defined as $\lambda = Ax$, where x is the vector containing the eigenvector centrality of each node, A is the adjacency matrix, and λ is the largest eigenvalue of the matrix (which results in nonnegative eigenvector centralities). A high eigenvector centrality implies that a node is strongly connected to nodes that themselves are highly connected.

Betweenness centrality is a nodal metric that is closely related to the global metric global efficiency and measures the fraction of all shortest paths in the network that pass through a given node[131]. It is defined as

$$b_i = \sum_{h \neq i \neq j} \frac{\sigma_{hj}(i)}{\sigma_{hj}},$$

where b_i is the betweenness centrality of node i, $\sigma_{hj}(i)$ is the number of shortest paths from node h to node j that pass through node i, and σ_{hj} is the total weighted path length between node h and node j. A high betweenness centrality suggests that the node acts as a central node in the shortest paths between many other nodes. The path lengths were weighted by the inverse of the values of the adjacency matrix as described above.

Control centrality is a local metric that measures the effect of each node on synchronizability. It is defined as

$$c_i = \frac{Sync_{new} - Sync_{old}}{Sync_{old}},$$

where c_i is the control centrality of node i, $Sync_{old}$ is the original synchronizability, and $Sync_{new}$ is the synchronizability of the network with the node removed[4]. Negative control centrality nodes are synchronizing, whereas positive control centrality nodes are desynchronizing.

Clustering coefficient is the nodal extension of transitivity that measures the amount of interaction between local triplets[132]. It is calculated by

$$cl_i = 2\sum_{k,j}\frac{(A_{ij}A_{kj}A_{ik})^{1/3}}{v(v-1)},$$

in which A is the adjacency matrix edge weight and v is the number of neighbors. Higher clustering coefficients reflect greater clustering of the node into tight groups.

4.3.4 Network Subsampling

To determine the sensitivity of network metrics to spatial sampling, we randomly identified electrode contacts for removal in each patient. We removed the rows and columns corresponding to these electrode contacts from the adjacency matrix representing the network. We recalculated each of the network metrics above. We performed this analysis removing 20%, 40%, 60%, and 80% of randomly selected electrode contacts. We repeated this process 1,000 times for each removal percentage to obtain a distribution of new metrics in the randomly subsampled network (Figure 4.1).

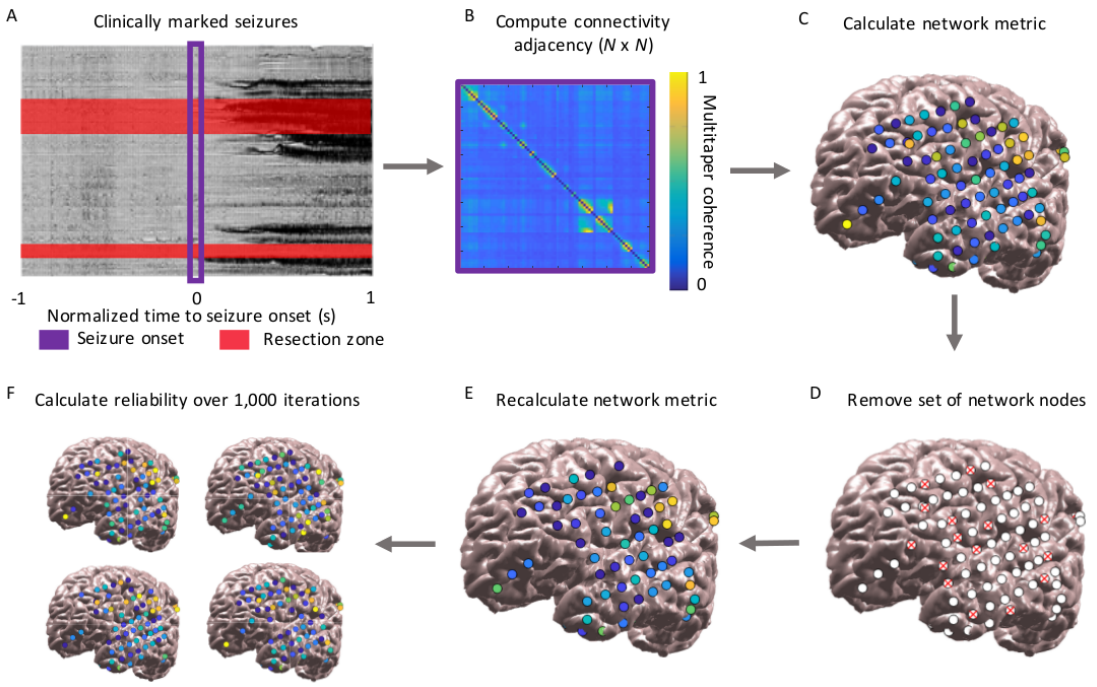


Figure 4.1: Network generation and resampling methods. A: Seizure onset times were marked by a board-certified epileptologist. Resected brain areas corresponding to EEG channels were determined through a semi-automated imaging technique and confirmed by a board-certified neuroradiologist. B: Multitaper coherence of a 1-second interval of EEG signal at seizure onset was used to create a functional adjacency matrix. C: Network metrics were calculated using the adjacency matrix with all nodes included. D: A subset of nodes were removed to simulate the effect of leaving out electrodes. E: Network metrics were recalculated from the resampled network. F: This process was repeated over 1,000 iterations and the reliability of each metric was quantified.

4.3.5 Determining Metric Reliability

To determine the stability of each network metric to subsampling, we calculated the reliability for each removal percentage[133]. Reliability is defined as

$$R = \frac{\sigma_T^2}{\sigma_X^2},$$

where $\sigma_X^2 = \sigma_T^2 + \sigma_E^2$, σ_T^2 is the variance of the true scores, σ_E^2 is the variance of the error, and σ_X^2 is the total variance. We defined the variance of the error to be the variance of the

subsampled metric across the 1,000 random subsamples, averaged across electrode contacts in the case of nodal metrics. For nodal metrics, we defined the variance in the true scores to be the variance of the subsampled metric across electrode contacts, averaged over all permutations. In the case of global metrics, we defined the variance in the true scores to be the variance in the subsampled metric across patients, averaged over all permutations. Reliability is constrained to be between 0 and 1, where 1 indicates that no variance is due to random subsampling, 0 indicates that all variance is due to random subsampling, and 0.5 indicates that the variance due to random subsampling equals the variance of the true metric. The goal of the reliability measure is to compare how much metrics vary across subsamples relative to how much they vary across patients (in the case of global metrics) or electrode contacts within the patient (in the case of nodal metrics). Lower reliabilities suggest that the variance across subsamples is higher than that across patients (global metrics) or electrode contacts (nodal metrics), suggesting that minor changes in electrode configurations could result in different orderings of highest-to-lowest metric values, thus decreasing our confidence in the result. Of note, we calculated the variance in the true scores in the subsampled networks, rather than the original network, to avoid a bias in which some network metrics (such as node strength) have larger values and larger variances across electrodes in larger networks, resulting in paradoxically *higher* reliabilities when removing more electrodes.

To determine whether some metrics were more robust to subsampling than others, we compared the metric reliability across all patients for the 20% removal percentage using separate Friedman tests, one for global metrics and one for nodal metrics[134] ($\alpha = 0.05$). In

the case of significant Friedman test results, we performed post hoc Dunn-Šidák multiple comparisons tests to identify significant differences between individual metrics[135,136]. We also determined the reliability of metrics for removal percentages other than 20%, which we report in our Supporting Information. We repeated this analysis for beta band coherence, alternate times relative to the EEC, removal of contiguous rather than random electrode contacts, alternate network densities, alternate time windows for calculating coherence, and different seizures, which we also report in our Supporting Information.

As an additional test of network stability to subsampling for nodal metrics, we calculated the Spearman rank correlation of the vector of nodal metrics across electrodes between the original network and each of the 1,000 subsampled metrics. We obtained the mean of the rank correlation across all 1,000 subsamples as a measure of the average correlation between the original set of nodal metrics and the subsampled metrics. We compared the metric rank correlations across all patients for the 20% removal percentage using a Friedman test ($\alpha = 0.05$), performing post hoc Dunn-Šidák multiple comparisons tests to identify significant differences between individual metrics in the case of significant Friedman test results.

We then determined whether there was a relationship between the network reliability and the number of electrode contacts comprising the original network. We obtained the reliability for each patient and each nodal and global metric at the 20% removal percentage of random electrode contacts, using the EEC time period and gamma band coherence. For each metric, we correlated the reliability with the original number of electrode contacts in the patient's network using Spearman rank correlation. We performed

Bonferroni correction as we were testing eight network metrics, yielding an α of 0.05/8 = 0.00625.

4.3.6 Influence of Seizure Onset Zone on Network Reliability

We next hypothesized that ictal network metrics may be more affected by removing electrode contacts near the seizure onset zone, as these contacts may have a stronger influence on epileptic networks. To test this, we again subsampled the network, this time systematically removing each electrode contact and its $N-1$ nearest neighbors, where N was equal to 20% of the total number of contacts in the network (we also calculated it for other removal percentages and report these results in our Supporting Information). We recalculated each of the global and nodal metrics in this systematically subsampled network. We obtained a measure of agreement between the original metric and the new metric in the subsampled network. For nodal metrics, the agreement measure a was defined as the Spearman's rank correlation coefficient across electrode contacts between the original and subsampled metric. For global metrics, the agreement measure was defined as the negative of the absolute value of the relative difference between the two metrics, represented by the equation

$$a = -\left|\frac{metric_{new} - metric_{old}}{metric_{old}}\right|.$$

The global agreement a was equal to 0 when there was perfect agreement between the new and original metric, and was increasingly negative with larger absolute differences.

To test whether there was larger metric agreement when the removed electrode contacts were further from the seizure onset zone, we obtained the Spearman's rank correlation coefficient between the agreement measure a with the distance between the

centroid of the removed electrode contacts and the centroid of the seizure onset zone. We obtained the Fisher's transformation of the rank coefficient for each patient, which is equal to the inverse hyperbolic tangent of the rank coefficient, in order to transform the coefficients to a variable whose distribution is approximately normal[137]. We aggregated these transformed rank coefficients across patients and performed a two-sided one-sample *t* test to determine whether the mean coefficient was significantly different from 0. We performed this test for each of the global and nodal metrics. We performed Bonferroni correction as we were testing eight network metrics, yielding an α of $0.05/8 = 0.00625$.

As an additional test of the hypothesis that removing seizure onset zone electrodes disproportionately affects network statistics, we performed two additional subsampling methods: a seizure onset zone-targeted subsampling and a seizure onset zone-sparing subsampling. In the seizure onset zone-targeted subsampling, we identified all electrodes forming the clinician-defined seizure onset zone and we removed all of these and only these electrodes. In the seizure onset zone-sparing subsampling, we identified a randomly selected subset of electrodes, equal in number to the number of seizure onset zone electrodes, but excluding the seizure onset zone (in one patient, Study022, the number of seizure onset electrodes was more than half of the total number of electrodes, and in this case we removed all other electrodes for the seizure onset zone-sparing subsampling). We repeated the seizure onset zone-sparing subsampling 1,000 times. For each subsampling, we again calculated the agreement, *a*, between the original and subsampled network statistics, where *a* is defined above for both global and nodal metrics. We took the mean agreement across all 1,000 subsamples in the case of seizure onset zone-sparing

subsampling. We compared the mean seizure onset zone-sparing agreement and the seizure onset zone-targeted agreement with a two-sided paired *t* test to determine whether the metric agreement when subsampling using a seizure onset zone-sparing method was significantly different from that using a seizure onset zone-targeted method. We performed this test for each of the global and nodal metrics. We performed Bonferroni correction as we tested eight network metrics, yielding an α of $0.05/8 = 0.00625$.

As an alternative approach, we also calculated for each patient the percentage of seizure onset zone-sparing agreements that were higher than the seizure onset zone-targeted agreement. We performed a one-sample two-sided *t* test to determine whether the mean percentile was significantly different from 50% (under the null hypothesis that if the seizure onset zone contacts were not of particular importance to the network metrics, half of patients would be expected to have higher seizure onset zone-sparing versus seizure onset zone-targeting agreements), using a Bonferroni correction for eight network metrics ($\alpha = 0.00625$). Of note, there were six patients for whom the number of seizure onset zone electrode contacts was large relative to the total number of electrode contacts (approaching half), and so for these patients there was likely a high interdependence between the 1,000 seizure onset zone-sparing subsamples. We expect that this makes this analysis less conservative than our primary analysis above.

We repeated the seizure onset zone analyses restricting analysis to patients with good (International League Against Epilepsy, ILAE = 1) outcomes ($N = 13$ patients), as it is possible that in the poor-outcome patients, the clinician-defined seizure onset zone did not accurately capture seizure generators. We also repeated these analyses using the

electrodes overlying the resected area of cortex, rather than the seizure onset zone, while restricting analysis to ILAE 1 outcome patients, under the assumption that the resected cortex in these good-outcome patients likely overlaps with seizure generators.

4.3.7 Deriving Patient-Specific Confidence in Network Results Using Jackknife Subsampling

We next utilized a jackknife subsampling method to generate patient-specific estimates in the confidence of the results of network analyses. Jackknife estimation is a method of sampling without replacement to derive statistical estimates[138–140]. It applies the same subsampling technique from our earlier analyses, but with the aim of obtaining patient-specific confidence rather than metric-specific reliabilities. Our goal was to determine how much a network result would be expected to change if a small number of electrode contacts had not been present. We randomly removed 20% of electrode contacts, recalculated the network statistic of interest for the random subsample, and repeated this process for 1,000 iterations. We chose a 20% removal percentage for this analysis to simulate minor variability in electrode implantation strategy. For each of the nodal metrics, we identified the electrode contact with the maximal metric value (minimal value for control centrality) in each of the 1,000 iterations. We identified the electrode contacts comprising 95% of all occurrences of the maximal metric value across the 1,000 iterations. We called this set of contacts the 95% jackknife confidence contact set. We also identified the 95% confidence contact set of the minimum *regional control centrality*, defined as the locations of an electrode contact and its $N - 1$ nearest neighbors, where N equals the number of resected electrode contacts that produces the largest negative change in synchronizability when

removed. Regional control centrality attempts to identify a region of a defined size—rather than a single electrode contact—with the largest control centrality, and thus a potential site for resection55. A larger 95% jackknife confidence set of electrode contacts implies greater sensitivity of the identity of the electrode contact with the maximal metric value to spatial subsampling, suggesting lower confidence in the patient-specific network result. For global metrics, we performed this method to obtain the 95% jackknife confidence interval for the value of the metric for a given patient, which was the interval containing 95% of all values obtained with jackknife subsampling. A larger 95% jackknife confidence interval for global metrics implies greater sensitivity of the global network statistic to spatial subsampling, suggesting lower confidence in the global network value. The runtime for the jackknife subsampling algorithm (1,000 iterations) for all metrics at a single time and frequency band was approximately ten minutes per patient when performed in MATLAB R2018a on an Intel Xeon processor (CPU E5-2698 v3 @ 2.30 GHz).

4.3.8 Statistical Analysis

All analyses were performed on MATLAB R2018a (MathWorks, Natick). Specific analyses are discussed in the four preceding sections, above.

4.4 Results

4.4.1 Patient and Electrode Information

Patients had a variety of clinical histories, electrode configurations, pathologies, and clinical outcomes (Supplemental Table 4.1). There were 28 patients (13 women), one of whom had two temporally distinct implantations, which were separately analyzed. The

average age at implantation was 33.9 years (range 5–57). The mean number of electrode contacts was 77 (range 16–118). The mean number of seizures was 6.8 (range 1–36). The median ILAE outcome score at 2 years was 2 (range 1–5).

4.4.2 Stability of Metrics to Random Subsampling

For all network measures, reliability to subsampling decreased as more electrode contacts were removed. The stability of network measures to subsampling varied across patients (Figure 4.2). The mean reliability was $R = 0.92$ for synchronizability, $R = 0.98$ for global efficiency, and $R = 0.98$ for transitivity, averaged over all patients when a random sample of 20% of electrode contacts was removed. In contrast, when a contiguous sample of 20% of electrode contacts was removed, the mean reliabilities were lower, with $R = 0.85$ for synchronizability, $R = 0.92$ for global efficiency, and $R = 0.93$ for transitivity. The reliability to random electrode contact removal was significantly different between global metrics (Friedman test: $\chi 2 = 36.5$, $p < 0.001$). Synchronizability was significantly less reliable than either global efficiency or transitivity (post hoc Dunn-Šidák multiple comparison test: $t = -3.02$, $p = 0.008$ compared with global efficiency and $t = -6.04$, $p < 0.001$ compared with transitivity). Global efficiency was also significantly less reliable than transitivity ($t = -3.02$, $p = 0.008$). The reliability for global efficiency was slightly lower than that for transitivity for 26 out of 29 patient implantations. However, for two implantations the reliability for global efficiency was substantially larger, explaining why global efficiency and transitivity have similar means despite global efficiency having significantly lower reliability by ordinal statistics.

When we examined the time periods 10 s before, 5 s before, 5 s after, and 10 s after the EEC (as opposed to the second at the EEC), synchronizability continued to have the lowest reliability of the global metrics. Control centrality continued to have the lowest reliability of the nodal metrics, and eigenvector centrality and node strength continued to have the highest reliabilities. The pattern remained when we examined beta frequency coherence rather than high gamma frequency coherence, when we removed contiguous as opposed to random sets of electrode contacts, when we examined the second seizure or the last seizure rather than the first seizure, when we thresholded the network weights to achieve a network density of 0.5, and when we used a 2-s time window for coherence calculations rather than a 1-s window (Supplemental Table 4.2). When we instead removed 40% or 60% of electrode contacts, control centrality and synchronizability continued to have the lowest reliability of nodal and global metrics, respectively, and node strength and eigenvector centrality continued to have the highest nodal metric reliabilities. When we removed 80% of electrode contacts, clustering coefficient instead demonstrated the lowest nodal metric reliability, and otherwise the pattern was unchanged (Supplemental Table 3).

For nodal metrics, the Spearman rank correlation between the original and subsampled met ric revealed the same ranking of metric robustness as with our primary approach. The mean Spearman rank correlation between the original metric and the subsampled metric, averaged across all subsamplings and all patients when 20% of electrode contacts were randomly removed, was $\rho = 0.84 \pm 0.08$ for control centrality, $\rho = 0.97 \pm 0.02$ for node strength, $\rho = 0.91 \pm 0.03$ for betweenness centrality, $\rho = 0.98 \pm 0.02$ for eigenvector centrality, and $\rho = 0.97 \pm 0.02$ for clustering coefficient. The Spearman

rank correlation significantly differed across nodal metrics (Friedman test: $\chi 2 = 99.5$, $p < 0.001$). The Spearman rank correlation between the original and subsampled control centrality metric was significantly lower than that for node strength (post hoc Dunn-Šidák test: $t = -5.81$, $p < 0.001$), eigenvector centrality ($t = -8.64$, $p < 0.001$), and clustering coefficient ($t = -5.98$, $p < 0.001$). The Spearman rank correlation between the original and subsampled node strength metric was significantly higher than that for betweenness centrality ($t = 4.24$, $p < 0.001$) and significantly lower than that for eigenvector centrality ($t = -2.82$, $p = 0.046$). The Spearman rank correlation between the original and subsampled betweenness centrality metric was significantly lower than that for eigenvector centrality ($t = -7.06$, $p < 0.001$) and for clustering coefficient ($t = -4.40$, $p < 0.001$). The comparisons between control centrality and betweenness centrality ($t = -1.58$, $p = -0.704$), node strength and clustering coefficient ($t = -0.17$, $p = 1.00$), and eigenvector centrality and clustering coefficient ($t = 2.66$, $p = 0.076$) were not significant.

We next examined the relationship between robustness to electrode contact subsampling and the number of electrode contacts in the original network. Among global measures, there was a significant positive relationship between reliability and number of contacts for synchronizability (Spearman rank correlation: $r_{27} = 0.68$, $p < 0.001$), but not for global efficiency ($r_{27} = 0.20$, $p = 0.299$) or transitivity ($r_{27} = 0.22$, $p = 0.260$). Among nodal measures, clustering coefficient ($r_{27} = 0.53$, $p = 0.003$) demonstrated a significant positive relationship, and relationships for control centrality ($r_{27} = -0.12$, $p = 0.551$), node strength ($r_{27} = 0.45$, $p = 0.015$), betweenness centrality ($r_{27} = 0.38$, $p = 0.043$), and eigenvector centrality ($r_{27} = 0.44$, $p = 0.016$) were nonsignificant ($\alpha = 0.00625$, Bonferroni

correction for eight measures). This pattern of findings suggests that, at least for synchronizability and clustering coefficient, patients with more electrode contacts implanted were less vulnerable to incomplete spatial sampling.

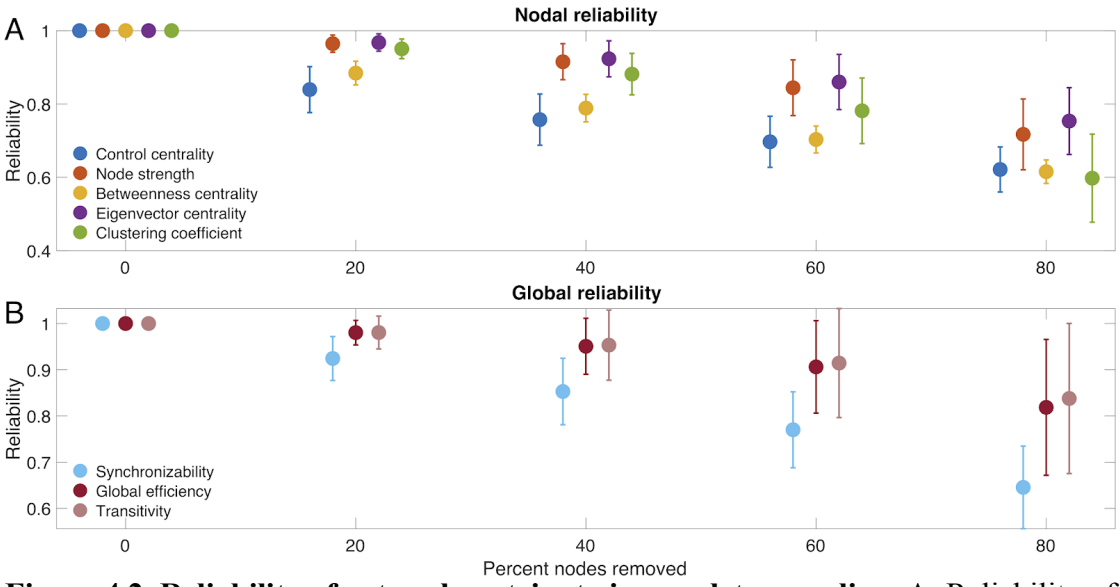


Figure 4.2. Reliability of network metrics to incomplete sampling. A: Reliability of nodal measures, averaged across all patients, when different percentages of electrodes were removed. B: Reliability of global measures, averaged across all patients, when different percentages of electrodes were removed. Error bars show the standard deviation of the reliability across patients. All data shown is for the EEC of first seizure, high gamma coherence, and random electrode removal. For all measures, reliability decreased as more electrodes were removed. Patients are heterogeneous in the reliability of their network measures, and certain network measures exhibit higher reliability than others.

4.4.3 Influence of Seizure Onset Zone on Sensitivity of Network Statistics to Subsampling

There was no significant association between metric agreement and distance of the removed electrode contacts from the seizure onset zone for any metric (one-sample two-sided t test: control centrality, $t = 0.80$, $p = 0.433$; node strength, $t = 1.25$, $p = 0.222$; betweenness centrality, $t = -0.95$, $p = 0.352$; eigenvector centrality, $t = 1.02$, $p = 0.318$;

clustering coefficient, $t = 1.23$, $p = 0.230$; synchronizability, $t = -0.26$, $p = 0.793$; global efficiency, $t = 0.74$, $p = 0.469$; transitivity, $t = 0.37$, $p = 0.717$). This pattern of findings implies that all metrics are equally sensitive to removing electrode contacts near versus distant from the seizure onset zone (Figure 4.3). This result was invariant to the choice of peri-ictal time window, choice of frequency band, choice of seizure, exclusion of non-ILAE 1 outcome patients, use of resection zone rather than seizure onset zone (excluding non-ILAE 1 patients), choice of network density, and choice of time window (Supplemental Table 4.4) as well as to the choice of removal percentage (Supplemental Table 4.5).

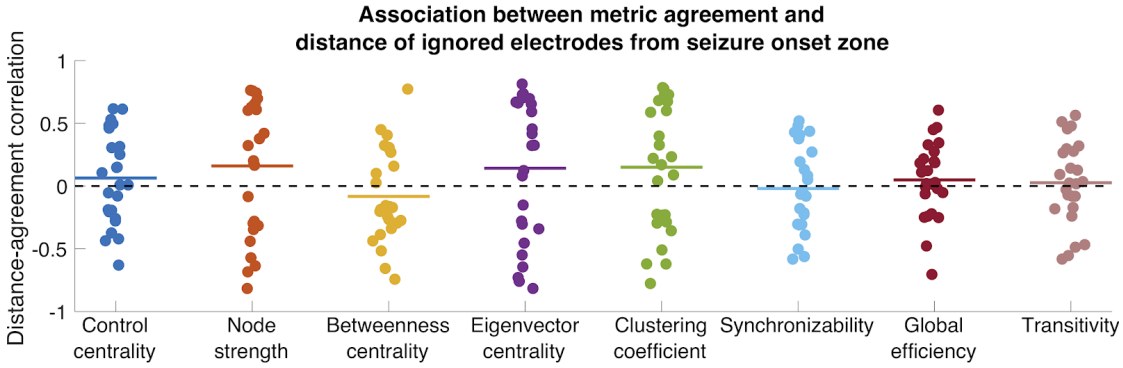


Figure 4.3. Association between metric agreement and distance of ignored electrodes from the seizure onset zone. Each nodal and global metric is shown. Each point represents the patient-specific Spearman rank correlation coefficient between the metric agreement and the distance of the ignored electrodes from the seizure onset zone. The metric agreement is defined for nodal metrics as the Spearman rank correlation coefficient between the original metric and the metric obtained from resampling, and for global metrics as the negative absolute value of the relative difference between the original and resampled metric. Horizontal lines show the average distance-agreement association across patients. No distance-agreement association was significantly different from zero (two-sided one-sample *t*-test, $\alpha = 0.05/8$. Bonferroni correction), signifying that all metrics are equally vulnerable to incomplete sampling near versus distant from the seizure onset zone.

When we compare metric agreement removing *all* seizure onset zone electrode contacts as opposed to removing only *non*-seizure onset zone electrode contacts, there was

again no significant difference in metric agreement between the seizure onset zone-sparing and seizure onset zone-targeted approach for any metric (paired two-sided *t* test: control centrality $t = -0.99$, $p = 0.331$; node strength, $t = 0.84$, $p = 0.408$; betweenness centrality, $t = 1.43$, $p = 0.167$; eigenvector centrality, $t = 0.81$, $p = 0.424$; clustering coefficient, $t = 1.10$, $p = 0.283$; synchronizability, $t = 2.33$, $p = 0.028$; global efficiency, $t = 1.74$, $p = 0.095$; transitivity, $t = 2.65$, $p = 0.014$; $\alpha = 0.00625$, Bonferonni correction for eight metrics). These findings suggest that sparing versus targeting seizure onset zone electrode contacts for removal has equivalent effects on most network statistics. Across conditions, transitivity generally displayed the largest differences between seizure onset zone-sparing agreement and seizure onset zone-targeted agreement, although these differences were significant only for the EEC + 10-s condition, the second seizure, and 50% network density (Supplemental Table 4.6). To determine whether the trend in transitivity perturbation was to increase or decrease network transitivity, we obtained the signed relative difference (as opposed to the metric ageement, which is unsigned) between the subsampled and the original transitivity measure, and compared this for the seizure onset zone-sparing and seizure onset zone-targeted subsampling methods. The relative difference in transitivity was nonsignificantly higher (more positive) when seizure onset zone-targeted subsampling was performed, which persisted across all time periods, frequency bands, and the second seizure (Supplemental Table 4.6, two-sided paired *t* test, $\alpha = 0.00625$, Bonferonni correction for eight metrics). This suggests that removing seizure onset zone electrode contacts may disproportionately *increase* transitivity (although this result was nonsignificant for most conditions).

When we used the approach calculating the percentage of seizure onset zone-sparing agreements larger than the seizure onset zone-targeted agreement, we found that the seizure onset zone-sparing agreement was significantly higher than the seizure onset zone-targeted agreement for both synchronizability (one sample two-sided t test, $t = 3.72$, $p = 0.001$) and transitivity ($t = 3.04$, $p = 0.005$) with no significant difference for other measures (Supplemental Table 4.6). The directions of these results were the same as those seen in the above analysis. However, given the dependence between seizure onset zone-sparing subsamples discussed in the methods section, we believe that this analysis is less conservative than our primary analysis above.

4.4.4 Jackknife Confidence Intervals

Both nodal and global metrics varied across patients with respect to jackknife confidence intervals produced by subsampling (Figure 4A–C). The median and range for the number of electrode contacts accounting for 95% of all jackknife instances of the maximum nodal metric (minimum for control centrality) was 3 (range 2–5) for node strength, 4 (3–9) for betweenness centrality, 3 (2–5) for eigenvector centrality, 3 (2–5) for clustering coefficient, and 9 (4–22) for control centrality. The median number of electrode contacts accounting for 95% of all jackknife instances of the minimum regional control centrality (where the set of electrode contacts with minimum regional control centrality is the set, equal in number to the number of resected electrode contacts, that together produces the largest negative change in synchronizability when removed) was 48 (range 12–93). The median ratio between this number and the number of electrode contacts forming the true minimum regional control centrality set was 4.0 (range 1.6–16.0; Figure 4.4D). Regarding

global metrics, the median width of the 95% jackknife confidence interval was 0.094 (range 0.045–0.192) for synchronizability, 0.016 (range 0.006–0.058) for global efficiency, and 0.012 (range 0.004–0.062) for transitivity (Figure 4.4E). These results demonstrate the heterogeneity among patients in the level of confidence in estimated network statistics that can be revealed by the jackknife algorithm. The locations of electrode contacts with the maximum or minimum metric values, as well as the results of jackknife subsampling, varied somewhat across time periods, choice of frequency band for coherence, seizure, network density, and time window for coherence calculations (Supplemental Figure 4.1, Supplemental Table 4.7).

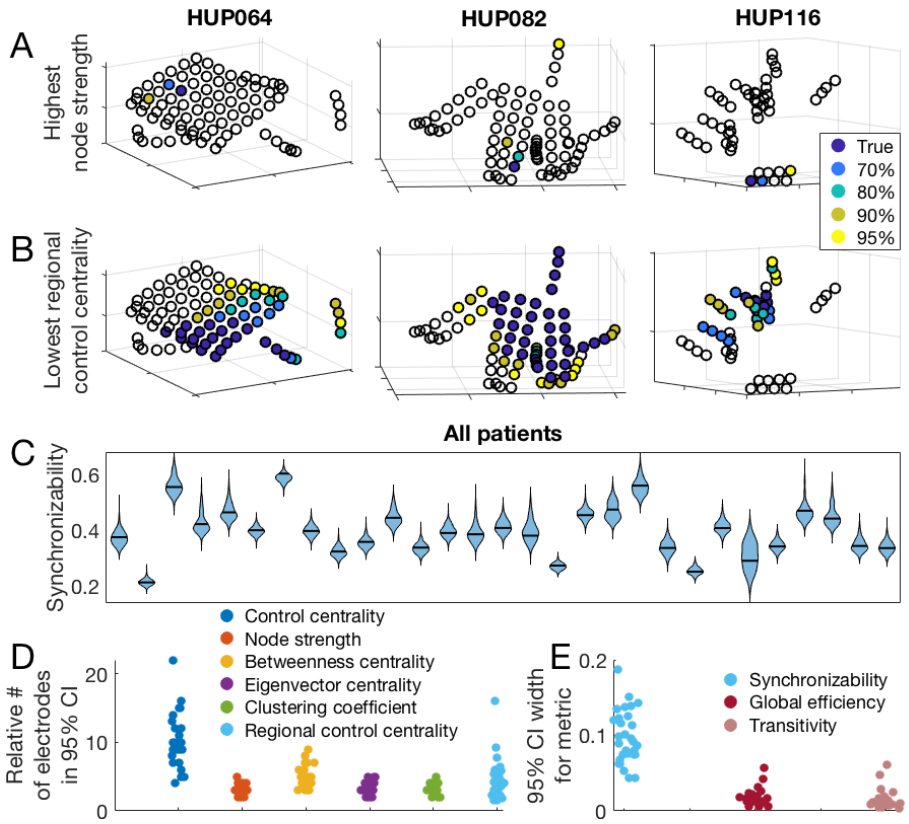


Figure 4.4. Jackknife resampling to estimate confidence regarding network metric values. A: The location of the electrode with the highest node strength, as well as the electrodes accounting for various percentages of highest node strength occurrences in 1,000 random jackknife resampling networks for three example patients. B: The location of the most synchronizing region (which is the region with the lowest regional control centrality), and the regions accounting for various percentages of the most synchronizing region occurrences in 1,000 random jackknife resampling networks for three patients. C: Patient-specific synchronizability distributions across resamples. Each separate violin represents a patient and shows the distribution of synchronizability values obtained across 1,000 random jackknife resamples. Horizontal black lines show the original value in the non-resampled network. D: Number of electrodes forming the 95% jackknife confidence electrode set for each nodal metric, for all patients. For each nodal metric, each dot shows the patient-specific number of electrodes accounting for 95% of all occurrences of the maximum (minimum for control centrality and regional control centrality) metric value in 1,000 random jackknife resampling networks. For regional control centrality, this number is divided by the number of electrodes forming the minimum regional control centrality in the original non-resampled network to obtain a ratio. E: Width of the 95% jackknife confidence interval for each global metric, for all patients. For each global metric, each dot shows the patient-specific width of the 95% jackknife confidence interval of the metric value across 1,000 random resamples. This figure demonstrates the variability in confidence of network theory results across patients that can be revealed by jackknife resampling.

4.5 Discussion

Handling missing data is a long-standing problem in science in general and is particularly problematic in network science, where a missing node may limit our understanding of the entire network[110]. In social networks, missing data can dramatically alter network statistics[109,141,142]. In the field of neuroscience, Jalili demonstrated that global efficiency of scalp EEG-based functional networks in healthy individuals was highly sensitive to the removal of certain nodes[143]. To our knowledge, this is the first study examining the reliability of network statistics in the epileptic brain and in iEEG data. We determined that network measures differ in robustness to spatial subsampling, and that the sensitivity to sampling does not depend on the distance from the seizure onset zone. We also found that more extensive implants were more robust to subsampling. Finally, we developed and applied an algorithm using jackknife subsampling of electrode contacts to estimate confidence in nodal and global statistics in patient brain networks.

4.5.1 Functional Network Metrics Exhibit Differential Reliability Under Spatial Subsampling

Metric reliability for all network measures decreased with a greater degree of missing data, which has previously been reported in social networks[142,144]. Among examined nodal metrics, node strength and eigenvector centrality were most reliable and control centrality was least reliable; among the global metrics we tested, transitivity was most reliable and synchronizability was least reliable. The difference in reliability across metrics reflects, in part, the underlying sensitivity of each metric to graph topology. Prior studies in social networks have also observed that node strength is more robust to

subsampling than betweenness centrality[144,145]. The relative robustness of node strength and eigenvector centrality compared with other nodal measures may reflect that metrics that primarily incorporate immediate connections to the node of interest are less sensitive to subsampling than metrics that more strongly weigh multistep connections. The preserved ordinality of network metric reliability across most patients, timescales, and frequency bands suggests that this result is generalizable. Clinically, applying network statistics that are more robust to spatial sampling may be preferable in cases in which the electrode coverage of important regions is uncertain. The ability of each metric to capture network behavior must be weighed against its spatial reliability if such personalized models are to be translated clinically.

Sensitivity to incomplete sampling depends somewhat on the number of electrode contacts forming the original implantation. Although synchronizability and clustering coefficient were the only global and nodal measures, respectively, to demonstrate a significant positive relationship between number of electrode contacts and metric reliability, all other measures except control centrality demonstrated nonsignificant positive relationships. This pattern of findings suggests that for most network measures, greater robustness can be achieved in part through more extensive electrode coverage. This agrees with work finding that more extensive electrode coverage results in better predictions of surgical outcome[146]. Mechanistically, this may imply that random subsampling is less likely to remove important hubs in larger networks. Alternatively, perhaps information about missing nodes and edges can be inferred from the remaining components of the network, which is facilitated by a larger starting network. Clinically,

this suggests that implanting a greater number of electrode contacts may increase our confidence in network statistics. This benefit would have to be weighed against the risks of more extensive coverage, including hemorrhage and infection[147]. Network metrics were generally more sensitive to contiguous removal than to random removal of electrodes. This may reflect spatial correlation of brain signals, such that information from a missing electrode contact may be more easily inferred from remaining neighbor contacts[51,148]. Alternatively, removing contiguous contacts may increase the probability of removing the entire set of critical electrode contacts that are needed to localize seizure generators. The analysis removing contiguous electrode contacts may better approximate the clinical scenario of leaving out an entire electrode or region from sampling, whereas the analysis removing random contacts better simulates choosing sparser network coverage. Consideration of which type of "missed coverage" applies in a given clinical context will inform how strongly that missing coverage may affect network results.

One limitation of the reliability measure as an approximation of metric robustness to subsampling is that nodal network metrics and their variances across subsampling are not normally distributed, and so the reliability measure may be disproportionately influenced by electrode contacts with more extreme values. The fact that nodal metrics had the same differential ordering in robustness to subsampling as measured by Spearman rank correlation between the original and subsampled metric supports the validity of the ordering of nodal metric reliabilities.

4.5.2 Metric Sensitivity to Incomplete Sampling Is Independent of Distance From the Seizure Onset Zone

All metrics tested were equally sensitive to removing nodes in close or far proximity from the clinician-defined seizure onset zone. Most metrics were also equally sensitive to subsampling that targeted removal of the complete seizure onset zone and subsampling that spared the seizure onset zone. This may be because the seizure-generating network is a relatively small subset of the full peri-ictal network, and thus perturbation of the seizure-generating network has a small effect on the network as a whole. These results are practically concerning in the case of metrics with lower reliability to spatial subsampling, such as control centrality and synchronizability, because placement of electrodes in the network periphery away from clinical zones of interest is often variable across patients and across epilepsy centers. To increase the clinical confidence in the results of these network measures, the incomplete network may be supplemented using structural connectivity data and atlas-based approaches[148–150]. Network theory also proposes several methods of predicting missing links[110,151]. The finding of nonsignificantly increased sensitivity of the transitivity measure to removing the entire seizure onset zone, along with a disproportionate *increase* in transitivity when seizure onset zone electrodes are removed, may reflect a tendency of nodes in the seizure onset zone to form widespread network connections peri-ictally. Practically, this finding also suggests that transitivity—and more specifically, subsampling-induced change in transitivity—is a promising measure to identify electrodes overlying the seizure onset zone. Of note, the seizure onset zone-targeted versus sparing subsampling method introduces a potential bias in that removing

the seizure onset zone as opposed to random non-seizure onset zone electrodes disproportionately targets electrodes that are spatially clustered together. If anything, we would expect this bias to act to decrease transitivity when seizure onset zone electrodes are removed, as we might expect electrodes in close proximity to have a higher within-group clustering and thus increase the overall transitivity of the network.

A limitation of this analysis is that clinical methods for identifying the seizure onset zone are imperfect, and so the clinician-defined seizure onset zone may not capture actual seizure generators. When redefining the hypothesized site of seizure generation as the resection zone (for good-outcome patients) we also found no difference in removing nodes in close or distant proximity from the resection zone. However, it is possible that neither method accurately captured true seizure generators.

4.5.3 Jackknife Network Subsampling Generates Confidence Intervals for Virtual Resection

Prior work has used global and nodal network measures to stratify surgical candidates and to select nodes for resection[52,60,65,93]. Here we provide a simple algorithm to augment these methods by determining patient-specific confidence in the robustness of estimated statistics to small perturbations in spatial sampling. Similar resampling-based approaches have been used in social networks[152–154], in gene expression data[155], and in resting fMRI data of healthy subjects. The heterogeneity in confidence across patients may be used to stratify patients into those for whom enough information of the epileptic network is captured to accurately use personalized network models and those for whom models are likely to be inaccurate because of implantation strategy. In this study we do not claim to

identify the ideal network model among the many published studies. The jackknife subsampling method may be easily applied to any of these network models.

4.5.4 Methodological Limitations and Future Directions

Our method of network subsampling can examine how networks change only with the removal of electrode contacts, and not upon the addition of contacts. For an intracranial EEG study that has parsimoniously captured seizure onset and spread, using our statistical subsampling method may miss critical electrode contacts and thus erroneously characterize the network as unreliable. However, for those patients who have very clear seizure onset and propagation, personalized network models of epilepsy may not be required to guide surgical planning. An additional limitation is that our measure of nodal metric similarity—the Spearman rank correlation—may not capture how network model predictions practically change with subsampling. It is conceivable that a network metric would demonstrate low reliability to subsampling as defined in this paper, but still generate consistent predictions about the optimal site of surgical intervention depending on how these predictions are formed. Different groups have proposed different methods to direct surgical targeting using network statistics[55,65,73,113], and our proposed subsampling method can be used to test the sensitivity of any targeting predictions to incomplete spatial sampling, similar to the jackknife procedure described in the section above.

The jackknife analysis results varied somewhat across time, frequency band, and choice of seizure, reflecting the different states of the network. These observations underscore the fact that spatial sampling is one of several sources of bias of the network statistics. Further exploration of the sensitivity of network statistics to the choice of seizure,

time, and frequency band will be necessary for useful clinical implementation. Also, additional steps toward translating this work into clinical care will require expanding our dataset to include larger numbers of patients, stereo EEG implantations, and those treated with focal laser ablation and perhaps brain stimulation devices.

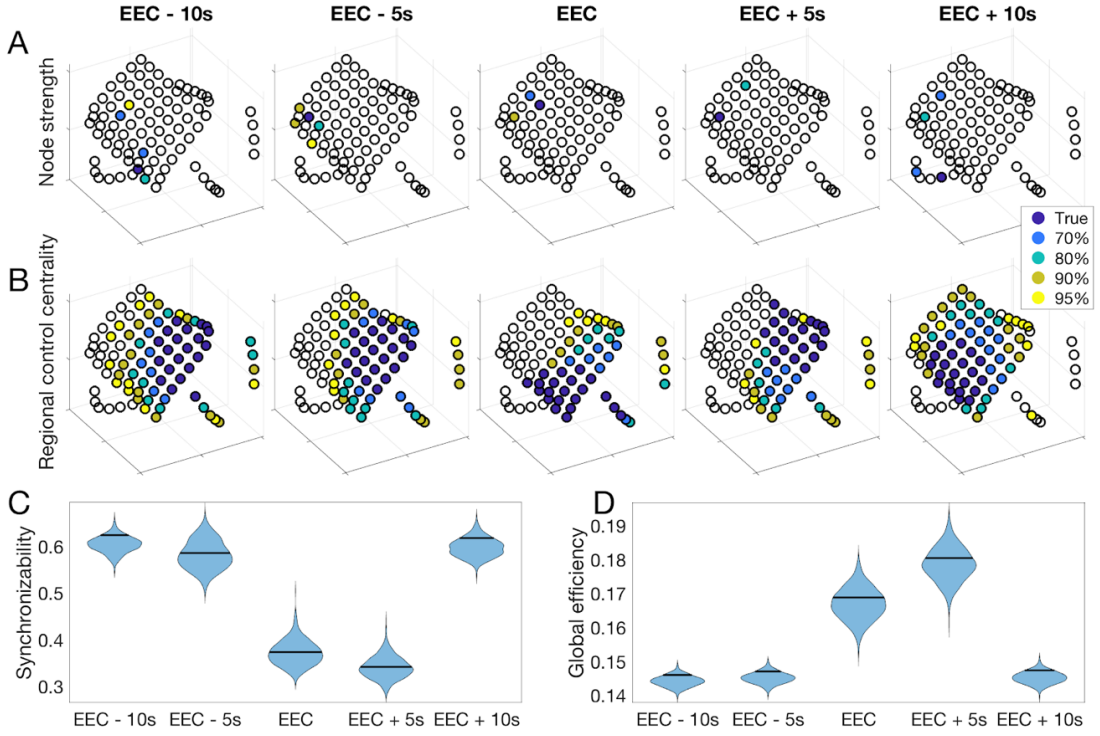
While in this study we implemented only data-driven metrics that describe underlying network properties, there is significant interest in fitting generative models of neural population dynamics to brain networks. One such example, the Epileptor model[81], is a neural mass model that describes many relevant epileptic dynamics and is currently under study as a clinical trial in Europe to guide epilepsy surgery[73]. Our network subsampling approach may also be used for generative neurophysiologic models and gives confidence to their clinical utility.

4.6 Conclusions

The field of network science provides a promising set of tools for understanding epilepsy dynamics and for surgical planning. However, the robustness of empirical estimates of network statistics to incomplete electrode sampling is not well understood. We have shown variability across network measures in robustness to incomplete sampling. Network measures are equally vulnerable to removing electrode contacts near versus distant from the seizure onset zone, and to removing electrode contacts within the seizure onset zone versus an equivalent number of non-seizure onset zone electrode contacts. Robustness to incomplete sampling is highly heterogeneous across patients, and jackknife subsampling is a simple algorithm to obtain patient-specific confidence in the results of network statistics. The choice of individual network models should be based upon the

intended application and on the clinical certainty that important seizure generators have been sampled by intracranial implants.

4.7 Supplemental materials



Supplemental Figure 4.1. Time dependence of jackknife resampling results. All data shown is for a single example patient (HUP064). A: The location of the electrode with the highest node strength, as well as the electrodes accounting for various percentages of highest node strength occurrences in 1,000 random jackknife resampling networks at five different time periods. B: The location of the most synchronizing region, and the regions accounting for various percentages of the most synchronizing region occurrences in 1,000 random jackknife resampling networks at five different time periods. C: The distribution of synchronizability values obtained with 1,000 random jackknife resamplings for each time period. Horizontal lines denote synchronizability in the original non-resampled network. D: The distribution of global efficiency values obtained with 1,000 random resamplings for each time period. Horizontal lines denote global efficiency in the original non-resampled network.

Patient ID	Sex	Age onset	Age at surgery	Localization	Pathology	Outcome	Grids	Strips	Depths
HUP064	M	3	22	LFL	MCD	1	64	22	0
HUP065	M	2	36	RTL	MCD	1	64	0	0
HUP068	F	13	28	RTL	MTS	1	63	20	0
HUP070	M	12	33	LFPL	MCD	2	63	0	0
HUP073	M	5	40	RFL	MCD	1	0	52	0
HUP074	F	5	25	LTL	MCD	1	63	29	22
HUP075	F	52	57	LTL	MTS	5	58	34	14
HUP078	M	2 mo	54	LTL	MTS	4	63	24	14
HUP080	F	35	41	LTL	Gliosis	2	62	18	16
HUP082	F	34	56	RTL	MTS	1	64	14	8
HUP083	M	8	29	LPL	Gliosis	2	59	22	0
HUP086	F	18	25	LTL	Gliosis	1	60	32	8
HUP087	M	19	24	LFL	T/V/I	3	60	12	12
HUP088	M	13 mo	24	LFL	MTS	1	0	43	11
HUP094	F	20	48	RTL	Gliosis	2	0	64	20
HUP105	M	27	39	RTL	T/V/I	1	0	43	12
HUP106	F	24	45	LTL	MTS	2	64	36	16
HUP107	M	5	36	RTL	MTS	1	64	38	16
HUP111A	F	28	40	RTL	MTS	1	0	32	16
HUP111B	F	28	40	RTL	MTS	1	47	42	12
HUP116	F	51	58	RTL	unknown	1	0	0	50
Study012	M	23	37	RFL	Gliosis	1	57	22	0
Study016	F	5	36	RFTL	Gliosis	4	47	14	0
Study017	M	?	?	RTL	unknown	4	0	0	16
Study019	M	31	33	LTL	Gliosis	5	60	28	8
Study020	M	5	10	RFL	Gliosis	4	40	16	0
Study022	F	11-20	21-30	LTL	Gliosis	5	42	12	0
Study028	M	4	5	LFPL	Gliosis	4	64	5	0
Study029	F	?	?	RTL	Gliosis	5	23	30	8

Supplemental Table 4.1. Clinical and electrode information. 2-year surgical outcomes use the International League Against Epilepsy (ILAE) system (class 1-6). LFL: left frontal lobe, LFPL: left frontoparietal lobe, LTL/RTL: left/right temporal lobe, RFTL: right frontotemporal lobe, RFL: right frontal lobe, MCD: malformation of cortical development; FCD: focal cortical dysplasia, MTS: mesial temporal sclerosis.

	EEC - 10 seconds	EEC - 5 seconds	EEC	EEC + 5 seconds	EEC + 10 seconds	EEC, beta frequency	EEC, contiguous removal	EEC, seizure #2
Control centrality	0.80	0.82	0.84	0.79	0.81	0.80	0.82	0.79
Node strength	0.95	0.95	0.96	0.96	0.96	0.95	0.95	0.96
Betweenness centrality	0.88	0.89	0.88	0.89	0.88	0.88	0.88	0.89
Eigenvector centrality	0.95	0.95	0.97	0.96	0.96	0.95	0.94	0.96
Clustering coefficient	0.94	0.94	0.95	0.95	0.95	0.94	0.91	0.95
Synchronizability	0.95	0.93	0.92	0.93	0.94	0.91	0.85	0.94
Global efficiency	0.98	0.97	0.98	0.99	0.99	0.98	0.92	0.98
Transitivity	0.98	0.97	0.98	0.99	0.99	0.98	0.93	0.97

Supplemental Table 4.2. Metric reliability for alternative conditions. The metric reliability, defined in the text, for alternative conditions including different time periods relative to the earliest electrographic change (EEC) (as opposed to the primary analysis on the EEC), beta frequency coherence (as opposed to the primary analysis on high gamma frequency coherence), removal of a contiguous set of electrodes (as opposed to random set), and the second seizure (as opposed to the first seizure).

	EEC - 10 seconds	EEC - 5 seconds	EEC	EEC + 5 seconds	EEC + 10 seconds	EEC, beta frequency	EEC, contiguous removal	EEC, seizure #2
Control centrality	0.80	0.82	0.84	0.79	0.81	0.80	0.82	0.79
Node strength	0.95	0.95	0.96	0.96	0.96	0.95	0.95	0.96
Betweenness centrality	0.88	0.89	0.88	0.89	0.88	0.88	0.88	0.89
Eigenvector centrality	0.95	0.95	0.97	0.96	0.96	0.95	0.94	0.96
Clustering coefficient	0.94	0.94	0.95	0.95	0.95	0.94	0.91	0.95
Synchronizability	0.95	0.93	0.92	0.93	0.94	0.91	0.85	0.94
Global efficiency	0.98	0.97	0.98	0.99	0.99	0.98	0.92	0.98
Transitivity	0.98	0.97	0.98	0.99	0.99	0.98	0.93	0.97

Supplemental Table 4.2. Metric reliability for alternative conditions. The metric reliability, defined in the text, for alternative conditions including different time periods relative to the earliest electrographic change (EEC) (as opposed to the primary analysis on the EEC), beta frequency coherence (as opposed to the primary analysis on high gamma frequency coherence), removal of a contiguous set of electrodes (as opposed to random set), and the second seizure (as opposed to the first seizure).

	EEC, 80%	EEC, 60%	EEC, 40%	EEC, 20%
Control centrality	0.62	0.70	0.76	0.84
Node strength	0.72	0.84	0.92	0.96
Betweenness centrality	0.62	0.70	0.79	0.88
Eigenvector centrality	0.75	0.86	0.92	0.97
Clustering coefficient	0.60	0.78	0.88	0.95
Synchronizability	0.65	0.77	0.85	0.92
Global efficiency	0.82	0.91	0.95	0.98
Transitivity	0.84	0.91	0.95	0.98

Supplemental Table 4.3. Metric reliability for all removal percentages. The metric reliability, defined in the text, for all removal percentages tested. All data is shown for the time period at the EEC.

	EEC - 10 seconds	EEC - 5 seconds	EEC	EEC + 5 seconds	EEC + 10 seconds	EEC, beta frequency	EEC, seizure #2
Control centrality	-0.09	0.02	0.80	2.64	1.15	0.49	1.06
Node strength	1.47	0.07	1.25	1.30	-0.83	1.76	0.95
Betweenness centrality	-0.83	-0.22	-0.95	-1.22	0.13	0.04	-0.31
Eigenvector centrality	1.60	0.22	1.02	1.72	-0.47	1.92	1.37
Clustering coefficient	1.75	0.02	1.23	1.53	-0.84	1.16	1.22
Synchronizability	-0.30	0.80	-0.26	0.85	0.84	-0.11	-1.06
Global efficiency	2.00	0.79	0.74	0.54	2.31	1.67	-0.61
Transitivity	-0.02	1.43	0.37	0.30	1.70	0.55	-1.39

Supplemental Table 4.4. Association between metric agreement and distance of removed electrodes from seizure onset zone for alternative conditions. Values denote the *t*-statistic evaluating the patient-aggregated Fisher's transformed Spearman rank correlations for the distance-agreement associations. The method for calculating the agreement-distance association is described in the text. Positive values indicate that the metric is more sensitive to removing electrodes near the resection zone. Values are shown for different time periods relative to the earliest electrographic change (EEC) (as opposed to the primary analysis on the EEC), beta frequency coherence (as opposed to the primary analysis on high gamma frequency coherence), and the second seizure (as opposed to the first seizure). No association was significant for $\alpha = 0.05/8$ (Bonferroni correction).

	EEC, 80%	EEC, 60%	EEC, 40%	EEC, 20%
Control centrality	-0.58	0.61	2.23	0.80
Node strength	0.42	1.53	1.53	1.25
Betweenness centrality	0.74	0.16	0.49	-0.95
Eigenvector centrality	0.47	1.72	1.46	1.02
Clustering coefficient	0.55	1.46	1.44	1.23
Synchronizability	1.28	2.25	0.95	-0.26
Global efficiency	0.34	0.89	0.90	0.74
Transitivity	0.54	0.54	0.55	0.37

Supplemental Table 4.5. Association between metric agreement and distance of removed electrodes from seizure onset zone for alternative removal percentages. Values denote the *t*-statistic evaluating the patient-aggregated Fisher's transformed Spearman rank correlations for the distance-agreement associations. The method for calculating the agreement-distance association is described in the text. Positive values indicate that the metric is more sensitive to removing electrodes near the resection zone. Values are shown for different percentages of removed electrodes. All results are for the time period at the earliest electrographic change (EEC), high gamma frequency coherence, and the first seizure. No association was significant for $\alpha = 0.05/8$ (Bonferroni correction).

	EEC - 10 seconds	EEC - 5 seconds	EEC	EEC + 5 seconds	EEC + 10 seconds	EEC, beta frequency	EEC, contiguous removal	EEC, seizure #2
Node strength	3	3	3	3	3	4	3	3
Betweenness centrality	4	4	4	5	5	5	4	4
Eigenvector centrality	3	4	3	3	3	4	3	3
Clustering coefficient	3	3	3	3	3	4	3	3
Control centrality	9.5	11	9	8	7	10	6	9
Regional control centrality	52	51	48	51	47.5	52	50	44
Synchronizability	0.09	0.10	0.09	0.09	0.09	0.09	0.13	0.09
Global efficiency	0.01	0.02	0.02	0.02	0.02	0.01	0.02	0.02
Transitivity	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01

Supplemental Table 4.6. Results of jackknife resampling method for alternative conditions. Each nodal value shows the average number of electrodes accounting for 95% of all occurrences of the maximum metric value (minimum for control centrality and regional control centrality) across 1,000 jackknife resamples. Each global value shows the width of the 95% jackknife confidence interval of the network metric across 1,000 jackknife resamples. Values shown are for alternative conditions including different time periods relative to the earliest electrographic change (EEC) (as opposed to the primary analysis on the EEC), beta frequency coherence (as opposed to the primary analysis on high gamma frequency coherence), removal of a contiguous set of electrodes (as opposed to random set), and the second seizure (as opposed to the first seizure).

CHAPTER 5: The effect of iEEG implant geometry on network models

5.1 Abstract

Brain network models derived from graph theory have the potential to guide functional neurosurgery, and to improve rates of post-operative seizure freedom for patients with epilepsy. A barrier to applying these models clinically is that intracranial EEG electrode implantation strategies vary by center, region and country, from cortical grid & strip electrodes (ECoG), to purely stereotactic depth electrodes (SEEG), to a mixture of both. To determine whether models derived from one type of study are broadly applicable to others, we investigate the differences in brain networks mapped by ECoG and SEEG in a cohort of patients who underwent surgery for temporal lobe epilepsy and achieved a favorable outcome. We show that networks derived from ECoG and SEEG define distinct relationships between resected and spared tissue, which may be driven by sampling bias of temporal depth electrodes in patients with predominantly cortical grids. We propose a method of correcting for the effect of internodal distance that is specific to electrode type and explore how additional methods for spatially correcting for sampling bias affect network models. Ultimately, we find that smaller surgical targets tend to have lower connectivity with respect to the surrounding network, challenging notions that abnormal connectivity in the epileptogenic zone is typically high. Our findings suggest that effectively applying computational models to localize epileptic networks requires accounting for the effects of spatial sampling, particularly when analyzing both ECoG and SEEG recordings in the same cohort, and that future network studies of epilepsy surgery should also account for differences in focality between resection and ablation. We propose

that these findings are broadly relevant to intracranial EEG network modeling in epilepsy and an important step in translating them clinically into patient care.

5.2 Introduction

Intracranial electrode recordings from patients with medically-refractory epilepsy characterize the brain's local function, widespread network organization, and guide surgical therapy. From the earliest days of intracranial EEG (iEEG), two major approaches have been used by clinicians for these purposes. In North America, Penfield's use of subdural grid and strip electrodes for electrocorticography (ECoG) persists at many major centers now decades after its initial use[11,156]. Meanwhile, centers in France and Italy still favor Talaraich and Bancaud's approach of using purely depth electrodes in stereoelectroencephalography (SEEG) pioneered at St. Anne's Hospital in Paris[12,156]. Recently, many centers in the United States have begun to favor SEEG due to its superior risk profile and tolerability, though some centers continue to use ECoG for its superior cortical spatial coverage[157]. Unfortunately, many patients do not become seizure free after epilepsy surgery, regardless of implant technique. The reasons for poor outcomes are unclear, but it is likely in part because the interpretation of intracranial recordings is complex, subjective, and plagued by sampling uncertainty[3]. It is also difficult to determine where and how much of the epileptic network must be resected or ablated to fully prevent seizures, particularly in cases where there are no obvious lesions on MRI[2]. Validated, quantitative methods to guide epilepsy surgery could lead to a greater rate of seizure freedom and greater clinical benefit to patients.

Recent evidence supports the hypothesis that epilepsy arises from disordered connectivity[46,158], and that mapping brain networks may aid in both selecting candidates for invasive treatment and identifying therapeutic targets for surgical resection, ablation or device implants[159]. In a brain network model, discrete 'nodes' exist either at the sensor-level for functional connectivity derived from iEEG signals, or at the atlas region-of-interest (ROI) level for structural connectivity derived from imaging[49]. Edges quantify the statistical relationships between nodes in functional connectivity approaches, or streamline count derived from DTI in structural connectivity. A variety of network methods are being explored to localize the epileptogenic zone from intracranial EEG data. Such approaches use interictal[51,52,65] or ictal recordings[4,55,66,77] and are derived from ECoG[52,55,65] or SEEG[105,160] to generate networks. These networks are analyzed using graph theory[52,53,55] or neural mass models, which simulate seizure-like activity and probe the effects of different surgical interventions[65,66]. Intracranial EEG network models are also used to study networks activated during normal brain function, for example in recent studies probing cognition[161] and attention[162]. The majority of these studies use patients implanted with ECoG, often supplemented with depth electrodes placed in the hippocampus. Unfortunately, because of the lack of standardization and difficulties in sharing iEEG data across centers, few studies test their methods in both ECoG and SEEG. In order to validate and translate network methods into clinical practice across centers, it is important to understand how these variations in electrode implantation impact estimate connectivity, subsequent network models and their clinical utility.

While networks derived from functional MRI and diffusion imaging easily generalize across patients and centers due to congruence in their full-brain spatial sampling, iEEG functional networks suffer from sparse sampling and implant heterogeneity. Still, the problem of spatial sampling bias which affects iEEG networks[71] may be offset by: (1) the superior spatiotemporal resolution of iEEG in implanted regions, compared to functional neuroimaging and (2) clinical experience that associates particular patterns in the EEG with typical onset regions, though it can sometimes be difficult to tell if these patterns are the result of seizure generation or spread. To better translate network models into patient care, we must better understand the extent of bias or sensitivity introduced by electrode implantation strategy, in this case ECoG versus SEEG, and its effect on network models. We must then either change implant strategy or develop computational methods to correct for this effect. It is important to note that tradition in specific centers is not the only thing that guides choice of electrode implantation strategy. Other issues, such as the need for stimulation mapping, characteristics of a lesion such as its type and location, ictal semiology and suspected clinical syndrome, as well surgeon and epileptologist experience and training may also factor into approach and electrode choice[163].

There are many differences in implantation strategy between ECoG and SEEG that arise from the electrode hardware itself. Patients implanted with a large ECoG grid will have regular spacing between contacts in the same electrode (e.g. an 8x8 contact grid), supplemented with a few additional strip and depth electrodes in other regions, as needed. The implantation strategy is much more heterogeneous in SEEG. This heterogeneity can manifest itself in the following ways: (i) a wide range in the number of depth electrodes &

electrode contacts from center to center, (ii) different spatial orientations and density of depth electrode implantation, (iii) different assortments of anatomical targets, electrode spacing and (iv) different levels of implant bilaterality[157]. Thus, translating network models into clinical care faces the challenge not only of resolving differences between ECoG and SEEG approaches, but also a high variability in SEEG approaches from center to center. It is unclear whether these distinct properties of electrodes and implant strategy preferentially degrade network representations derived from one approach vs the other, and whether they preclude using the same analysis for networks derived from SEEG and ECoG.

In this study, we explore the effect of implant strategy in a retrospective cohort of patients with drug-resistant temporal lobe epilepsy who were evaluated with either ECoG or SEEG for invasive treatment. We hypothesize the following: (i) ECoG and SEEG networks have distinct properties due to different patterns of spatial sampling, and (ii) differences in network properties between ECoG and SEEG will impact the observed relationship between resected and spared tissue. We aim for our findings to help translate personalized network models of epilepsy into clinical practice, and to inform other applications of intracranial EEG connectivity analysis.

5.3 Materials and Methods

5.3.1 Patient data acquisition

We retrospectively analyzed a data set consisting of 33 patients who underwent intracranial recording during evaluation for epilepsy surgery at the Hospital of the University of Pennsylvania (HUP). Sixteen of these patients had implants with grid, strip, and a small number of depth electrodes, while the remaining patients had only

stereotactically-placed depth electrodes. In this study we refer to cortical-predominant patients as the 'ECoG' group while patients with only depth electrodes constitute the 'SEEG' group. All patients underwent either resection or laser ablation after electrode explant, however in subsequent sections we use the term 'resected' tissue to include ablation patients as well. We chose only patients that achieved good outcome, assessed at 6 months post-operatively, to maximize the likelihood that tissue removed in surgery contained the epileptogenic zone. Table 5.1 lists subject demographics and characteristics of therapy and electrode implants. All subjects provided consent to have their full-length intracranial EEG recordings and anonymized imaging and metadata publicly released on the ieeg.org portal, an open-source online repository for electrophysiologic studies[69,70].

	ECoG	SEEG	p-value
Total number of subjects	16	17	
Number of female subjects	10	8	0.49[a]
MRI*			0.75[a]
Lesional	8	10	
Non-Lesional	8	8	
Type of surgery			**0.0033[a]**
Resection	14	7	
Laser ablation	2	11	
Node counts			
Total GM contacts			
Mean ± std. dev.	92.1 ± 21.2	88.6 ± 34.7	0.72[b]
Depth GM contacts			
Mean ± std. dev.	10.9 ± 9.3	88.6 ± 35.4	**1.5e-6[b]**
Total GM resected / ablated			
Mean ± std. dev.	16.9 ± 14.0	9.1 ± 6.0	0.08[b]
Depth GM resected / ablated			
Mean ± std. dev.	4.1 ± 4.8	9.1 ± 6.0	**0.01[b]**

Table 1: Clinical and demographic information. We analyzed a retrospective cohort of 33 patients with drug-resistant epilepsy who underwent surgery of the temporal lobe and achieved seizure freedom at 6 months post-operatively. Abbreviations: GM: grey matter. [a] = Fisher's exact test; [b] = Wilcoxon rank-sum test.

Each patient underwent a standard epilepsy imaging protocol including pre-implant MRI, post implant CT & MRI, and post-resection MRI. We have previously described our method for localizing electrode locations in detail[55,87], and briefly summarize them in Figure 5.1. Post-iEEG-implant MRI (Figure 5.1D) was registered to pre-implant MRI (Figure 5.1C) using ANTs[164] and electrodes are segmented to derive their coordinates using ITK-SNAP[89]. Any electrode contacts with centroids outside of the brain in the native MRI space were eliminated. We then non-linearly registered the pre-implant MRI into Montreal Neurological Institute (MNI) space for use with neuroimaging atlases, and visually inspected results for accuracy in each subject. We chose a 90 region AAL atlas[165] to assign each electrode contact location a brain region of interest (ROI). Any electrodes in white matter with centroids not overlapping with any atlas region were eliminated. Finally, we used a semi-automated algorithm previously described and validated[55] to perform resection and ablation zone segmentations, which allow the electrode contacts targeted by surgery to be determined.

5.3.2 Connectivity calculation

We calculated functional connectivity using a pipeline that we have previously described and validated[52,55,77]. We randomly selected an interictal segment 1 hour in length for each patient, occurring at least 1 hour away from clinically-annotated seizures. We divided the interictal epoch into one-second intervals (Figure 5.1A) and for each window calculated connectivity using coherence in the beta (15-25 Hz) and low-gamma (30-40 Hz) bands as well as using broadband cross correlation (after applying 5-115 Hz bandpass and 60 Hz notch filters) as they previously have yielded significant results in interictal network

studies in ECoG subjects[52]. We computed the median of each edge over time to obtain a single adjacency matrix for each patient (Figure 5.1B). Together with the results of our imaging pipeline, this process yielded networks in which each node is either resected/ablated or spared in both patients with ECoG (Figure 5.1E) and SEEG (Figure 5.1F).

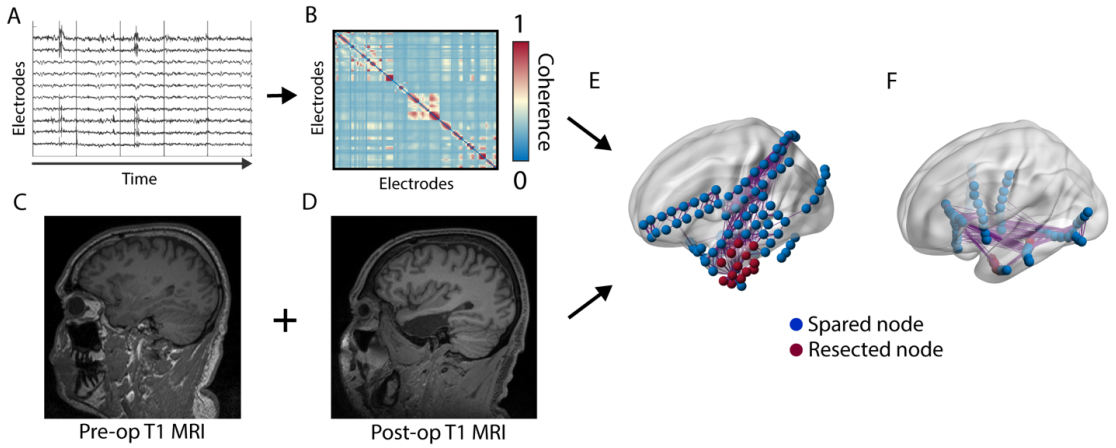


Figure 5.1. Imaging and network methods: (A) We use artifact-free clips of interictal iEEG to calculate (B) mean adjacency matrices using multitaper coherence. (C) Pre-operative and (D) post-operative T1 weighted MRI are used to segment the resection cavity which is used to determine resected nodes. (E) Together, we construct networks with the resected nodes determined in ECoG. (F) SEEG implantations using only depth electrodes appear distinct even for similar anatomic targets.

5.3.3 Network methods

To probe how network structure differs between ECoG and SEEG we detected communities using modularity maximization, which labels nodes so that each community consists of nodes that are more connected to each other and relatively less connected to all other nodes outside of their community[166]. We used a Louvain-like method[167] to maximize modularity, which is represented by $Q = \frac{1}{2m}\sum_{ij}\left[A_{ij} - \frac{k_i k_j}{2m}\right]\delta(c_i c_j)$, where A_{ij} is link between nodes i and j, k are edge weights, m is the sum of all edge weights in the graph,

and δ is the Kronecker delta function. To compare community structure across patients, we computed the participation coefficient[168] which measures the ratio of a node's connectivity strength external versus internal to its module. Averaging participation coefficient across nodes within each patient yielded an estimate of whether networks are (i) integrated, with high connectivity between modules, or (ii) segregated, with lower connectivity between modules.

To illustrate the importance of differences in the way ECoG and SEEG represent epileptic networks in-vivo, we compared the ability of connectivity derived from these modalities to distinguish resected and spared tissue. We chose the simple network metric of node strength, computed as the sum of all edge weights connecting it to all other nodes and is computed as $s_i = \sum_{j=1}^{N} A_{ij}$ in which s_i is the strength of node i, A_{ij} is the adjacency matrix element containing the edge weight between node j and node i, and N is the number of nodes. Our group and others have previously demonstrated that high node strength localizes the epileptogenic zone and predicts surgical outcome in patients implanted with ECoG[52,53], however its translatability to SEEG is not well-established.

5.3.4 Statistical Approach

In our comparisons of network properties between ECoG and SEEG groups we primarily used nonparametric statistical tests such as the Wilxocon rank-sum test, as they do not assume that data are normally distributed. To assess the ability for node strength to detect epileptogenic regions, we used a metric known as the distinguishability statistic (D_{rs}) which quantifies the area under the receiver operating characteristic curve for classifying nodes as either resected or non-resected[169]. The quantity D_{rs} has been previously studied

and validated for its ability to quantify whether networks have sufficient information to determine resected or non-resected regions[146,169]. This value is calculated as the normalized U-statistic, and ranges from 0-1. In our study, a value of 1 implies that all EZ nodes are lower in strength than all non-resected nodes; a value of 0 means that all EZ nodes are stronger than all non-resected nodes; and a value of 0.5 implies that node strength is unable to distinguish between resected and spared nodes.

5.3.5 Data availability

One of our primary goals is to aid in the translation of epilepsy network models into clinical practice. To this end, we shared all raw intracranial EEG and imaging data for HUP patients at iEEG.org, a free cloud sharing platform for electrophysiological data. Each subject's recordings are associated with the ID listed in Supplementary Table 5.1 and can be accessed through the web interface or the open-source iEEG.org MATLAB & python toolboxes. The code for calculating adjacency matrices is available at GitHub.com/Akhambhati/echobase, processed adjacency matrices, and code for comparing networks between ECoG and SEEG is hosted at GitHub.com/jbernabei/ecog_vs_seeg.

5.4 Results

Here we compared networks mapped by ECoG and SEEG in a cohort of temporal lobe epilepsy patients. We aimed to describe how each implant approach is biased towards distinct network properties. We then compared how these distinct network properties affected how well connectivity could distinguish epileptogenic and non-epileptogenic regions with the ultimate goal of improved surgical planning.

5.4.1 Anatomical sampling is similar between modalities

We first asked what differences in the location and extent of anatomic sampling exist between ECoG and SEEG. After quantifying the top 15 anatomical targets implanted by each approach (Supplemental Figure 5.1A), we found a similar distribution of electrode contacts with temporal gyri, hippocampus, and inferior frontal gyri highly sampled. While the total number of implanted nodes was higher in SEEG compared to ECoG (120.6 ± 41.5 vs. 94.5 ± 22.3, rank-sum $p < 0.05$), many depth electrodes localized to white matter and ultimately the number of nodes in grey matter (GM) across ECoG and SEEG was similar (92.1 ± 21.2 vs. 88.6 ± 34.7, rank-sum $p = 0.7$, Table 5.1). Ensuring similar node count was critical for comparing networks and thus for all subsequent analyses we considered only GM nodes. We observed a slight bias of ECoG to favor ipsilateral sampling with more nodes than SEEG implanted in the same hemisphere as the resection zone (Supplemental Figure 5.1B, rank-sum test $p = 0.02$). The median number of contralateral nodes was higher in SEEG than ECoG (25 vs. 4), however this did not reach statistical significance (Supplemental Figure 5.1C, rank-sum $p = 0.06$). Despite the modest differences in hemispheric differences in GM nodes, we observed similar median internodal distances between ECoG and SEEG (Figure S1D, S1E, rank-sum $p = 0.2$). Overall, the targets sampled by ECoG and SEEG for patients with temporal lobe epilepsy were similar, implying that differences in anatomy and internodal distance alone would not primarily drive any subsequent differences in network models.

5.4.2 Differences in mapping resected versus spared tissue

Although ECoG and SEEG sample from similar brain regions, they may not represent the epileptogenic regions similarly from a network perspective. We aimed to gauge the ability of each implant strategy to distinguish resected and spared tissue using the distinguishability statistic (D_{rs}). This value is high when resected nodes are weaker than the spared network and low when they are stronger (Figure 5.2A). Across our cohort, ECoG patients tended to have a low D_{rs} value while SEEG had higher and more variable values (rank-sum test, $p < 0.01$), which was unexpected given that all patients had temporal lobe epilepsy and achieved good surgical outcome (Figure 5.2B). We then sought to determine whether the difference network relationship between resected and spared tissue could result from the frequent placement of depth temporal depth electrodes in ECoG subjects (Figure 5.2C). In these patients we found that resected nodes from surface electrodes had higher normalized strength than non-resected surface electrodes (rank-sum test, $p < 0.01$), however resected depth electrodes were not higher in strength than non-resected depth electrodes (rank-sum test, $p = 0.1$). Furthermore, non-resected depth electrodes were higher in strength than non-resected surface electrodes (rank-sum test, $p < 0.01$), and resected depth electrodes were higher in strength than resected surface electrodes (rank-sum test, $p < 0.01$). These findings, and the sizable proportion of resected depth electrode contacts in ECoG could account for the observed difference in D_{rs} values between the two implantation strategies.

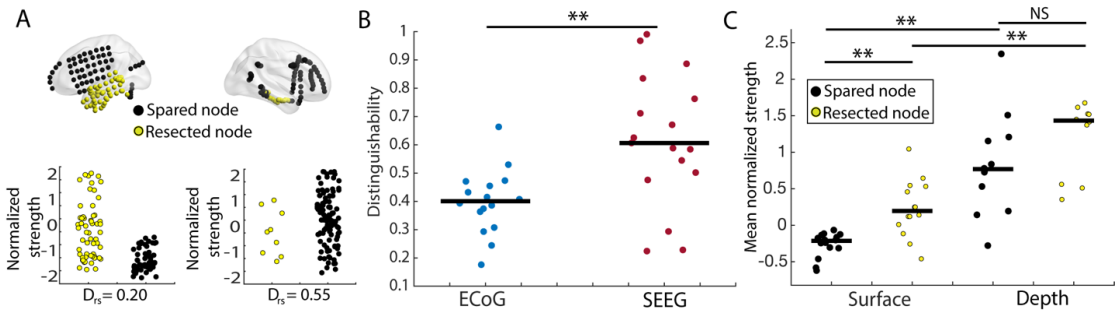


Figure 5.2. Network localization. (A) Distinguishability statistic calculated for an ECoG patient (left) and a SEEG patient (right). In cases where resected node strength is higher than the remaining network on average, D_{rs} will have a low value, in cases where resected node strength is lower than the remaining network D_{rs} will be high. A D_{rs} value of 0.5 means that node strength cannot distinguish resected and spared tissue (B) In networks of grey matter nodes, D_{rs} of resected and spared tissue is higher in SEEG compared to ECoG (rank-sum test, $p = 0.0026$). (C) In patients with ECoG we found resected nodes from surface electrodes to be higher in strength than non-resected surface electrodes (rank-sum test, $p = 0.0065$). Non-resected depth electrodes were higher in strength than non-resected surface electrodes (rank-sum test, $p = 0.0013$). Resected depth electrodes were higher in strength than resected surface electrodes (rank-sum test, $p = 0.0031$). Resected depth electrodes were not higher in strength than non-resected depth electrodes (rank-sum test, $p = 0.14$). ** = $p < 0.01$.

5.4.3 Distinct network properties between ECoG and SEEG

We sought next to determine whether we could adequately correct for our findings of D_{rs} differences between SEEG and ECoG by regressing for internodal distance in an electrode-specific manner. We fit a nonlinear regression model to ECoG and SEEG separately (Figure 5.3A) using a rational polynomial (rat11 in MATLAB) which has been previously validated in interictal network analysis for epilepsy[146]. Within ECoG we also used different models for depth-depth connections, surface-depth connections, and surface-surface connections. Even after correcting for internodal distance, network heterogeneity was higher in SEEG (Figure 5.3B), represented by a higher standard deviation of edge weight residuals across the network (rank-sum test, $p < 0.01$). Furthermore, we calculated

modularity in distance-corrected networks (Figure 5.3C) and found a higher median participation coefficient in SEEG representing higher network integration (rank-sum test, $p < 0.01$). Analogous results for broadband cross – correlation and low – gamma coherence are found in Supplementary Figure 5.3. These results indicate that despite accounting for internodal distance and the effects of distinct electrode types on connectivity, differences in global network properties remain.

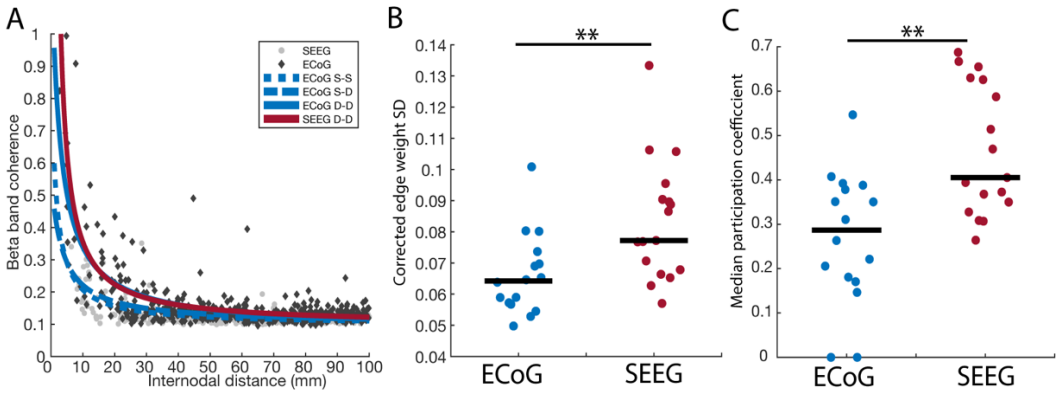


Figure 5.3. Global network structure is impacted by sampling differences between ECoG and SEEG. (A) We fit a nonlinear regression model to ECoG surface – surface (dotted blue line), surface – depth (dashed blue line), and depth – depth connections (solid blue line), as well as SEEG depth – depth connections (solid red line). (B) After correcting for internodal distance, the standard deviation of edge weights remained higher in SEEG versus ECoG (rank-sum test $p = 0.0052$). (C) After correcting for internodal distance, the median participation coefficient remained higher in SEEG versus ECoG (rank-sum test $p = 0.0018$). Abbreviations: S-S: surface – surface, S-D: surface – depth, D-D: depth – depth, SD: standard deviation, ** = $p < 0.01$.

5.4.4 Modifying ECoG and SEEG networks to correct for sampling bias

We finally asked how our internodal distance correction would affect resection zone distinguishability in ECoG and SEEG, and whether we could correct for any remaining differences in network localization by simplifying initial networks to reduce sampling differences. Due to the different balance of ipsilateral and contralateral nodes in

ECoG and SEEG (Supplemental Figure 5.1B & C) and the distinct connectivity of intra- vs inter-hemispheric edges, we hypothesized that eliminating nodes contralateral to the resection zone of distance-corrected networks could improve localization. Based on the distinct modular structure of ECoG and SEEG, we additionally hypothesized that averaging all edges between pairs of brain regions and thus reducing nodes from representing electrode contacts to representing atlas-level regions of interest (ROI) could correct for different balances of connectivity within and between modules. We performed each of these steps for all ECoG and SEEG patients (Figure 5.4A).

We then calculated D_{rs} for each of the three modifications (i) correction for internodal distance (DC), (ii) using only unilateral nodes ipsilateral to the resection zone (UL), and (iii) atlas-level ROI (AR) (Figure 5.4B). For each condition, SEEG networks had a higher D_{rs} value than ECoG: GM ECoG vs. GM SEEG, as in Fig. 2B: rank-sum test $p < 0.01$, DC ECoG vs. DC SEEG: rank-sum test $p < 0.01$, UL ECoG vs. UL SEEG: rank-sum test $p < 0.01$, AR ECoG vs. AR SEEG: rank-sum test $p = 0.02$. Each condition in ECoG and SEEG also had a higher D_{rs} value than in uncorrected networks (DC / UL / AR ECoG vs GM ECoG: sign-rank test $p < 0.01$ for each, DC / UL / AR SEEG vs GM SEEG: sign-rank test $p < 0.001$ / 0.001 / 0.01). Analogous results for broadband cross – correlation and low – gamma coherence are found in Supplementary Figure 5.2. The atlas-ROI representation had the highest median distinguishability in ECoG compared to the base network, however SEEG patients had higher D_{rs} values compared to ECoG even for this condition. As ECoG and SEEG networks with distance-corrected, unilateral, ROI-level nodes should be as similar as possible, we hypothesized that remaining differences were a

result of differences in the extent of surgical intervention between groups (Table 5.1). Indeed, across ECoG and SEEG, we found a strong, negative correlation between D_{rs} values and the number of atlas ROI nodes resected (Pearson correlation rho = -0.48, $p < 0.005$, Figure 5.4C). This finding suggests that the network relationship between resected and spared tissue depends on the focality of the surgical approach.

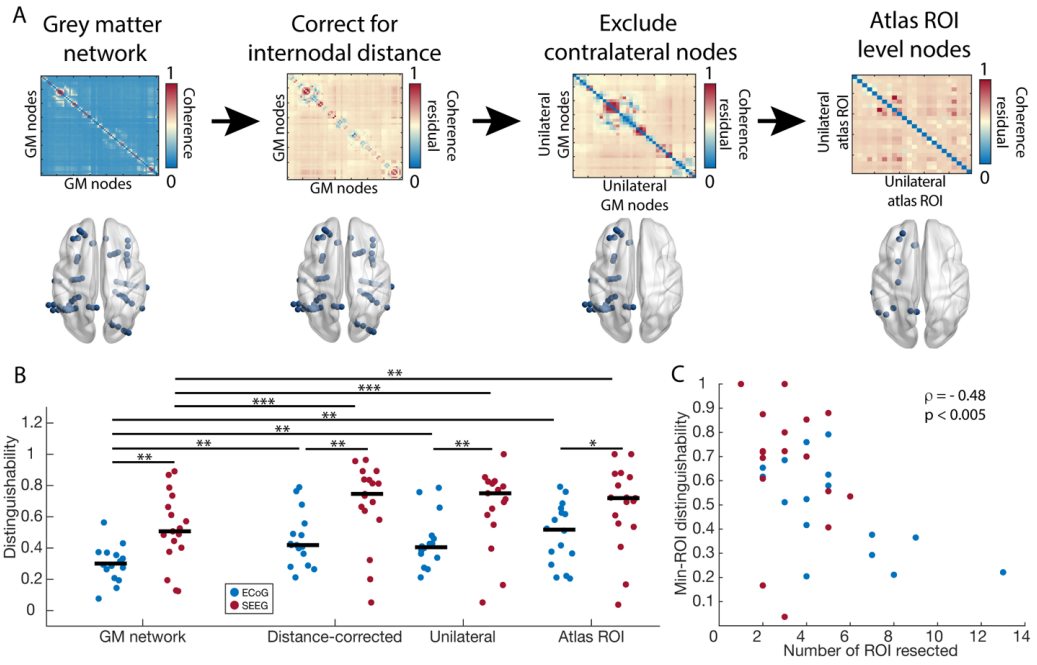


Figure 5.4. ECoG and SEEG have distinct representations of the epileptogenic zone. (A) We used three approaches to modifying networks to probe sampling differences between ECoG and SEEG and their effect on distinguishing resected and spared tissue. We corrected for the effects of internodal distance (DC). Then, we eliminated nodes contralateral to the resection zone (UL). Finally, we averaged edges between pairs of brain regions to have a single node per atlas-level region of interest (AR). (B) We compared the effect of correcting for internodal distance to unilateral to atlas ROI representations. For each condition, SEEG networks had a higher D_{rs} value than ECoG. Each condition in ECoG and SEEG also had a higher D_{rs} value than not accounting for internodal distance. GM ECoG vs GM SEEG, as in Fig. 2B: (rank-sum test $p = 0.0065$), DC ECoG vs DC SEEG: (rank-sum test $p = 0.0059$), UL ECoG vs UL SEEG: (rank-sum test $p = 0.0047$), MR ECoG vs MR SEEG: (rank-sum test $p = 0.022$). DC/UL/AR ECoG vs GM ECoG: (sign-rank test $p = 0.0027/0.0061/0.0011$), DC/UL/AR SEEG vs GM SEEG: (sign-rank test $p = 0.0006/0.0008/0.0031$). (C) D_{rs} values of min-ROI networks were negatively correlated with the number of ROI that contained electrode contacts in the resection zone (Pearson correlation $p < 0.005$).

5.5 Discussion

Understanding the sampling differences of different implant approaches is critically important when applying network models to interpret iEEG data. Here, we showed how unique characteristics of ECoG and SEEG sampling result in distinct properties of derived networks despite similar clinical targets. Node strength, a frequently studied network metric in epilepsy, had an unpredictable relationship between resected and spared tissue, and accounting for internodal distance & electrode type still resulted in distinct network properties. We also establish that these general patterns are present in different frequency bands and in both coherence and correlation measures of functional connectivity. Finally, we provided several methods to partially mitigate the effects of sampling bias introduced by implantation strategy on network models, and showed that remaining differences are associated with the focality of the subsequent resection or ablation.

Our study adds to the growing body of literature on methodological considerations for applying network models clinically[71,146]. From these studies, we recognize that a major challenge in applying network models to the epileptic brain is determining whether observed patterns in brain activity and therefore network structure truly capture the phenomena of interest, and to what extent they arise from sampling artifact. To this end, we must acknowledge that the sampling bias in both ECoG and SEEG is distinct, and that subtle differences in the arrangement of electrodes can determine whether connectivity can accurately uncover true pathology. Thus, sampling bias is more complex than just whether or not a particular target was sampled, and we must take special care to ensure that models

are not biased by the locations of electrodes selected by physicians to simply confirm *a priori* suspicions present before implant.

While the finding that ECoG and SEEG have distinct network connectivity patterns in resected versus spared tissue is significant, it may in part reflect conceptual differences underlying these implant strategies. ECoG attempts to map the boundaries of epileptogenic cortical regions by assessing seizures and interictal activity[12], while SEEG focuses on 'electro-anatomo-clinical' correlations[12], in which broader network mapping and the relationship of anatomical spread to seizure semiology is important. For these reasons, as well as the typical use of a single electrode type & geometry, that SEEG may be superior from a network perspective as the inherent conceptualization of the modality takes the network approach in mind[170]. In particular, reducing networks to atlas ROI nodes as we present here may be an appealing approach for this type of mapping in the future, since these regions and their connections correspond to anatomically relevant and interpretable structures. Such an approach could also facilitate the integration of findings from intracranial EEG networks with studies of quantitative imaging such as fMRI and DTI which typically use atlas ROI nodes, or through the use of intracranial EEG atlases for the prediction of missing information[148,171].

While others have reported differences in connectivity values between depth electrodes and surface electrodes[172], the potential scientific and clinical relevance of network differences that arise as a result of these are significant. It is likely that this finding underlies the results of Figure 5.2B, that in uncorrected networks of ECoG which often include temporal depths, connectivity of resected tissue is relatively strong whereas in

SEEG it is variable. Indeed, in cases of suspected temporal lobe epilepsy mapped by ECoG, the chance that depth electrodes will capture the seizure onset zone is high. Furthermore, the different physical size and cylindrical shape of each electrode contact in SEEG compared to ECoG records local fields from different types of neural populations which could have distinct coherence values. Overall, this observation is likely fixed by performing separate internodal distance corrections for depth and cortical electrodes, which adds to the literature that regressing for internodal distance improves outcome prediction[146].

The results of our sampling correction process (Figure 5.4) reveal interesting aspects of sampling differences between ECoG and SEEG. The finding that eliminating nodes contralateral to the resection zone doesn't significantly change localization from bilateral distance-corrected networks, implies that this issue is not a major factor driving why certain models may succeed or fail in some patients. Indeed, most subjects do not have symmetric implants but rather have a bias towards the hemisphere with the most clinical correlates. Contralateral electrodes are often placed to address clinical hypotheses of lateralization, and due to their relative isolation from the bulk of electrodes, it is possible that they are already outliers in the network and do not contribute highly to the outcome of the distinguishability statistic. On the other hand, if the true epileptogenic zone is in the hemisphere with fewer electrodes, network models may struggle with localizing it. Furthermore, our finding that averaging edges between pairs of brain regions to convert nodes to atlas ROIs maintains similar performance implies that network models may not

need dense sampling within regions, but instead may benefit from sampling inter-regional connections.

Another important consideration to the application of network models is the issue of surgical focality which may differ significantly between resection and ablation patients. While many previous studies focus on resection patients, which may extend to natural anatomic boundaries therefore removing additional, non-epileptogenic tissue, a large part of our cohort underwent ablation in which lesions are relatively small and more specific to the epileptogenic zone. Indeed, others have found that a large number of nodes within and outside of the resection zone is important for accurate outcome prediction from interictal connectivity, which may be impossible for ablation patients. In this context, our finding of a negative correlation between D_{rs} and number of regions targeted may imply that the truest representation of the epileptogenic zone is of low node strength relative to the rest of the brain. This notion is supported by studies demonstrating cellular loss in these regions, particularly the hippocampus, in temporal lobe epilepsy[173,174]. As minimally-invasive approaches such as laser ablation become more common, it is important that our notions of network abnormalities and methods of localizing the epileptogenic zone and outcome prediction do not rely too heavily on findings from resection patients alone.

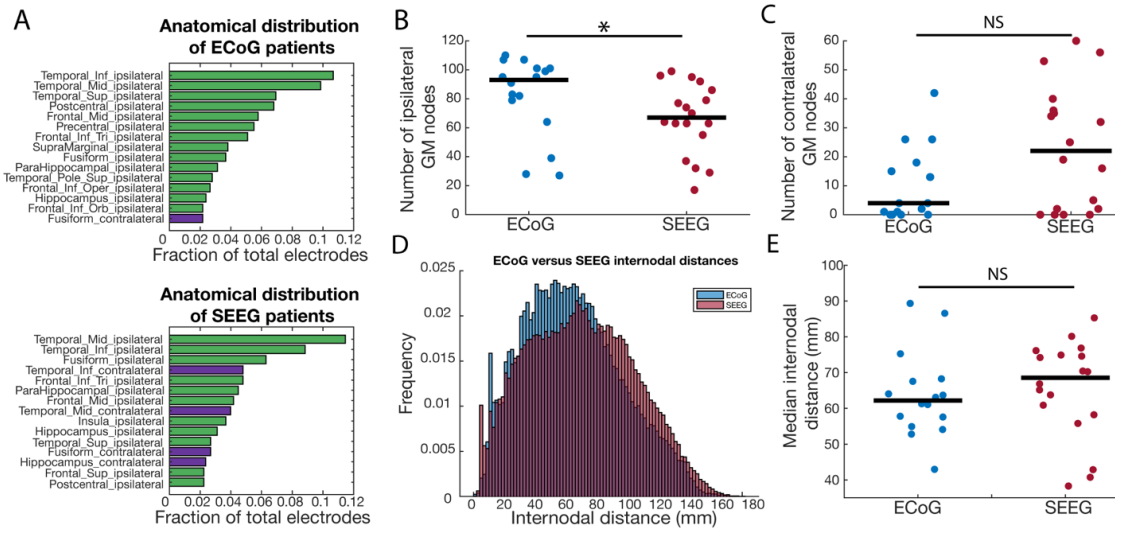
Despite its encouraging results in illuminating the differences between ECoG and SEEG networks, our study has several key limitations. We focused our analysis on node strength, one of the simplest graph theory metrics that has been studied frequently in epilepsy, so this measure may not fully represent the complexity of abnormal networks in this disorder. However, given that node strength is among the least sensitive network

metrics to sampling bias[71], we felt that it was a reasonable choice to compare these approaches. Furthermore, It is well known that node strength is correlated with other network centrality metrics[175] and even certain phenomenological network models which employ dynamical systems linked by functional connectivity[65]. Thus, many of the principles we highlight here may be broadly generalizable to other network studies in epilepsy, and future work should account for sampling bias where possible. Another limitation is our consideration of a single 1-hour iEEG segment, which does not capture the variability in interictal activity and thus connectivity which is known to follow circadian[172] and slower timescales[176]. However, sleep-wake cycles may be interrupted and difficult to estimate in the epilepsy monitoring unit when sleep deprivation and medication withdrawal are common, and hospital admissions are too short to capture predominant multi-day cycles which are close to a month long in many patients. A final limitation is our analysis of only temporal lobe epilepsy patients. We chose this cohort to minimize variability within and across groups and because they represent the largest number of patients at our center. However, distinct patterns of sampling bias may exist in extratemporal epilepsies. Future work should address whether the relationship of connectivity in other anatomical locations of the epileptogenic has a similar pattern to temporal lobe epilepsy, and whether these patterns are also different in ECoG versus SEEG.

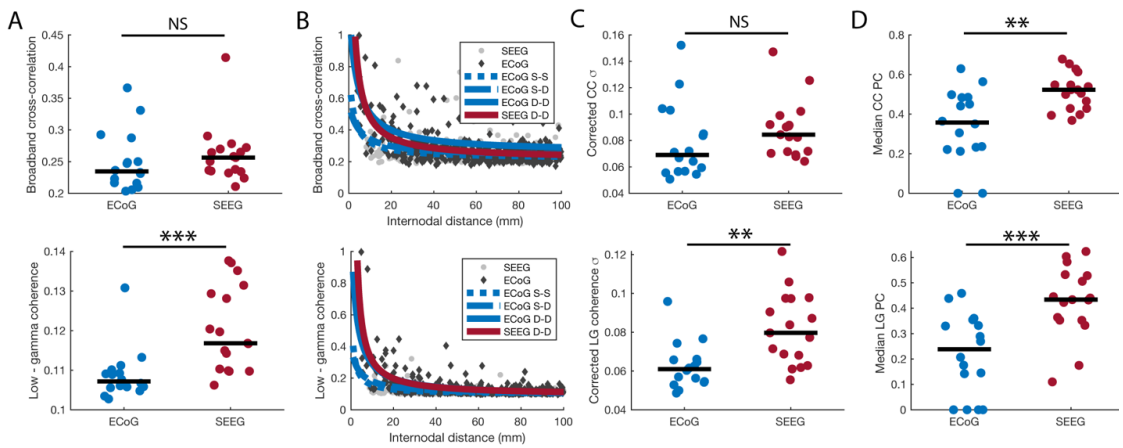
Ultimately, clinical judgement of risk and reward will drive the choice between ECoG and SEEG for individual patients. While the path of clinical translation for network models is complicated by inherent sampling biases across techniques and varying surgical

practice among institutions, we believe it is vital to compare and contrast studies of ECoG and SEEG and recognizing that each provides a distinct view of the brain's underlying connectivity but that neither are 'correct'. Finally, we believe that carefully understanding the sampling properties of networks mapped by intracranial EEG can extend the use of graph theory to broader problems in translational human neuroscience.

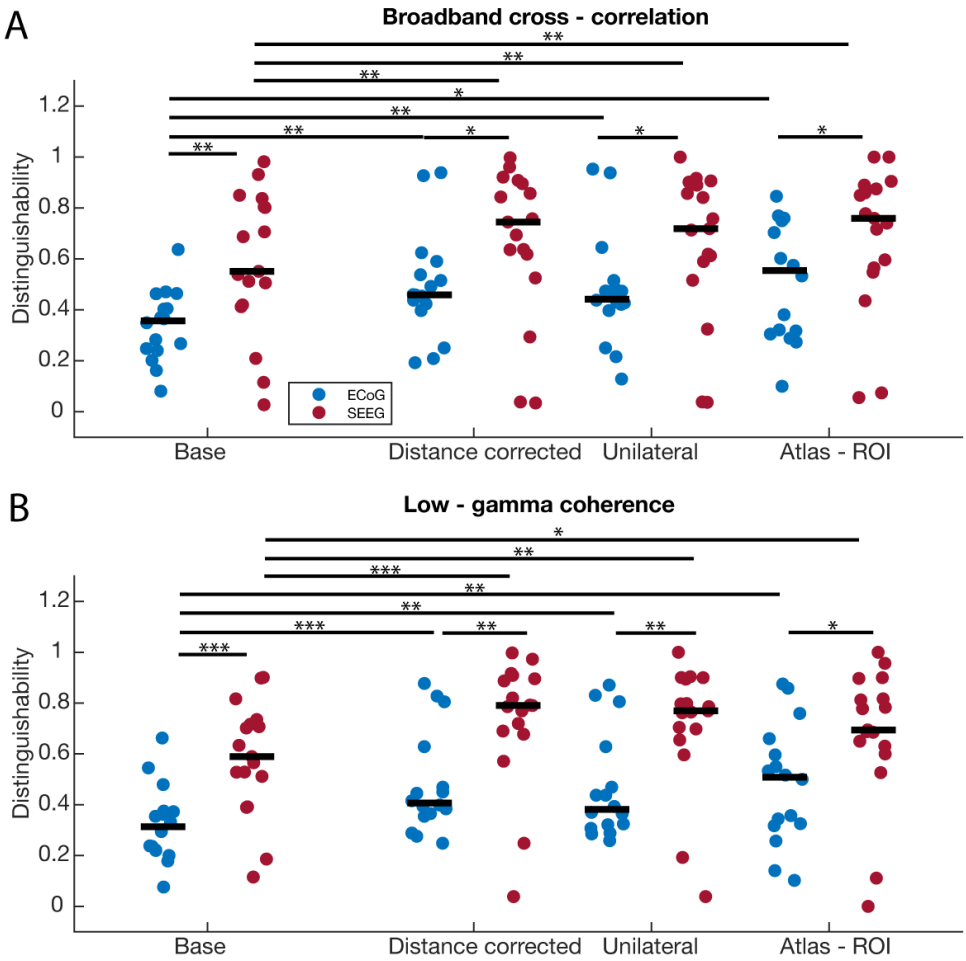
5.6 Supplemental materials



Supplemental Figure 5.1. Anatomic distribution of electrodes is similar in ECoG and SEEG: (A) The top 15 ranked AAL regions in terms of total number of electrodes across patients. Green: ipsilateral, Purple: contralateral. Abbreviations: Mid: middle, Inf: inferior, Sup: superior, Tri: pars triangularis, Oper: pars opercularis, Orb: pars orbitalis. (B) The number of grey matter nodes ipsilateral to the resection zone was higher in ECoG (median: 93) versus SEEG (median: 64), (rank-sum test, $p = 0.022$). (C) Comparing the number of grey matter nodes contralateral to the resection zone between ECoG (median: 4), and SEEG (median: 25) did not reach statistical significance (rank-sum test, $p = 0.057$). (D) Distribution histogram of internodal distances in ECoG (blue) and SEEG (red). (E) The median internodal distance in each patient was not significantly different between ECoG (median: 62.2) versus SEEG (median: 70.2), (rank-sum test, $p = 0.24$).



Supplemental Figure 5.2. Network connectivity in alternate frequency bands: Data for broadband cross-correlation (CC) are on the top row, and low-gamma (LG) coherence is on the bottom row. (A) Median connectivity values were not significantly different between ECoG and SEEG for broadband cross-correlation (rank-sum test p = 0.17), but SEEG had higher low-gamma coherence (rank-sum test p = 1.9e-4). (B) We fit a nonlinear regression model to ECoG surface – surface (dotted blue line), surface – depth (dashed blue line), and depth – depth connections (solid blue line), as well as SEEG depth – depth connections (solid red line). (C) After correcting for internodal distance, the standard deviation of edge weights remained higher in SEEG versus ECoG for low-gamma coherence (LG rank-sum test p = 0.0013) but not broadband cross-correlation (CC rank-sum test p = 0.10). (D) After correcting for internodal distance, the median participation coefficient remained higher in SEEG versus ECoG (CC rank-sum test p = 0.0026, LG rank-sum test p = 8.5e-4). Abbreviations: NS = not significant, S-S: surface – surface, S-D: surface – depth, D-D: depth – depth, SD: standard deviation, ** = $p < 0.01$, *** = $p < 0.001$.



Supplemental Figure 5.3. Distinguishability of resected versus spared tissue in alternate frequency bands: We computed distinguishability for the different networks derived from ECoG and SEEG (Figure 4A), including only grey matter (GM), distance corrected (DC), networks with contralateral nodes eliminated (UL), and networks with atlas-level ROI (AR). (A) For broadband cross-correlation in each condition, SEEG networks had a higher D_{rs} value than ECoG. Each condition in ECoG and SEEG also had a higher D_{rs} value than not accounting for internodal distance. GM ECoG vs GM SEEG: (rank-sum test p = 0.0042), DC ECoG vs DC SEEG: (rank-sum test p = 0.024), UL ECoG vs UL SEEG: (rank-sum test p = 0.029), AR ECoG vs AR SEEG: (rank-sum test p = 0.029), DR/UL/MR ECoG vs GM ECoG: (sign-rank test p = 0.0023/0.0072/0.0131), DC/UL/AR SEEG vs GM SEEG: (sign-rank test p = 0.0012/0.0019/0.0075). (B) For low – gamma coherence: SEEG networks had a higher D_{rs} value than ECoG. Each condition in ECoG and SEEG also had a higher D_{rs} value than not accounting for internodal distance. GM ECoG vs GM SEEG: (rank-sum test p = 9.8e-4), DC ECoG vs DC SEEG: (rank-sum test p = 0.0059), UL ECoG vs UL SEEG: (rank-sum test p = 0.0090), AR ECoG vs AR SEEG: (rank-sum test p = 0.014). DC/UL/AR ECoG vs GM ECoG: (sign-rank test p = 0.0009/0.0027/0.0052), DC/UL/AR SEEG vs GM SEEG: (sign-rank test p = 0.0007/0.0010/0.0245).

CHAPTER 6: An atlas of normative iEEG activity and connectivity

6.1 Abstract

Recording seizures using intracranial EEG (iEEG) is an essential tool for surgical planning for patients with refractory epilepsy. Quantitative measures of interictal iEEG are potentially appealing biomarkers, however their utility is limited by the sparsity electrode implantation as well as the confounds of normal neural activity and connectivity which vary spatiotemporally. We propose that leveraging a large number of patients to construct a normative atlas of intracranial EEG activity and connectivity will allow us to reliably map abnormal regions, thereby serving as a tool to increase our understanding of epilepsy and identify better targets for epilepsy surgery. We aggregated interictal iEEG retrospectively across 166 subjects comprising >5000 channels. For each channel, we calculated normalized spectral power, wavelet entropy, Pearson correlation, and coherence. We constructed an iEEG atlas by mapping the distribution of each feature across the brain, and test the atlas by generating a Z-score for each channel of novel patients. This procedure can reliably identify quantitative abnormalities in clinically relevant areas such as the seizure onset zone (SOZ) and irritative zone. We show that for SOZ within the amygdala and hippocampus, measures of connectivity abnormality are more enhanced than univariate measures of abnormal neural activity. We also find that patients with lesional MRI have a greater level of abnormality in the SOZ compared to non-lesional patients. We input the Z-scores of each metric to a random forest classifier to determine whether channels are likely to be in the SOZ. By integrating measures of both single-channel activity and inter-regional functional connectivity, we find a better accuracy in predicting the SOZ versus normal brain (area under the curve = 0.81) compared to either

group of features alone. The findings of this study serve to directly increase our understanding of the relationship between abnormal neural activity, connectivity and epilepsy, and our methods establish a framework for leveraging big data in surgical planning. Finally, we publicly share our atlas so that others may build upon the data and methods which we present here.

6.2 Introduction

Over 20 million individuals worldwide have drug-resistant epilepsy[10], and for these patients surgical resection provides the best chance of seizure freedom[177]. However, up to 40% of patients relapse after surgery, indicating an insufficient understanding of how seizures arise from the epileptic brain as well as the difficulty of interpreting multimodal data including intracranial EEG (iEEG) which is traditionally used in surgical planning. Another limitation of iEEG the cost and morbidity associated with lengthy hospital stays in specialized epilepsy monitoring units to wait for seizures to occur. Furthermore, as such seizures are typically rare events, the seizure onset zones observed during iEEG may not constitute all brain regions capable of generating seizures, known as the epileptogenic zone[12], a mismatch which could lead to the sub-optimal selection of surgical targets and poor surgical outcomes. A reliable, interictal biomarker of the epileptogenic zone could reduce the need for extended recording and the reliance on precipitating seizures to guide epilepsy surgery.

Epilepsy results in specific patterns of interictal activity that may occur in epileptogenic regions, such as spikes[178] and high frequency oscillations[179]. While these events may provide some localizing value in surgical evaluation[180,181], they have not proven

to be specific enough to serve as reliable biomarkers alone[182]. Spikes can be generated across many regions which are unnecessary to resect for a good surgical outcome[183]. Similarly, HFOs are not always pathologic and may localize outside of epileptogenic tissue[184], making them unlikely to replace seizure onset zone identification as the gold standard of surgical evaluation. Beyond these paroxysmal events, epilepsy may also result in focal differences in baseline rhythmic activity such as slowing[185]. More work is needed to establish sensitive and specific interictal biomarkers of the epileptogenic zone that generalize across a wide range of etiologies and locations of focal epilepsy.

Besides manifesting abnormal neural activity, the seizure onset zone also comprises part of an abnormal network that is increasingly recognized as an essential aspect of epilepsy pathophysiology[47]. This pattern exists across spatial scales at both the structural[186] and functional[51,52] level, and occur not only during seizures[54,55] but also at rest[65]. Methods of mapping connectivity for surgical planning based on neuroimaging suffer from poor temporal resolution, and those based on intracranial EEG suffer from sampling bias introduced by electrode number and placement[71,187]. Beyond sampling bias, quantitative methods of mapping epileptic networks are confounded by normal spatial patterns of neural activity[171] and neural connectivity[148] which vary across the brain. Neuroimaging studies avoid this problem by leveraging identical, full-brain spatial sampling across patients and comparing subjects with epilepsy to healthy controls. Conversely, iEEG is very rarely used in patients without epilepsy and always has incomplete coverage, so traditional normative mapping is impossible. However, recent work has provided methods for aggregating multi-patient iEEG data to attempt normative analyses of full-brain univariate neural

activity[171,188], and bivariate neural connectivity[148], which may serve as a powerful methods of leveraging big data to map epileptogenic abnormalities using iEEG in patients with drug-resistant epilepsy.

Here, we propose and validate a framework for mapping epileptogenic abnormalities across the brain by leveraging two large cohorts of iEEG. Using a previously published, multi-center dataset[171] and one that is novel to our study, we construct a normative atlas of iEEG activity and connectivity. We hypothesize that deviations in these metrics from the normative estimates are reliable markers of the seizure onset zone, and that capturing both focal and inter-regional abnormalities provides the most localizing value. While we expect that the atlas approach will require more data before it can serve as a reliable clinical tool, we aim to highlight its utility as an alternative to traditional iEEG analysis, and also release our data for others to build upon.

6.3 Methods

We retrospectively analyzed data from 166 patients with drug resistant epilepsy. Sixty of these subjects underwent iEEG implantation as part of epilepsy surgery at our center, the Hospital of the University of Pennsylvania (HUP). As no single center has a high enough volume to support the construction of an iEEG atlas, we leveraged a high-quality, publicly available dataset of 106 subjects to increase the amount of normative data. These patients received iEEG at the Montreal Neurologic Institute (MNI) and two other centers, and 1772 channels of their data were judged as clinically normal and released as the MNI Open iEEG Atlas (https://mni-open-ieegatlas.research.mcgill.ca)[171]. Across the multicenter cohort, each patient underwent implantation with subdural grid & strip

electrodes (ECoG), only depth electrodes (SEEG), or a mixture of both. At HUP, all subjects underwent either resection or laser ablation after electrode explant.

	Latest surgical outcome		
	Engel 1	Engel 2+	*p*-value
Number of patients	38	22	
Sex			0.46[a]
Female	17	12	
Male	21	10	
Laterality			0.50[a]
Right	19	9	
Left	19	13	
Pre-surgical MRI			0.05[a]
Lesional	22	7	
Non-Lesional	16	15	
Age onset	14.9 ± 11.7	17.5 ± 14.5	0.63[b]
Age surgery	38.8 ± 10.8	35.2 ± 11.8	0.69[b]
Target			0.41[a]
Temporal	28	14	
Frontal	7	6	
Insular	1	2	
Implant type			**0.008[a]**
ECoG	20	4	
SEEG	18	18	
Surgery type			**0.01[a]**
Resection	25	7	
Ablation	13	15	
Node count			
Total in grey matter	82.1 ± 25.1	91.8 ± 30.5	0.16[b]
Removed	12.3 ± 10.7	8.5 ± 6.2	0.12[b]
Initially Engel 1 before relapse	N/A	8	

Table 6.1. Atlas patient dataset. Patients were grouped by initial surgical outcome. Left column shows patients that achieved Engel 1 outcome at 6 months Right column shows that had poor surgical outcome at 6 months. Statistical tests: *a* rank-sum test; *b* Pearson chi-square test.

We also determined surgical outcome on the Engel scale at a minimum of 6 months after

surgery, and at the 12 & 24 month post-operative interval for the majority of patients. All subjects consented to data collection & sharing, and we performed all research in concordance with protocols approved by the institutional review board of the University of Pennsylvania. The overall characteristics of the HUP cohort are described in Table 6.1..

6.3.1 Data selection

To ensure that data from our center could be compared and combined with MNI data, we implemented a rigorous data selection process that closely followed the gold-standard methods of Frauscher et al., 2018[171]. We selected 60 seconds of data (Figure 6.1A) in two 30-second clips for each subject that met the following criteria in order of priority: (i) Free of artifact, (ii) at least two hours before the beginning of a seizure, and at least two hours after a subclinical seizure, six hours after a focal seizure, or twelve hours after a generalized seizure, (iii) free of spikes, and (iv) not within the first 72 hours of recording to minimize immediate implant and anesthesia effects. All selected clips met conditions (i) and (ii), and the vast majority of clips met conditions (iii) and (iv). We used clips that captured awake brain activity, determined both by the selection of daytime epochs and the use of a custom non-REM sleep detector developed using MNI data and validated on HUP data. Specific details of the sleep detector approach and performance are included in Supplementary Methods 1. As our study centers around identifying clinically normal and abnormal regions, we also used a validated algorithm to detect and quantify spikes[180]. We consider a region that generates at least 1 spike per hour to be part of the irritative zone. A sample of 50 spike detections was manually examined by a fellowship-trained neurologist for positive predictive value. We also recorded clinically-determined seizure onset zones

that were used in surgical evaluation. Similar to the MNI dataset, we define an abnormal channel as one that is within the seizure onset zone, irritative zone, or was within the resection zone. Between HUP and MNI atlases, we have 2630 channels which are clinically normal, and 2577 HUP channels which are clinically abnormal (Figure 6.1B).

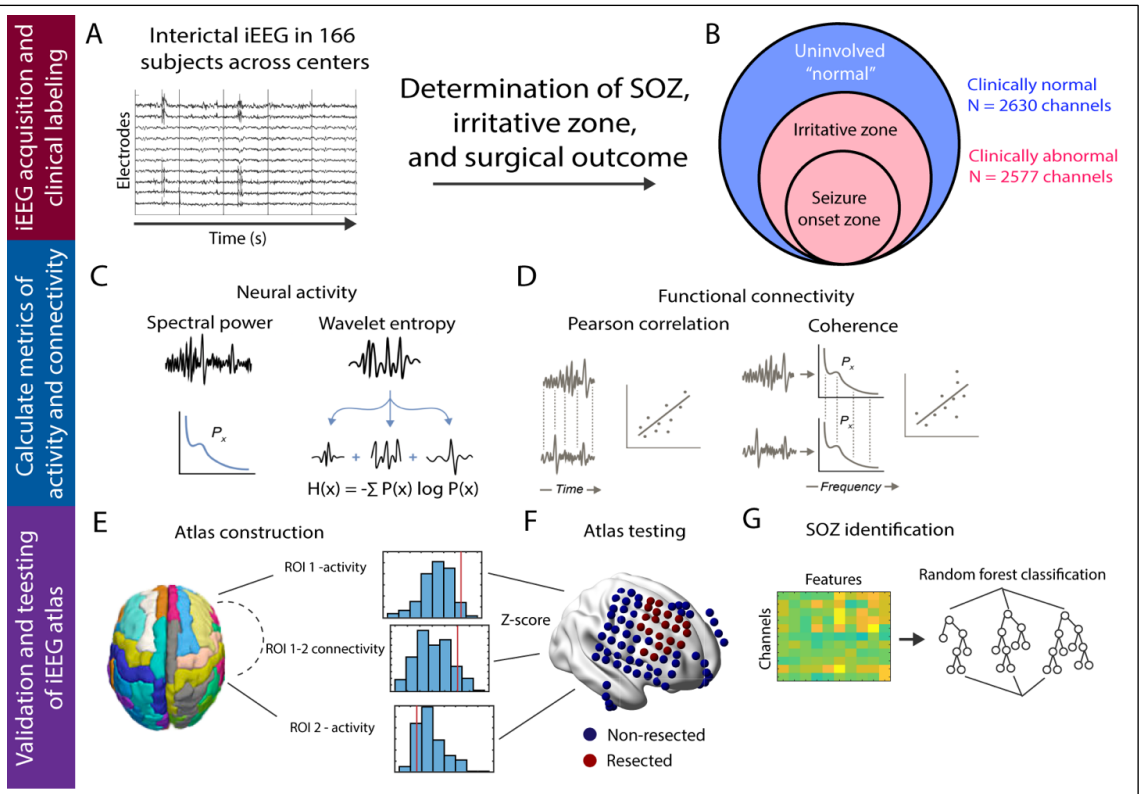


Figure 6.1. Project overview. (A) Interictal iEEG in 166 subjects (106 MNI, 60 HUP). All data was determined to come from awake subjects, SOZ, irritative zone, and surgical outcome. (B) All channels from Engel 1 patients that were outside of the irritative zone and seizure onset zone were considered clinically normal, amounting to 2305 channels. All channels from Engel 2+ subjects, and all channels from irritative zone and seizure onset zone were considered clinically abnormal, comprising 1851 channels. (C) For each channel, we calculated univariate features of normalized spectra power from 0.5-80 Hz, as well as wavelet entropy (D) We also calculate connectivity using band-limited Pearson correlation and coherence (E)We construct an intracranial EEG functional atlas by mapping the distribution of each univariate feature within an anatomical ROI, and each bivariate feature between pairs of ROI (F)We test the atlas by generating a Z-score for each feature in each node (or pair of nodes for bivariate features) of a test patient. (G) We use the Z-scores of each metric as input for a random forest classifier to determine whether individual channels are likely to be in the SOZ.

6.3.2 Patient imaging

Each HUP patient underwent a standard epilepsy imaging protocol including pre-implant MRI, post implant CT, and post-resection MRI. Post-iEEG-implant MRI is registered to pre-implant MRI using ANTs[164] and electrodes are segmented and to derive coordinates using VoxTools[87] and ITK-SNAP[89]. We then register the pre-implant MRI into Montreal Neurological Institute (MNI) space for use with neuroimaging atlases. We finally use a semi-automated algorithm previously described and validated to perform resection and ablation zone segmentations which determine the electrode contacts targeted by surgery. The MNI Open iEEG Atlas data contained MNI coordinates of each electrode as well and all electrode contacts were localized using the AAL atlas[165]. However, we eliminated regions that are not typically targeted by iEEG, including the cerebellum and basal ganglia, and further aggregated neighboring gyri in regions with low sampling (such as occipital lobe) to increase the samples in each region, thus reducing the sparsity of our atlas. This process and the resulting 'reduced' atlas which we will use for the rest of this study is shown in Supplementary Figure 6.1, and contains 20 regions in each hemisphere.

6.3.3 Signal processing

We first montaged the HUP data into a bipolar configuration to match that of the MNI Open iEEG Atlas. We used a first order Butterworth filter with the passband between 0.5 and 80 Hz to remove high-frequency oscillations as well as a 60 Hz IIR notch filter to remove noise, and down-sampled the data to 200 Hz from its original 512 or 1024 Hz sample rate to match the MNI atlas. We excluded any channels contaminated by excessive

noise as well as bipolar pairs in which both contacts were located in white matter, or either contact was located outside of the brain.

For each channel, we calculated the power spectral density (Figure 6.1C) using Welch's method[189] with a 2-second Hamming window and a 1 second overlap. We normalized the spectral density to have a sum of 1 as surface and depth electrodes may have different signal amplitude, but have preserved frequency content[171]. We calculated normalized spectral content in the following frequency bands: delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-12 Hz), beta (12-30 Hz), and gamma (30-80 Hz). We also applied separate bandpass filters (first order Butterworth) to extract the oscillatory components of the signal corresponding to each of the aforementioned canonical frequency bands and calculated median Shannon wavelet entropy[190] across 1-second windows of each channel's 60 seconds of data.

6.3.4 Connectivity calculation

We calculated functional networks within each patient (Figure 6.1D) using Pearson correlation and coherence between each pair of N channels to yield N x N adjacency matrices. For Pearson correlation, we used first-order Butterworth filters to limit connectivity to each of the five frequency bands as discussed above, with resulting connectivity values ranging from -1 to 1. We calculated band-limited coherence, which could range from 0 to 1. For each method, we divided the interictal epoch into one-second intervals and computed the median of each edge over time to obtain a single adjacency matrix for each of the ten feature-connectivity pairs (five frequency bands each of correlation and coherence) in each patient.

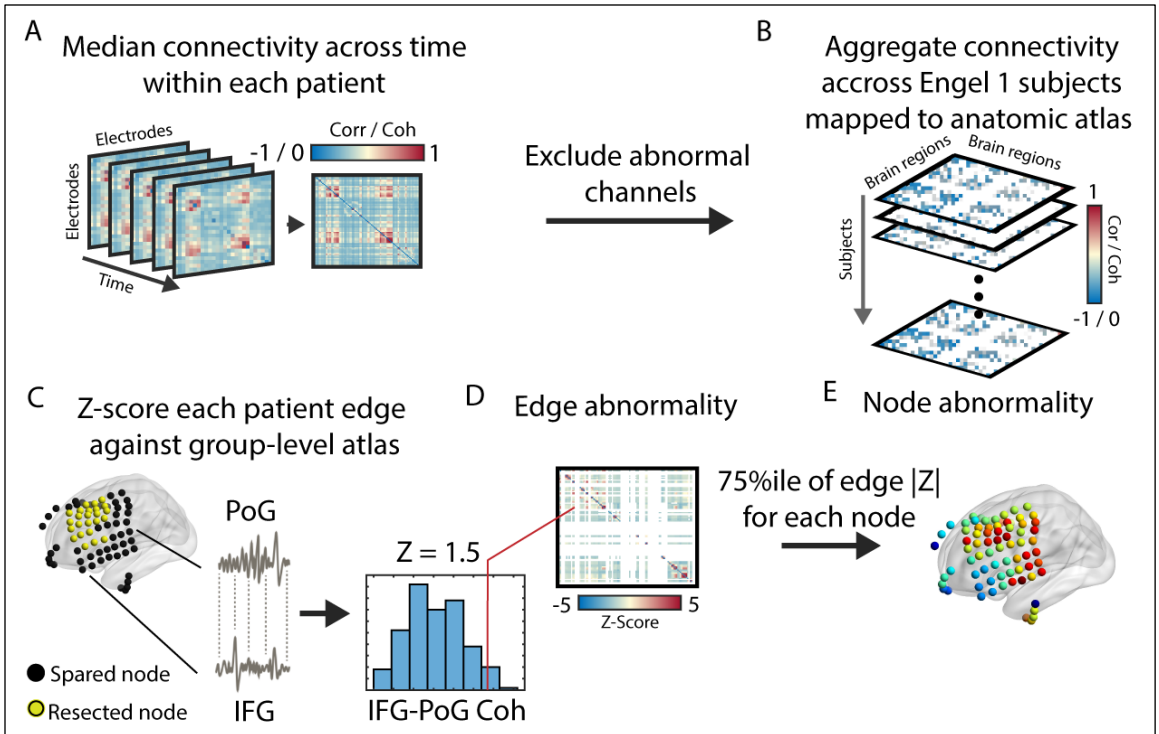


Figure 6.2. Construction of a multi-subject functional connectivity template. (A) The median connectivity (band-limited coherence and correlation) is calculated within each patient across 1-second windows (B) After excluding clinically abnormal channels, all edges between each pair of ROI are aggregated across subjects to construct a functional connectivity atlas. (C) For each patient, each edge is Z-scored against the distribution of edges corresponding to the appropriate brain regions in the atlas. Any edges in the 'test' patient that are not captured in the atlas are not assigned a Z-score and are excluded from further analysis. (D) This process yields an adjacency matrix of edge abnormality Z-scores. (E) Taking the 75th percentile of all absolute edge Z-scores for each node gives a single measure of nodal connectivity abnormality for each connectivity metric.

6.3.5 Atlas construction & testing

Our approach to normative iEEG mapping requires the separation of clinically normal channels and abnormal channels. We assigned each channel to a region of interest (ROI) in the reduced atlas we described above. Each region therefore has a distribution of values for each univariate feature comprised of all normal channels localized to that ROI (Figure 6.1E). We describe our connectivity-based methods figure 2. From the median

adjacency matrix of each patient (Figure 2A), representing the connections between each electrode contact, we use assign each node a location in the common MNI 152 space, and subsequently a ROI of the reduced atlas (Figure 6.2B). For any region pairs which only have fewer than five edges measured across patients we eliminate them from further analysis as their variance estimate is unlikely to be accurate. We further exclude edges which are partially or fully involve nodes which were judged to be clinically abnormal. Our final step in atlas construction is to take the median and standard deviation of each edge across all good outcome patients to yield final matrices corresponding to the expected iEEG connectivity and its variability across patients which serves as our atlas.

6.3.6 Atlas testing & statistical methods

To measure central tendency, we typically use the median as it is more sensitive to outliers. We also log-transformed entropy and coherence features so they would approximate normal distributions, while spectral density and correlation features did not require this step. To map the abnormality across intracranial EEG functional connections, we use the atlas as a look-up table to determine whether each individual connection in a 'test' patient is likely to be normal or abnormal. For each inter-regional edge, we determine its Z-score against the atlas median and standard deviation of connectivity between that same pair of brain regions (Figure 6.2C). Any connections which are unmeasured in the atlas are not assigned Z-scores and are left blank. As we are primarily concerned about the level of abnormality, we took the absolute value of all Z scores to yield a magnitude. As connectivity measures operate at the edge level (Figure 6.2D) rather than the node level which would complicate analysis, we reduced connectivity abnormality to a single value

for each feature-node pair by taking the 75th percentile of absolute Z-score across all edges for each node (Figure 6.2E). This allows the abnormality scores of connectivity and activity to be interpreted similarly. From all absolute Z-scores of univariate and bivariate features (Figure 6.1F), we use a random forest model to identify the seizure onset zone (Figure 6.1G). We use 10-fold cross-validation over normal versus seizure onset zone channels, in which the 20 input features for each channel correspond to 5 bands for each of normalized spectral density, log entropy, correlation, and coherence. We quantify performance using the receiver operating characteristic curve, and compare the performance between the full feature set and using only univariate or only bivariate features.

6.3.7 Code & Data sharing

In the interest of helping network techniques such as ours reach clinical practice, we share all code and data from this study. All pre-implant, post-implant, and post-resection or ablation imaging as well as full intracranial EEG records are available at ieeg.org under the project iEEG_atlas. All code as well as .mat files containing the processed atlas adjacency matrices, and information about clinical metadata is available at GitHub.com/iEEG_atlas. The HUP Open iEEG Atlas file is available at Pennsieve, with a similar structure to the MNI Open iEEG Atlas. A description of each of the data fields contained within the file is available in Supplemental Table 6.1.

6.4 Results

6.4.1 Validating the HUP Open iEEG Atlas

We first sought to determine whether features of the HUP atlas were similar to the MNI atlas to test whether their normative data and ours could be seamlessly combined. We

compared the median normalized spectral density in each frequency band and each brain region (aggregating nodes from the same region in each hemisphere) between datasets using the rank-sum test and adjusted the significance level to $\alpha = 0.0005$ to correct for 100 comparisons (20 regions x 5 frequencies). In general, few region-frequency pairs in MNI and HUP significantly differed (Supplementary Figure 6.3). Therefore, we combined the MNI atlas (1772 channels) with regions in the HUP cohort which we judged to be likely clinically normal (858 channels), yielding a composite normative atlas total of 2630 channels for all subsequent analyses.

We then compared spectral features of the composite atlas against those of channels in the irritative and seizure onset zones, and found that this process identifies a large number of region-frequency pairs as significantly different (Supplementary Figure 6.4). After correcting for 100 comparisons we identified 23 and 19 significantly different region-frequency pairs for irritative and seizure onset zones respectively, validating this approach to normative atlas mapping. However, this analysis could not identify potentially epileptogenic channels in some of the most clinically relevant ROI. While abnormalities in temporal and frontal neocortex were quantitatively evident, median spectral density in all bands was similar in the combined amygdala-hippocampus region between channels in the seizure onset zone and those judged as normal. Thus, additional features are required as mesial temporal lobe epilepsy are some of the most important focal epilepsy syndromes and often may significantly benefit from epilepsy surgery.

To complement spectral density in detecting abnormalities using normative atlas mapping, we asked whether signal entropy significantly differed between clinically normal

and abnormal channels. As values of wavelet entropy in our dataset are negative and skewed, we plotted and analyzed log negative Shannon entropy (Supplementary Figure 6.5) so we may estimate normal distributions for subsequent analyses. Shannon wavelet entropy is a commonly-used metric in quantitative EEG analysis that quantifies the level of disorder in time-series signals. We expected that clinically abnormal channels would generally have less disorder given that they may lose their normal function. Similar to spectral density, we tested whether the distribution of entropy was significantly different in each region between normal, irritative, and SOZ channels. We found that wavelet entropy could distinguish irritative and seizure onset zones from normal particularly in frontal ROI, and in particular could distinguish SOZ from both normal and irritative zones in superior and middle frontal gyri. In these regions, log negative entropy was higher in SOZ versus normal channels implying the seizure onset zone has signals that are more ordered and that normal variability is interrupted. These findings suggest that entropy is an effective metric for normative atlas mapping and may provide complementary information to traditional spectral features.

6.4.2 Mapping abnormality with univariate versus bivariate features

To build upon mapping neural activity with univariate features, we then sought to determine whether inter-regional activity abnormalities identify seizure onset zones using the normative atlas. We first confirmed that HUP and MNI components of the normative atlas provided similar quantifications of brain connectivity by examining the Pearson correlation of their adjacency matrices ($r = 0.43$, $p = 4.6e\text{-}19$). We used the atlas to compute within-patient connectivity and use edges between uninvolved channels to estimate

normative distributions for each inter-regional connection. This process allowed us to generate Z-scores for each connection and we took the maximum Z score across all 10 bivariate features and all 10 univariate features for each node. As clinical abnormalities in the amygdala and hippocampus were not easily distinguished from normal by univariate Z-scores, we first asked whether the seizure onset zone had abnormal connectivity when located in the mesial temporal lobe (Figure 6.3A). We analyzed differences in Z-scores for non-involved, irritative zone, and seizure onset zone separately for the maximum bivariate and maximum univariate measure. We found that both the seizure onset zone and exclusive irritative zone had significantly higher median connectivity abnormality compared to normal channels when located amygdala or hippocampus (rank-sum $p < 0.001$). However, as expected, we observed no significant difference in maximum univariate measures. This finding confirms the importance of mapping not only abnormalities in activity but also those in connectivity, and together with the previous univariate results suggests that normative atlas mapping is able to detect abnormalities across the brain regardless of ROI.

We then asked whether patients with lesional imaging have higher abnormality than those without lesions (Figure 6.3B). In patients with lesions on imaging, we found higher bivariate abnormality within versus outside the SOZ ($p < 0.001$). The SOZ of lesional subjects was higher than the SOZ of non-lesional subjects ($p < 0.001$), outside the SOZ of lesional subjects was also more abnormal than outside the SOZ of non-lesional subjects ($p < 0.001$). For univariate features, there was higher abnormality outside the SOZ compared to within it for lesional subjects ($p<0.001$), whereas non-lesional subjects had higher abnormality within ($p<0.001$) and outside the SOZ ($p<0.05$) compared to lesional subjects.

These results demonstrate that lesions may affect widespread patterns of activity and connectivity in different ways.

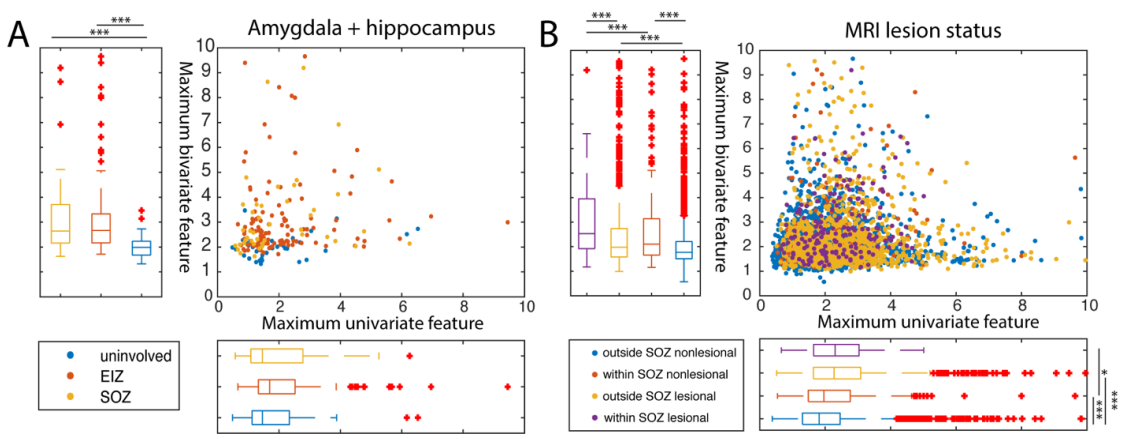


Figure 6.3. Maximum univariate and bivariate abnormality. (A) In the MTL, a region in which univariate features performed poorly, connectivity features are higher in abnormal regions (both EIZ and SOZ) compared to normal brain (rank-sum $p < 0.001$) however they cannot separate channels in the amygdala + hippocampus which spike from those which generate seizures (rank-sum $p = 0.4$). Univariate features in this region could not distinguish SOZ, EIZ, and uninvolved channels (B) In patients with lesions on imaging, we found higher bivariate abnormality within versus outside the SOZ ($p < 0.001$). The SOZ of lesional subjects was higher than the SOZ of non-lesional subjects ($p < 0.001$), outside the SOZ of lesional subjects was also more abnormal than outside the SOZ of non-lesional subjects ($p < 0.001$). For univariate features, there was higher abnormality outside the SOZ compared to within it for lesional subjects ($p<0.001$), whereas non-lesional subjects had higher abnormality within ($p<0.001$) and outside the SOZ ($p<0.05$) compared to lesional subjects.

6.4.3 Clinical applications of normative atlas mapping

We sought to establish the potential clinical utility of iEEG atlas mapping. To illustrate the approach, we highlight two clinical examples from patients with temporal lobe epilepsy. Figure 6.4A shows a patient that achieved seizure freedom from surgery. One electrode contact, in superior frontal gyrus, showed low levels of abnormality across all features and clinically determined to be normal as well. Another in inferior frontal gyrus

showed higher abnormality and localized to the irritative zone. A final electrode in hippocampus showed the highest abnormality in alpha coherence, and was successfully resected as part of the seizure onset zone. Figure 4B shows a patient which experienced a reduction in seizures but not seizure freedom. Here, an irritative zone channel in middle temporal gyrus had higher abnormality scores than a non-involved channel in middle frontal gyrus, or a resected seizure onset zone channel in inferior temporal gyrus. Overall, our approach represents a relatively intuitive method for mapping interictal abnormalities that may not be evident from examining intracranial recordings directly.

We finally sought to quantify the accuracy with which atlas abnormality Z-scores could distinguish seizure onset zone versus normal brain. Using Z-scores from all 20 features we trained a random forest classifier, chosen for its ease of training and robustness to correlated features, and quantified performance using 10-fold cross validation. This process achieved an AUC of 0.81, versus AUC of 0.79 and 0.73 for only univariate and only bivariate features respectively (Figure 6.4C). These results represent good classification performance showing that normative mapping scores can interictally distinguish SOZ from normal channels.

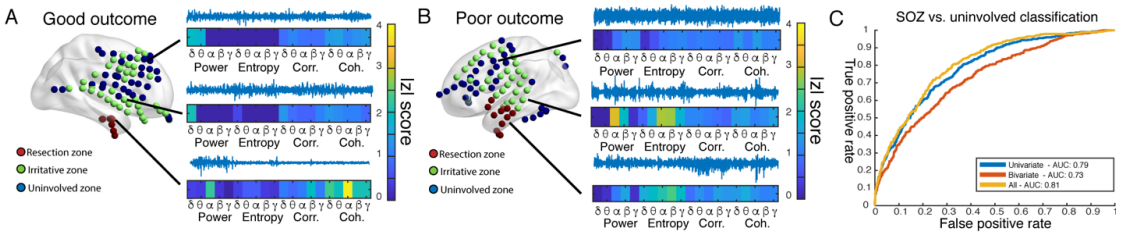


Figure 6.4. Clinical application of abnormality mapping. (A) In a good surgical outcome patient, we show resection zone, irritative zone, and uninvolved zone on the left. One electrode contact, in superior frontal gyrus, showed low levels of abnormality across all features and clinically determined to be normal as well. Another in inferior frontal gyrus showed higher abnormality and localized to the irritative zone. A final electrode in hippocampus showed the highest abnormality in alpha coherence, and was successfully resected as part of the seizure onset zone. (B) In a poor outcome patient, an irritative zone channel in middle temporal gyrus had higher abnormality scores than a non-involved channel in middle frontal gyrus, or a resected seizure onset zone channel in inferior temporal gyrus. (C) We cross-validated a random forest model using both univariate and bivariate features for predicting seizure onset zone versus uninvolved brain. We found an AUC of 0.81 which was higher than either using only univariate features (AUC = 0.79) or only bivariate features (AUC = 0.73).

6.5 Discussion

In this study we construct and validate an iEEG atlas, illustrating its potential utility towards data-driven mapping of abnormal brain regions in epilepsy. By comparing features of activity and connectivity in test subjects to those of putatively normal brain, we reliably identify clinically abnormal regions such as the seizure onset zone. We show that lesional patients have a higher level of overall abnormality than non-lesional patients, that in different ROI the seizure onset zone may be identified from normal brain by unique features, and that abnormalities in functional connectivity are particularly useful for identifying mesial temporal lobe SOZs. Finally, we publicly share our validated iEEG atlas containing both normal and abnormal channels to allow others to build upon our methods.

Developing new interventions for patients with drug-resistant seizures requires a better understanding of the relationship between the abnormal brain and epilepsy. While

distinct epilepsy syndromes and etiologies may result in different electrographic seizure onset patterns[19], recognizing their impact on interictal activity is more difficult. For our current study in which we probe quantitative abnormality of uninvolved, irritative, and seizure onset zones, we find that univariate spectral content is most useful in temporal and frontal neocortex, entropy is useful in frontal cortex, and connectivity is useful in mesial temporal lobe. Thus, the ability of normative atlas mapping to leverage different quantitative metrics in a single, flexible framework may allow it to serve as a generalizable tool that works in many clinical scenarios. It is unlikely that abnormalities that underlie seizure onset regions across the brain manifest a single, unified characteristic. Furthermore, our finding that univariate and bivariate measures provide complementary information light on an important debate in epilepsy surgery: It is not always clear whether epilepsy should be viewed as a focal problem within a network (and thus probed by univariate methods) or as a broader, integrated network disorder (and thus studied through connectivity)[191]. Our atlas provides methods for assessing the extent of each of these hypotheses and whether certain patients may have pathology that is predominantly focal or distributed.

Our study adds to a growing body of work on normative atlas analysis for intracranial EEG. We directly build upon the work of Frauscher et al. 2018, in which they demonstrated the approach of aggregating curated normative data and showed that spectral features including peak frequencies differ throughout cortical regions. Our extension demonstrates that this technique can indeed detect abnormality. We also adapt multi-patient functional connectivity approaches from Betzel et al., 2018, in which they

demonstrated that integrating adjacency matrices across patients provides group-level ECoG connectivity that closely mirrors functional MRI connectivity. Here, we have confirmed that deviations from expected, normative connectivity indeed can distinguish potentially epileptogenic regions. Overall, our work of normative mapping is closest in methodology to Taylor et al., 2021 in which they use an atlas approach of univariate spectral features to identify epileptogenic regions by Z-score as we have. However, besides confirming aspects of their results on a completely separate dataset we also extend this by showing the utility of other univariate features such as entropy and that connectivity may improve the overall approach. Ultimately, the large, multi-center nature of these efforts including ours may lay the groundwork for normative iEEG mapping to serve as an important clinical tool in the near future.

6.5.1 Limitations & future directions

Our study represents a significant foray into using full-brain atlas approaches of iEEG to guide epilepsy surgery, but comes with several limitations. One of these is that the assumption of normal activity and connectivity in uninvolved regions may not be universally valid. For example, uninvolved regions may exert inhibitory activity on a seizure focus as part of an inhibitory surround and could have abnormal activity as a result of this. Indeed, even in good outcome patients, abnormalities in structural connectivity are observed outside of the resection zone[186] and may not be significant enough to cause seizures. Future work and data sharing across centers could permit the incorporation of patients that received iEEG for purposes such as facial pain[192] or other conditions in which no epileptic foci exist.

Another significant limitation is the spatial scale at which we map activity and connectivity. In this study we focus on connectivity between gross anatomic regions, while many of the aberrations in connectivity exist within ROI and thus could not be probed by our current approach. However, the natural step would be to break up AAL regions into even smaller parcellations, and such atlases of up to 600 regions have been used by our group in different contexts[193]. Even with 166 patients which represents one of the largest published patient cohorts in a computational iEEG study, we did not have dense enough coverage in every region to have greater spatial granularity. However, as we show the power of building upon previously established datasets, we hope for others to further build upon our normative atlas mapping which could allow this approach to be taken in the future. This step could also ensure that all possible connections in the functional connectivity atlas are represented and have sufficient samples for the median and variance in connectivity to be more accurately estimated.

To expand the utility of our atlas of iEEG connectivity epilepsy surgery, the natural next step is to join our approach with similar methods in neuroimaging. Doing so may allow for identifying abnormal connections in regions that aren't typically implanted by iEEG such as central gray matter structures including the thalamus which is implicated in seizure generation and propagation. Furthermore, combining our approach with DTI could help answer whether functional connectivity abnormalities as measured by iEEG are highly correlated to structural abnormalities, or whether these phenomena are only loosely related. Using full-brain iEEG atlas approaches could enable clinicians to better understand the

relationship between structure and function and might ultimately allow iEEG to be replaced with fully non-invasive studies in some patients.

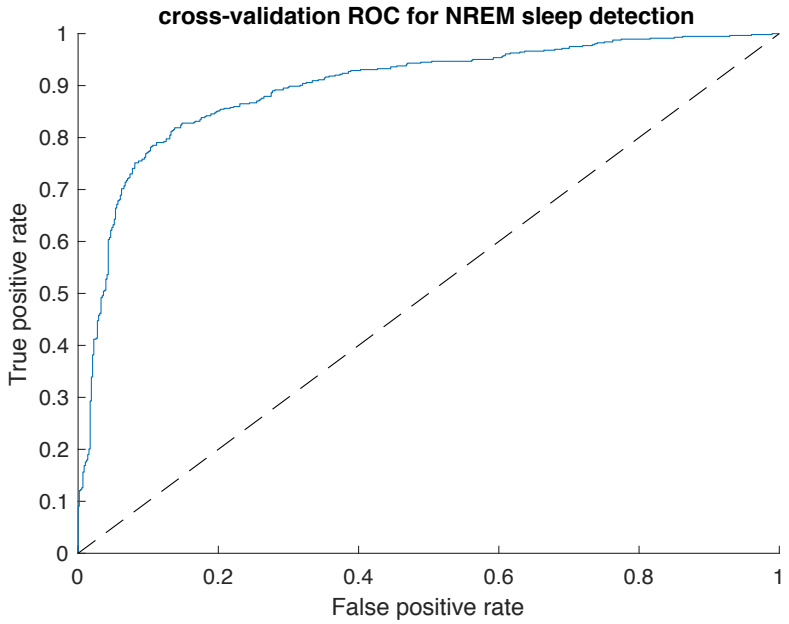
6.5.2 Conclusions

In conclusion, we have demonstrated the feasibility of using an atlas of iEEG for brain mapping in patients with drug-resistant epilepsy undergoing surgical evaluation. We validate the concordance of a multi-center atlas, and illustrate that clinically abnormal regions including seizure onset zone and irritative zone are detectable by Z-score. We show that connectivity augments univariate measures of activity, particularly in the mesial temporal lobe, and show good classification of seizure onset zone versus normal channels. Through extensive data sharing we may soon reach adequate accuracy and brain coverage to use the atlas method as a preferred approach as a quantitative method for identifying the epileptogenic zone from intracranial EEG which could offer a substantial improvement for epilepsy surgery outcomes.

6.6 Supplemental Materials

Supplemental Methods 1: Wake/Sleep Classification.

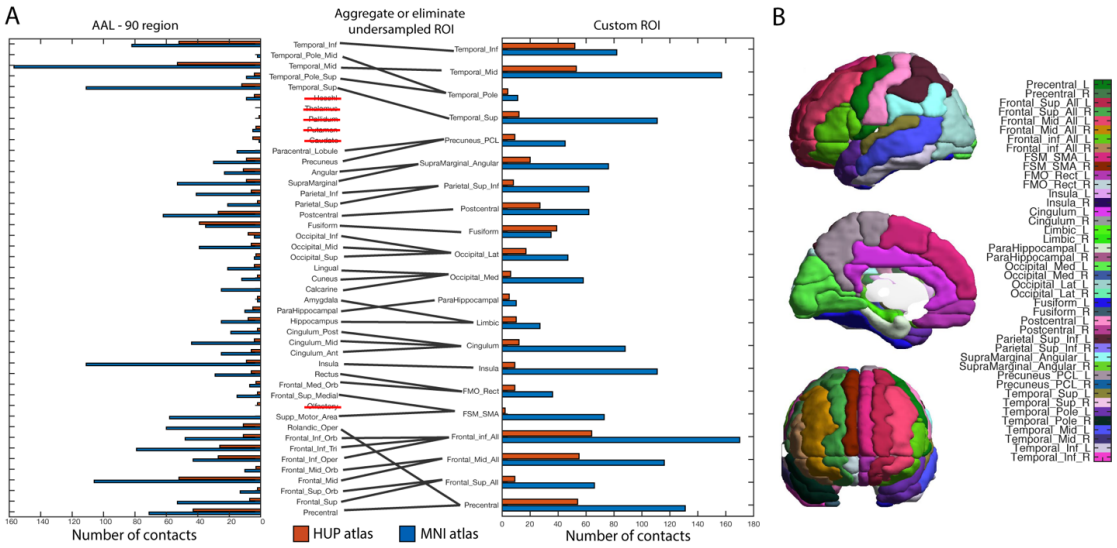
In many intracranial EEG (iEEG) patients at the Hospital of the University of Pennsylvania (HUP), we did not have simultaneous scalp EEG and/or adequate clinical staging of sleep. Thus, we needed to develop a sleep detector based on only iEEG signals. We used the data from Montreal Neurological Institute to train a detector of NREM sleep based on bandpower in the canonical frequency bands (delta: 0.5-4 Hz, theta: 4-8 Hz, alpha 8-13 Hz, beta 13-30 Hz). We used a logistic regression model and achieved 0.90 AUC on a per-channel basis in the MNI dataset. In a subset of our patients, limited clinical sleep annotations were available. We tested the accuracy of our sleep detector in these patients, and found a performance of 0.77 AUC on a per-channel basis and 0.80 AUC on a per-time basis. We also selected time windows during normal waking hours where possible (8 am – 8 pm) to minimize the effects of any misclassification on the part of the NREM sleep detector.



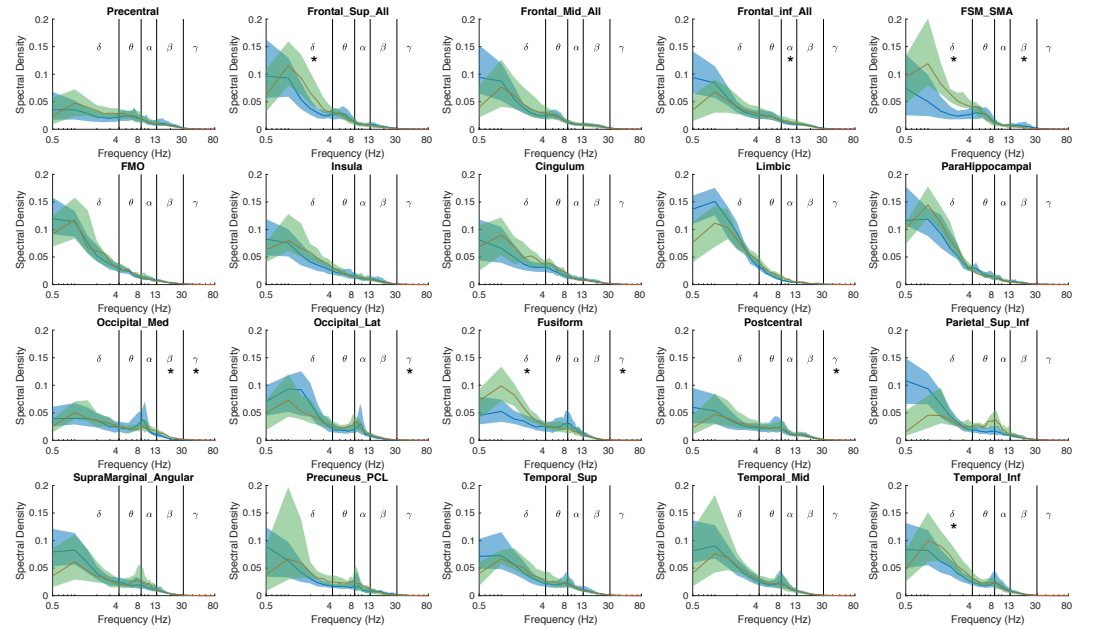
Supplementary Figure 6.1. Cross-validation receiver operating characteristic for NREM sleep detector. Area under the receiver-operating characteristic curve on cross-validation is 0.90.

Structure name	Rows	Columns	Unit/descriptor
wake_clip	12000 = samples (60 seconds at 200 Hz)	3431 channels	iEEG data in bipolar montage, filtered 0.5-80 Hz
mni_coords	3431 = channels	3 coordinates (X,Y,Z)	MNI-152 coordinates of each bipolar centroid
patient_no	3431 = channels	1	Patient number for each channel
resected_ch	3431 = channels	1	Whether or not each channel was resected/ablated (1 = resected/ablated)
soz_ch	3431 = channels	1	Whether or not each channel was in the seizure onset zone (1 = within SOZ)
spike_24h	3431 = channels	1	Estimated spike rate per 24 hours (we defined irritative zone as spike_24h>24)

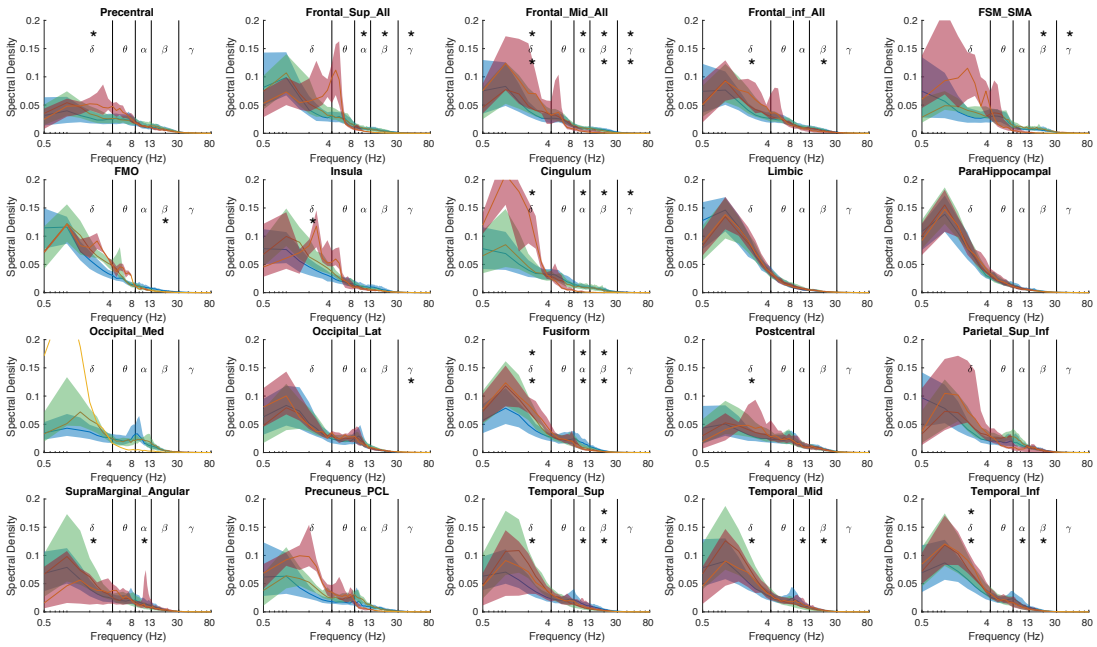
Supplementary Table 6.1. Data fields for the HUP iEEG Open Atlas. The Matlab file contains the structures listed in the leftmost column. Rows and columns of each matrix are described above.



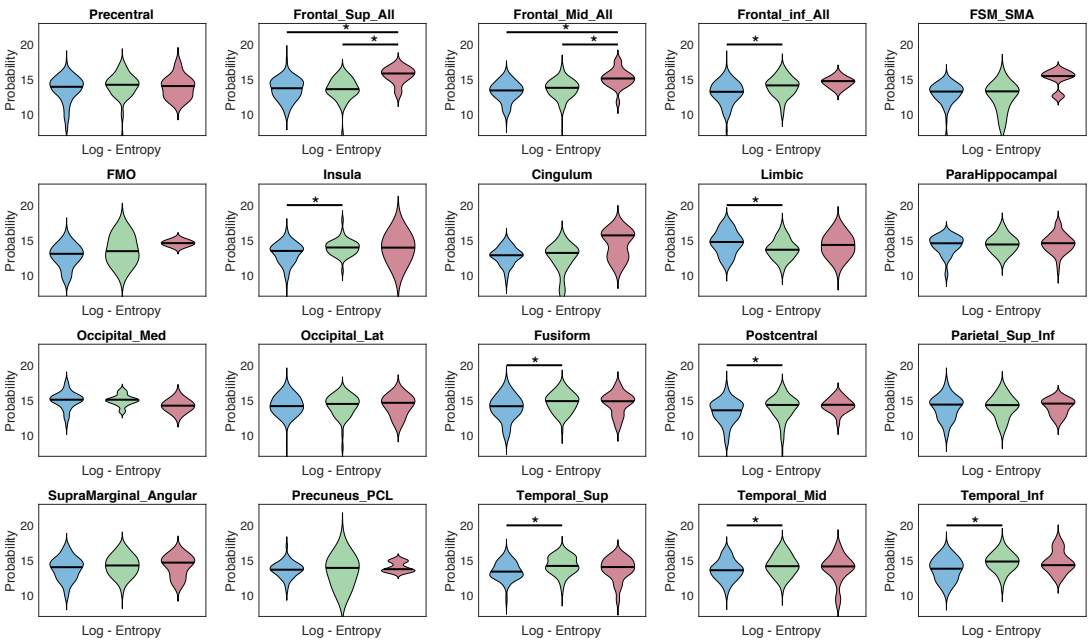
Supplementary Figure 6.2. Reduced AAL atlas. (A) In the original 90 regions of non-cerebellar AAL atlas, we eliminated regions which had very few channels, and combined others to reduce the total number of ROI and increase the samples per region. (B) A rendering of the ROI in our reduced atlas.



Supplementary Figure 6.3: Normal power spectra in HUP vs MNI. Blue: MNI, Green: HUP. * = $p < 0.0005$ as we corrected for 100 comparisons. In most regions, power spectral density between HUP and MNI atlases were similar and thus we combined them into a composite atlas.



Supplementary Figure 6.4: Power spectra in irritative and seizure onset zone versus normative composite atlas. Blue: Composite normative atlas, Green: Irritative zone, Red: Seizure onset zone. * = $p < 0.0005$ as we corrected for 100 comparisons. The * is located above the Greek letters indicating frequency band if SOZ vs. composite atlas is significant, and below the Greek letters if irritative zone vs. composite atlas is significant.



Supplementary Figure 6.5: Entropy in irritative and seizure onset zone versus normative composite atlas. Blue: Composite normative atlas, Green: Irritative zone, Red: Seizure onset zone. * = $p < 0.0005$ as we corrected for 100 comparisons. We are plotting log negative entropy here and use that in our atlas as it approaches a normal distribution that can be estimated more accurately. Higher values correspond to signals that are more ordered (and therefore less entropy)

CHAPTER 7: Conclusions and Future Directions

In this thesis, I developed and validated methods for the quantitative localization of the epileptogenic zone from intracranial EEG. These methods, which leverage both interictal and ictal data could eventually serve as supplementary biomarkers during epilepsy surgical evaluation. Overall, I have made the following contributions with this dissertation:

Contribution 1: Methods for simulating different surgical approaches and outcome prediction

The virtual resection method is a flexible, data-driven framework for pre-surgical prognostication using the change of synchronizability at seizure onset as a network biomarker of surgical outcome. Control centrality can serve as a method to simulate different interventions and select the surgical approach most likely to render the patient seizure free. Furthermore, I have validated that virtual resection can reveal subtle aspects of seizure dynamics that are not evident from traditional evaluation of onset patterns.

Contribution 2: Methods for assessing the impact of sampling bias

One of the primary limitations in deploying network models as clinical tools is the interaction between which and how many brain regions are sampled and the values of different metrics. From the work presented here, we now have ways of accounting for sampling bias, such as using jackknife resampling as in Chapter 4. Furthermore, I also show that correcting for electrode type and internodal distance as in Chapter 5 can help mitigate bias introduced by differences in ECoG and SEEG though the epileptogenic zone still has distinct characteristics in patients which underwent resection versus ablation.

These tools will be essential in the future study and eventual clinical deployment of iEEG network models.

Contribution 3: An intracranial EEG atlas for assessing abnormalities in neural activity and connectivity

Another major limitation of previous work is that there are no healthy controls with iEEG so it is difficult to estimate what normal activity and connectivity should be throughout the brain. By aggregating and carefully selecting channels without spikes, seizure onset zones, and other abnormalities, I developed an iEEG atlas and validated it against a similar dataset published by another center. I also provide methods for probing the relationship between activity and connectivity by mapping the abnormality of each node. This approach is likely to become even more powerful as datasets get larger over time.

Contribution 4: Open source data and code

Epilepsy surgery and intracranial EEG is highly specialized and is only available at a relatively small number of hospitals worldwide. Even at these centers, clinical volumes are not high, therefore the amount of data that any one research group can access is typically too small to account for the high level of variability across patients. Therefore, extensive data-sharing efforts are required to advance quantitative iEEG towards clinical practice. As part of this thesis, we are releasing full datasets from over 60 patients, including imaging and clinical metadata, which may serve as the 'gold-standard' dataset for future studies. I am also releasing code associated with each method found in this dissertation so that others may build upon our efforts.

Future directions

The work presented in this thesis provides the foundation for multiple areas of future research. One important direction, built upon the work of Chapter 3 could use the framework of synchronizability and control centrality to understand why some seizures generalize and others do not, and determine how long before seizures synchronizability elevates from its baseline levels. Another important direction is extending the iEEG atlas framework presented in Chapter 6 to other periods of time, such as during sleep, or before seizures. Overall, many of the methods I have developed must also be compared and contrasted with findings form quantitative image analysis including diffusion imaging and functional MRI in order to gain some information about epileptic networks non-invasively. Finally, we must construct hardware and software to allow for the real-time deployment of these models in a clinical setting to allow for trials to translate these methods from bench to bedside.

References

1. Engel, J. Surgery for Seizures. *N. Engl. J. Med.* **334,** 647–653 (1996).

2. Téllez-Zenteno, J. F., Ronquillo, L. H., Moien-Afshari, F. & Wiebe, S. Surgical outcomes in lesional and non-lesional epilepsy: A systematic review and meta-analysis. *Epilepsy Res.* **89,** 310–318 (2010).

3. Parvizi, J. & Kastner, S. Human Intracranial EEG: Promises and Limitations. *Nat. Neurosci.* **21,** 474–483 (2018).

4. Khambhati, A. N., Davis, K. A., Lucas, T. H., Litt, B. & Bassett, D. S. Virtual Cortical Resection Reveals Push-Pull Network Control Preceding Seizure Evolution. *Neuron* **91,** 1170–1182 (2016).

5. Stam, C. J. & Reijneveld, J. C. Graph theoretical analysis of complex networks in the brain. *Nonlinear Biomed. Phys.* **1,** (2007).

6. Schindler, K. A., Bialonski, S., Horstmann, M. T., Elger, C. E. & Lehnertz, K. Evolving functional network properties and synchronizability during human epileptic seizures. *Chaos* **18,** (2008).

7. Jeha, L. E. *et al.* Surgical outcome and prognostic factors of frontal lobe epilepsy surgery. *Brain* **130,** 574–584 (2007).

8. Kwan, P. & Sperling, M. R. Refractory seizures: Try additional antiepileptic drugs (after two have failed) or go directly to early surgery evaluation? *Epilepsia* **50,** 57–62

9. Rosenow, F. & Lüders, H. Presurgical evaluation of epilepsy. *Brain* **124,** 1683–1700 (2001).

10. Kwan, P., Schachter, S. C. & Brodie, M. J. *Drug-Resistant Epilepsy*. *The New England Journal of Medicine* (2011).

11. Lüders, H. O., Najm, I., Nair, D., Widdess-Walsh, P. & Bingman, W. Definition and localization of the epileptogenic zone The epileptogenic zone: general principles. *Epileptic Disord.* **8,** S1-9 (2006).

12. Jehi, L. The Epileptogenic Zone: Concept and Definition. *Epilepsy Curr.* **18,** 12–16 (2018).

13. Holmes, M. D., Kutsy, R. L., Ojemann, G. A., Wilensky, A. J. & Ojemann, L. M. Interictal, unifocal spikes in refractory extratemporal epilepsy predict ictal origin and postsurgical outcome. *Clin. Neurophysiol.* **111,** 1802–1808 (2000).

14. Krendl, R., Lurger, S. & Baumgartner, C. Absolute spike frequency predicts surgical outcome in TLE with unilateral hippocampal atrophy. *Neurology* **71,** 413–418 (2008).

15. Malow, B. A., Selwa, L. M., Ross, D. & Aldrich, M. S. Lateralizing Value of Interictal Spikes on Overnight Sleep-EEG Studies in Temporal Lobe Epilepsy. *Epilepsia* **40,** 1587–1592 (1999).

16. Iwasaki, M. *et al.* Surgical Implications of Neuromagnetic Spike Localization in Temporal Lobe Epilepsy. *Epilepsia* **43,** 415–424 (2002).

17. Hufnagel, A., Dumpelmann, M., Zentner, J., Schijns, O. & Elger, C. E. Clinical Relevance of Quantified Intracranial Interictal Spike Activity in Presurgical Evaluation of Epilepsy. *Epilepsia* **41,** 467–478 (2000).

18. Cai, Z. *et al.* Noninvasive high-frequency oscillations riding spikes delineates epileptogenic sources. *Proc. Natl. Acad. Sci. U. S. A.* **118,** (2021).

19. Lagarde, S. *et al.* The repertoire of seizure onset patterns in human focal epilepsies: Determinants and prognostic values. *Epilepsia* **60,** 85–95 (2019).

20. Perucca, P., Dubeau, F. & Gotman, J. Intracranial electroencephalographic seizure-onset patterns: Effect of underlying pathology. *Brain* **137,** 183–196 (2014).

21. Miller, B. A., Salehi, A., Limbrick, D. D. & Smyth, M. D. Applications of a robotic stereotactic arm for pediatric epilepsy and neurooncology surgery. *J. Neurosurg. Pediatr.* **20,** 364–370 (2017).

22. Campbell Arnold, T. *et al.* Remote effects of temporal lobe epilepsy surgery: long-term morphological changes after surgical resection. *medRxiv* 2020.11.05.20224873 (2020). doi:10.1101/2020.11.05.20224873

23. Nuwer, M. Assessment of digital EEG, quantitative EEG, and EEG brain mapping: Report of the American Academy of Neurology and the American Clinical Neurophysiology Society. *Neurology* **49,** 277–292 (1997).

24. Lieb, J. P. *et al.* Quantitative analysis of depth spiking in relation to seizure foci in patients with temporal lobe epilepsy. *Electroencephalogr. Clin. Neurophysiol.* **44,** 641–663 (1978).

25. Alarcon, G., Binnie, C. D., Elwes, R. D. C. & Polkey, C. E. Power spectrum and intracranial EEG patterns at seizure onset in partial epilepsy. *Electroencephalogr. Clin. Neurophysiol.* **94,** 326–337 (1995).

26. Esteller, R., Echauz, J., Tcheng, T., Litt, B. & Pless, B. Line length: An efficient feature for seizure onset detection. *Annu. Int. Conf. IEEE Eng. Med. Biol.* **2,** 1707–

1710 (2001).

27. Rummel, C. *et al.* Resected Brain Tissue, Seizure Onset Zone and Quantitative EEG Measures: Towards Prediction of Post-Surgical Seizure Control. *PLoS One* **10,** (2015).

28. Groppe, D. M. *et al.* Dominant frequencies of resting human brain activity as measured by the electrocorticogram. *Neuroimage* **79,** 223–233 (2013).

29. Andrzejak, R. G. *et al.* Improved spatial characterization of the epileptic brain by focusing on nonlinearity. *Epilepsy Res.* **69,** 30–44 (2006).

30. Grinenko, O. *et al.* A fingerprint of the epileptogenic zone in human epilepsies. *Brain* **141,** 117–131 (2018).

31. Li, J. *et al.* Learning to define an electrical biomarker of the epileptogenic zone. *Hum. Brain Mapp.* **41,** 429–441 (2020).

32. Bartolomei, F., Chauvel, P. & Wendling, F. Epileptogenicity of brain structures in human temporal lobe epilepsy: a quantified study from intracerebral EEG. *Brain* **131,** 1818–30 (2008).

33. David, O. *et al.* Imaging the seizure onset zone with stereo-electroencephalography. *Brain* **134,** 2898–2911 (2011).

34. Sevy, A. *et al.* Beyond the lesion: The epileptogenic networks around cavernous angiomas. *Epilepsy Res.* **108,** 701–708 (2014).

35. Bonini, F. *et al.* Epileptogenic networks in seizures arising from motor systems. *Epilepsy Res.* **106,** 92–102 (2013).

36. Marchi, A. *et al.* Occipital and occipital "plus" epilepsies: A study of involved epileptogenic networks through SEEG quantification. *Epilepsy Behav.* **62,** 104–114 (2016).

37. Wendling, F., Bartolomei, F., Bellanger, J. J., Bourien, J. & Chauvel, P. Epileptic fast intracerebral EEG activity: evidence for spatial decorrelation at seizure onset. doi:10.1093/brain/awg144

38. Antony, A. R. *et al.* Functional Connectivity Estimated from Intracranial EEG Predicts Surgical Outcome in Intractable Temporal Lobe Epilepsy. *PLoS One* **8,** e77916 (2013).

39. Bartolomei, F. *et al.* Seizures of temporal lobe epilepsy: identification of subtypes by coherence analysis using stereo-electro-encephalography. *Clin. Neurophysiol.* **110,** 1741–54 (1999).

40. Varotto, G., Tassi, L., Franceschetti, S., Spreafico, R. & Panzica, F. Epileptogenic networks of type II focal cortical dysplasia: A stereo-EEG study. *Neuroimage* **61,** 591–598 (2012).

41. Cotic, M. *et al.* Mapping the coherence of ictal high frequency oscillations in human extratemporal lobe epilepsy. *Epilepsia* **56,** 393–402 (2015).

42. Weiss, S. A. *et al.* Seizure localization using ictal phase-locked high gamma A retrospective surgical outcome study. *Neurology* **84,** 2320–8 (2015).

43. Elahian, B., Yeasin, M., Mudigoudar, B., Wheless, J. W. & Babajani-Feremi, A. Identifying seizure onset zone from electrocorticographic recordings: A machine learning approach based on phase locking value. *Seizure* **51,** 35–42 (2017).

44. Narasimhan, S. *et al.* Seizure-onset regions demonstrate high inward directed connectivity during resting-state: An SEEG study in focal epilepsy. *Epilepsia* **61,** 2534–2544 (2020).

45. Bink, H. *et al.* Spatiotemporal evolution of focal epileptiform activity from surface and laminar field recordings in cat neocortex. *J. Neurophysiol.* **119,** 2068–2081 (2018).

46. Bernhardt, B. C., Bonilha, L. & Gross, D. W. Network analysis for a network disorder: The emerging role of graph theory in the study of epilepsy. *Epilepsy and Behavior* **50,** (2015).

47. Kramer, M. A. & Cash, S. S. Epilepsy as a Disorder of Cortical Network Organization. *Neurosci.* (2012). doi:10.1177/1073858411422754

48. Bernhardt, B. C., Hong, S., Bernasconi, A. & Bernasconi, N. Imaging structural and functional brain networks in temporal lobe epilepsy. *Front. Hum. Neurosci.* **7,** 624 (2013).

49. Bassett, D. S., Zurn, P. & Gold, J. I. On the nature and use of models in network neuroscience. *Nat. Rev. Neurosci.* **19,** 566–578 (2018).

50. Bullmore, E. & Sporns, O. Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10,** 186–198 (2009).

51. Lagarde, S. *et al.* Interictal stereotactic-EEG functional connectivity in refractory focal epilepsies. *Brain* **141,** 2966–2980 (2018).

52. Shah, P. *et al.* High interictal connectivity within the resection zone is associated with favorable post-surgical outcomes in focal epilepsy patients. *NeuroImage Clin.* **23,** 101908 (2019).

53. Wilke, C., Worrell, G. & He, B. Graph analysis of epileptogenic networks in human partial epilepsy. *Epilepsia* **52,** 84–93 (2011).

54. Burns, S. P. *et al.* Network dynamics of the brain and influence of the epileptic seizure onset zone. *Proc. Natl. Acad. Sci. U. S. A.* **111,** E5321–E5330 (2014).

55. Kini, L. G. *et al.* Virtual resection predicts surgical outcome for drug-resistant epilepsy. *Brain* **142,** 3892–3905 (2019).

56. Khambhati, A. N., Sizemore, A. E., Betzel, R. F. & Bassett, D. S. Modeling and interpreting mesoscale network dynamics. *Neuroimage* (2017). doi:10.1016/j.neuroimage.2017.06.029

57. Lee, D. D. & Seung, H. S. Algorithms for Non-negative Matrix Factorization. in *Conference on Neural Information Processing Systems* (2000).

58. Lee, D. D. & Seung, H. S. Learning the parts of objects by non-negative matrix factorization. *Nature* **401,** 788–791 (1999).

59. Khambhati, A. N. *et al.* Recurring functional interactions predict network architecture of interictal and ictal states in neocortical epilepsy. *eNeuro* 100000001–1554488 (2017). doi:10.1523/ENEURO.0091-16.2017

60. Lopes, M. A. *et al.* An optimal strategy for epilepsy surgery: Disruption of the rich-club? *PLOS Comput. Biol.* **13,** e1005637 (2017).

61. Goodfellow, M., Schindler, K. & Baier, G. Self-organised transients in a neural mass model of epileptogenic tissue dynamics. *Neuroimage* **59,** 2644–2660 (2012).

62. Touboul, J., Wendling, F., Chauvel, P. & Faugeras, O. Neural mass activity, bifurcations, and epilepsy. *Neural Computation* **23,** 3232–3286 (2011).

63. Wendling, F., Benquet, P., Bartolomei, F. & Jirsa, V. Computational models of epileptiform activity. *Journal of Neuroscience Methods* **260,** 233–251 (2016).

64. Id, S. A., Bartolomei, F., Guye, M. & Jirsa Id, V. Optimization of surgical intervention outside the epileptogenic zone in the Virtual Epileptic Patient (VEP). *PLOS Comput. Biol.* **15,** (2019).

65. Sinha, N. *et al.* Predicting neurosurgical outcomes in focal epilepsy patients using computational modelling. *Brain* **140,** 319–332 (2017).

66. Goodfellow, M. *et al.* Estimation of brain network ictogenicity predicts outcome from epilepsy surgery. *Sci. Rep.* **6,** 29215 (2016).

67. Wendling, F., Bartolomei, F., Bellanger, J. J. & Chauvel, P. Epileptic fast activity

can be explained by a model of impaired GABAergic dendritic inhibition. *Eur. J. Neurosci.* **15,** 1499–1508 (2002).

68. Blenkinsop, A., Valentin, A., Richardson, M. P. & Terry, J. R. The dynamic evolution of focal-onset epilepsies - combining theoretical and clinical observations. *Eur. J. Neurosci.* **36,** 2188–2200 (2012).

69. Kini, L. G., Davis, K. A. & Wagenaar, J. B. Data integration: Combined Imaging and Electrophysiology data in the cloud. *Neuroimage* **124,** 1175–1181 (2016).

70. Wagenaar, J. B., Brinkmann, B. H., Ives, Z., Worrell, G. A. & Litt, B. A multimodal platform for cloud-based collaborative research. in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)* 1386–1389 (IEEE, 2013). doi:10.1109/NER.2013.6696201

71. Conrad, E. C. *et al.* The sensitivity of network statistics to incomplete electrode sampling on intracranial EEG. *Netw. Neurosci.* **4,** 484–506 (2020).

72. Owen, T. W. *et al.* Multivariate white matter alterations are associated with epilepsy duration. *Eur. J. Neurosci.* ejn.15055 (2020). doi:10.1111/ejn.15055

73. Jirsa, V. K. *et al.* The Virtual Epileptic Patient: Individualized whole-brain models of epilepsy spread. *Neuroimage* **145,** 377–388 (2017).

74. Attiah, M. A., Paulo, D. L., Danish, S. F., Stein, S. C. & Mani, R. Anterior temporal lobectomy compared with laser thermal hippocampectomy for mesial temporal epilepsy: A threshold analysis study. *Epilepsy Res.* **115,** 1–7 (2015).

75. Willie, J. T. *et al.* Real-time magnetic resonance-guided stereotactic laser amygdalohippocampotomy for mesial temporal lobe epilepsy. *Neurosurgery* **74,** 569–584 (2014).

76. Thomas, G. P. & Jobst, B. C. Critical appraisal of the neuropace rns system in medically. *Med. Devices Evid. Res.* **8,** 405–411 (2015).

77. Khambhati, A. N. *et al.* Dynamic Network Drivers of Seizure Generation, Propagation and Termination in Human Neocortical Epilepsy. *PLOS Comput. Biol.* **11,** e1004608 (2015).

78. Besson, P. *et al.* Anatomic consistencies across epilepsies: a stereotactic-EEG informed high-resolution structural connectivity study. *Brain* **140,** 2639–2652 (2017).

79. Burns, Samuel P.; Santaniello, Sabato; Yaffe, Robert B.; Jouny, Christophe C.; Crone, Nathan E.; Bergey, Gregory K., Anderson, William S.; Sarma, S. V. . Network dynamics of the brain and influence of the epileptic seizure onset zone.

Proc. Natl. Acad. Sci. (2014). doi:10.1073/pnas.1401752111

80. Taylor, P. N. *et al.* Optimal control based seizure abatement using patient derived connectivity. *Front. Neurosci.* **9,** 202 (2015).

81. Jirsa, V. K., Stacey, W. C., Quilichini, P. P., Ivanov, A. I. & Bernard, C. On the nature of seizure dynamics. *A J. Neurol.* doi:10.1093/brain/awu147

82. Litt, B. *et al.* Epileptic Seizures May Begin Hours in Advance of Clinical Onset: A Report of Five Patients. *Neuron* **30,** 51–64 (2001).

83. Vakharia, V. N. *et al.* Getting the Best Outcomes from Epilepsy Surgery. *Ann. Neurol.* **83,** 676–690 (2018).

84. Jin, B., So, N. K. & Wang, S. Advances of Intracranial Electroencephalography in Localizing the Epileptogenic Zone. *Neurosci. Bull.* **32,**

85. Lagarde, S. *et al.* Seizure-onset patterns in focal cortical dysplasia and neurodevelopmental tumors: Relationship with surgical prognosis and neuropathologic subtypes. *Epilepsia* **57,** 1426–1435 (2016).

86. Noe, K. *et al.* Long-term Outcomes After Nonlesional Extratemporal Lobe Epilepsy Surgery. *JAMA Neurol.* **70,** 1003 (2013).

87. Azarion, A. A. *et al.* An open-source automated platform for three-dimensional visualization of subdural electrodes using CT-MRI coregistration. *Epilepsia* **55,** 2028–2037 (2014).

88. Avants, B. B. *et al.* A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage* **54,** 2033–2044 (2011).

89. Yushkevich, P. A. *et al.* User-guided 3D active contour segmentation of anatomical structures: Significantly improved efficiency and reliability. *Neuroimage* **31,** 1116–1128 (2006).

90. Buzsáki, G. & Mizuseki, K. The log-dynamic brain: how skewed distributions affect network operations. *Nat. Rev. Neurosci.* (2014). doi:10.1038/nrn3687

91. Ramsay, J. & Silverman, B. *Functional Data Analysis*. (Springer, 1997).

92. DeLong, E. R., DeLong, D. M. & Clarke-Pearson, D. L. Comparing the Areas under Two or More Correlated Receiver Operating Characteristic Curves: A Nonparametric Approach. *Biometrics* **44,** 837 (1988).

93. Tomlinson, S. B., Porter, B. E. & Marsh, E. D. Interictal network synchrony and local heterogeneity predict epilepsy surgery outcome among pediatric patients.

Epilepsia **58,** 402–411 (2017).

94. Kramer, M. A., Kolaczyk, E. D. & Kirsch, H. E. Emergent network topology at seizure onset in humans. *Epilepsy Res.* **79,** 173–186 (2008).

95. Ray, S., Crone, N. E., Niebur, E., Franaszczuk, P. J. & Hsiao, S. S. Neural correlates of high-gamma oscillations (60-200 Hz) in macaque local field potentials and their potential implications in electrocorticography. *J. Neurosci.* **28,** 11526–36 (2008).

96. Baldassano, S. *et al.* Cloud computing for seizure detection in implanted neural devices. *J. Neural Eng.* **16,** (2019).

97. Kalitzin, S., Koppert, M., Petkov, G., Velis, D. & da Silva, F. L. Computational model prospective on the observation of proictal states in epileptic neuronal systems. *Epilepsy Behav.* **22,** S102–S109 (2011).

98. Hudgins, E. *et al.* Focal Seizures Induced by Intracranial Electroencephalogram Grids. *Cureus* **8,** e831 (2016).

99. Sperling, M. R. Clinical challenges in invasive monitoring in epilepsy surgery. *Epilepsia* **38 Suppl 4,** S6-12 (1997).

100. Englot, D. J., Birk, H. & Chang, E. F. Seizure outcomes in non-resective epilepsy surgery: An update. *Neurosurg. rev.* (2017). doi:10.1007/s10143-016-0725-8

101. Proix, T. E., Bartolomei, F., Guye, M. & Jirsa, V. K. Individual brain structure and modelling predict seizure propagation. *Brain* 641–654 (2017). doi:10.1093/awx018

102. Liao, W. *et al.* Altered Functional Connectivity and Small-World in Mesial Temporal Lobe Epilepsy. *PLoS One* **5,** e8525 (2010).

103. Pedersen, M., Omidvarnia, A. H., Walz, J. M. & Jackson, G. D. Increased segregation of brain networks in focal epilepsy: An fMRI graph theory finding. *NeuroImage. Clin.* **8,** 536–42 (2015).

104. Paldino, M. J., Zhang, W., Chu, Z. D. & Golriz, F. Metrics of brain network architecture capture the impact of disease in children with epilepsy. *NeuroImage Clin.* **13,** 201–208 (2017).

105. Panzica, F., Varotto, G., Rotondi, F., Spreafico, R. & Franceschetti, S. Identification of the epileptogenic zone from stereo-EEG signals: A connectivity-graph theory approach. *Front. Neurol.* **4 NOV,** 6–11 (2013).

106. Fletcher, J. M. K. & Wennekers, T. From Structure to Activity: Using Centrality

Measures to Predict Neuronal Activity. *Int. J. Neural Syst.* **28,** (2018).

107. Ponten, S. C., Bartolomei, F. & Stam, C. J. Small-world networks and epilepsy: Graph theoretical analysis of intracerebrally recorded mesial temporal lobe seizures. *Clin. Neurophysiol.* **118,** 918–927 (2007).

108. Albert, R., Albert, I. & Nakarado, G. L. Structural vulnerability of the North American power grid. *Phys. Rev. E* **69,** 025103 (2004).

109. Albert, R., Jeong, H. & Barabási, A.-L. Error and attack tolerance of complex networks. *Nature* **406,** 378–382 (2000).

110. Guimerà, R. & Sales-Pardo, M. Missing and spurious interactions and the reconstruction of complex networks. *Proc. Natl. Acad. Sci. U. S. A.* **106,** 22073–22078 (2009).

111. Lu, L. & Zhou, T. Link Prediction in Complex Networks: A Survey. *Phys. A Stat. Mech. its Appl.* **390,** 1150–1170 (2010).

112. Proix, T., Jirsa, V. K., Bartolomei, F., Guye, M. & Truccolo, W. Predicting the spatiotemporal diversity of seizure propagation and termination in human focal epilepsy. doi:10.1038/s41467-018-02973-y

113. Shah, P. *et al.* High interictal connectivity within the resection zone is associated with favorable post-surgical outcomes in focal epilepsy patients. *NeuroImage Clin.* **23,** (2019).

114. Avants, B. B., Tustison, N. J., Wu, J., Cook, P. A. & Gee, J. C. An Open Source Multivariate Framework for n-Tissue Segmentation with Evaluation on Public Data. *Neuroinformatics* **9,** 381–400 (2011).

115. Sun, F. T., Arcot Desai, S., Tcheng, T. K. & Morrell, M. J. Changes in the electrocorticogram after implantation of intracranial electrodes in humans: The implant effect. *Clin. Neurophysiol.* **129,** 676–686 (2018).

116. Schroeder, G. M. *et al.* Seizure pathways change on circadian and slower timescales in individual patients with focal epilepsy. *Proc. Natl. Acad. Sci. U. S. A.* **117,** 11048–11058 (2020).

117. Kramer, M. A. *et al.* Behavioral/Systems/Cognitive Emergence of Persistent Networks in Long-Term Intracranial EEG Recordings. *J. Neurosci.* (2011). doi:10.1523/JNEUROSCI.2287-11.2011

118. Arbabshirani, M. R. *et al.* Impact of autocorrelation on functional connectivity. *Neuroimage* **102,** 294–308 (2014).

119. Towle, V. L., Carder, R. K., Khorasani, L. & Lindberg, D. Electrocorticographic coherence patterns. *J. Clin. Neurophysiol.* **16,** 528–547 (1999).

120. Mitra, P. P. & Pesaran, B. Analysis of dynamic brain imaging data. *Biophys. J.* **76,** 691–708 (1999).

121. Bettus, G. *et al.* Interictal functional connectivity of human epileptic networks assessed by intracerebral EEG and BOLD signal fluctuations. *PLoS One* **6,** 20071 (2011).

122. Rubinov, M. & Sporns, O. Complex network measures of brain connectivity: Uses and interpretations. *Neuroimage* **52,** 1059–1069 (2010).

123. Latora, V. & Marchiori, M. Efficient Behavior of Small-World Networks. *Phys. Rev. Lett.* **87,** (2001).

124. Dijkstra, E. W. A note on two problems in connexion with graphs. *Numer. Math.* **1,** 269–271 (1959).

125. Opsahl, T., Agneessens, F. & Skvoretz, J. Node centrality in weighted networks: Generalizing degree and shortest paths. *Soc. Networks* **32,** 245–251 (2010).

126. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D.-U. Complex networks: Structure and dynamics. *Phys. Rep.* **424,** 175–308 (2006).

127. Holland, P. W. & Leinhardt, S. Transitivity in structural models of small groups. *Small Gr. Res.* **2,** 107–124 (1971).

128. Opsahl, T. & Panzarasa, P. Clustering in weighted networks. *Soc. Networks* **31,** 155–163 (2009).

129. Watts, D. J. & Strogatz, S. H. Collective dynamics of 'small-world9 networks. *Nature* **393,** 440–442 (1998).

130. Fornito, A., Zalesky, A. & Bullmore, E. T. *Fundamentals of Brain Network Analysis*. *Fundamentals of Brain Network Analysis* (Elsevier, 2016). doi:10.1016/b978-0-12-407908-3.09999-4

131. Freeman, L. C. A Set of Measures of Centrality Based on Betweenness. *Sociometry* **40,** 35 (1977).

132. Barrat, A., Marc, B. & Vespignani, A. *Dynamical Processes on Complex Networks*. (2008).

133. Murphy, K. R. & Davidshofer, C. O. Psychological Testing: Principles and Applications. (2005). Available at: https://www.pearson.com/us/higher-

education/program/Murphy-Psychological-Testing-Principles-and-Applications-6th-Edition/PGM190901.html. (Accessed: 21st June 2021)

134. Friedman, M. The Use of Ranks to Avoid the Assumption of Normality Implicit in the Analysis of Variance. *J. Am. Stat. Assoc.* **32,** 675–701 (1937).

135. Dunn, O. J. Multiple Comparisons Using Rank Sums. *Technometrics* **6,** 241–252 (1964).

136. Sidak, Z. Rectangular Confidence Regions for the Means of Multivariate Normal Distributions. *J. Am. Stat. Assoc.* **62,** 626 (1967).

137. Fisher, R. A. Frequency Distribution of the Values of the Correlation Coefficient in Samples from an Indefinitely Large Population. *Biometrika* **10,** 507 (1915).

138. Quenouille, M. H. Notes on Bias in Estimation. *Biometrika* **43,** 353–360 (1956).

139. Quenouille, M. H. Problems in Plane Sampling. *Ann. Math. Stat.* **20,** 355–375 (1949).

140. Tukey, J. Bias and Confidence in Not-quite Large Samples. *Ann. Math. Stat.* **29,** 614–623 (1958).

141. Albert, R., Albert, I. & Nakarado, G. L. Structural vulnerability of the North American power grid. *Phys. Rev. E - Stat. Nonlinear, Soft Matter Phys.* **69,** 025103 (2004).

142. Kossinets, G. Effects of missing data in social networks. *Soc. Networks* **28,** 247–268 (2006).

143. Jalili, M. Resiliency of EEG-based brain functional networks. *PLoS One* **10,** (2015).

144. Smith, J. A. & Moody, J. Structural effects of network sampling coverage I: Nodes missing at random. *Soc. Networks* **35,** 652–668 (2013).

145. Costenbader, E. & Valente, T. W. The stability of centrality measures when networks are sampled. *Soc. Networks* **25,** 283–307 (2003).

146. Wang, Y. *et al.* Interictal intracranial electroencephalography for predicting surgical success: The importance of space and time. *Epilepsia* epi.16580 (2020). doi:10.1111/epi.16580

147. Mullin, J. P., Sexton, D., Al-Omar, S., Bingaman, W. & Gonzalez-Martinez, J. Outcomes of Subdural Grid Electrode Monitoring in the Stereoelectroencephalography Era. *World Neurosurg.* **89,** 255–258 (2016).

148. Betzel, R. F. *et al.* Structural, geometric and genetic factors predict interregional brain connectivity patterns probed by electrocorticography. *Nat. Biomed. Eng.* 1 (2019). doi:10.1038/s41551-019-0404-5

149. Fan, L. *et al.* The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. *Cereb. Cortex* **26,** 3508–3526 (2016).

150. Greicius, M. D., Supekar, K., Menon, V. & Dougherty, R. F. Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cereb. Cortex* **19,** 72–78 (2009).

151. Pan, L., Zhou, T., Lü, L. & Hu, C. K. Predicting missing links and identifying spurious links via likelihood analysis. *Sci. Rep.* **6,** 1–10 (2016).

152. Duval, R. D., Christensen, K. & Spahiu, A. Bootstrapping a Terrorist Network. in *Political Networks Conference* (2010).

153. Ohara, K., Saito, K., Kimura, M. & Motoda, H. Critical Node Identification based on Articulation Point Detection for Uncertain Network. *Int. J. Netw. Comput.* **9,** 201–216 (2019).

154. Lusseau, D., Whitehead, H. & Gero, S. Incorporating uncertainty into the study of animal social networks. *Animal Behaviour* **75,** 1809–1815 (2008).

155. de Matos Simoes, R. & Emmert-Streib, F. Bagging statistical network inference from large-scale gene expression data. *PLoS One* **7,** 33624 (2012).

156. Reif, P. S., Strzelczyk, A. & Rosenow, F. The history of invasive EEG evaluation in epilepsy patients. *Seizure* **41,** 191–195 (2016).

157. Youngerman, B. E., Khan, F. A. & McKhann, G. M. Stereoelectroencephalography in epilepsy, cognitive neurophysiology, and psychiatric disease: Safety, efficacy, and place in therapy. *Neuropsychiatric Disease and Treatment* **15,** 1701–1716 (2019).

158. Neuroscientist, C., Kramer, M. A. & Cash, S. S. Overview of Epilepsy and Seizures, Rhythms, and Networks Different Kinds of Seizures and Implications for Spatial Characterization of Epilepsy Epilepsy as a Disorder of Cortical Network Organization. *Neurosci.* **18,** 360–372 (2012).

159. Scheid, B. H. *et al.* Synchronizability predicts effective responsive neurostimulation for epilepsy prior to treatment. *medRxiv* 2021.02.05.21250075 (2021). doi:10.1101/2021.02.05.21250075

160. Pizzo, F. *et al.* Epileptogenic networks in nodular heterotopia: A stereoelectroencephalography study. *Epilepsia* **58,** 2112–2123 (2017).

161. Cruzat, J. *et al.* The dynamics of human cognition: Increasing global integration coupled with decreasing segregation found using iEEG. *Neuroimage* **172,** 492–505 (2018).

162. Buch, V. P. *et al.* Network Brain-Computer Interface (nBCI): An Alternative Approach for Cognitive Prosthetics. *Front. Neurosci.* **12,** (2018).

163. Katz, J. S. & Abel, T. J. Stereoelectroencephalography Versus Subdural Electrodes for Localization of the Epileptogenic Zone: What Is the Evidence? *Neurotherapeutics* **16,** 59–66 (2019).

164. Avants, B., Tustison, N. & Song, G. Advanced Normalization Tools (ANTS). *Insight J.* 1–35 (2009).

165. Tzourio-Mazoyer, N. *et al.* Automated Anatomical Labeling of Activations in SPM Using a Macroscopic Anatomical Parcellation of the MNI MRI Single-Subject Brain. *Neuroimage* **15,** 273–289 (2002).

166. Newman, M. E. J. Modularity and community structure in networks. *Proc. Natl. Acad. Sci. U. S. A.* **103,** 8577–8582 (2006).

167. De Meo, P., Ferrara, E., Fiumara, G. & Provetti, A. Generalized Louvain method for community detection in large networks. *arXiv* (2012).

168. Guimerà, R. & Nunes Amaral, L. A. Cartography of complex networks: modules and universal roles. *J. Stat. Mech. Theory Exp.* **2005,** P02001-1 (2005).

169. Ramaraju, S. *et al.* Removal of Interictal MEG-Derived Network Hubs Is Associated With Postoperative Seizure Freedom. *Front. Neurol.* **11,** (2020).

170. Bartolomei, F. *et al.* Defining epileptogenic networks: Contribution of SEEG and signal analysis. *Epilepsia* **58,** 1131–1147 (2017).

171. Frauscher, B. *et al.* Atlas of the normal intracranial electroencephalogram: neurophysiological awake activity in different cortical areas. *Brain* **141,** 1130–1144 (2018).

172. Sanz-Garcia, A., Rings, T. & Lehnertz, K. Impact of type of intracranial EEG sensors on link strengths of evolving functional brain networks. *Physiol. Meas.* **39,** (2018).

173. Dam, A. M. Epilepsy and Neuron Loss in the Hippocampus. *Epilepsia* **21,** 617–629 (1980).

174. Lopim, G. M. *et al.* Relationship between seizure frequency and number of neuronal and non-neuronal cells in the hippocampus throughout the life of rats

with epilepsy. *Brain Res.* **1634,** 179–186 (2016).

175. Oldham, S. *et al.* Consistency and differences between centrality measures across distinct classes of networks. *PLoS One* **14,** e0220061 (2019).

176. Baud, M.O., Kleen, J.K., Mirro, E. A. Multi-day rhythms modulate seizure risk in epilepsy. *Nat. Commun.* **9,** (2018).

177. Jobst, B. C. & Cascino, G. D. Resective epilepsy surgery for drug-resistant focal epilepsy: A review. *JAMA - J. Am. Med. Assoc.* (2015). doi:10.1001/jama.2014.17426

178. Ayala, G. F., Dichter, M., Gumnit, R. J., Matsumoto, H. & Spencer, W. A. Genesis of epileptic interictal spikes. New knowledge of cortical feedback systems suggests a neurophysiological explanation of brief paroxysms. *Brain Res.* **52,** 1–17 (1973).

179. Zijlmans, M. *et al.* High-frequency oscillations as a new biomarker in epilepsy. *Ann. Neurol.* **71,** 169–178 (2012).

180. Conrad, E. C. *et al.* Spatial distribution of interictal spikes fluctuates over time and localizes seizure onset. *Brain* **143,** 554–569 (2020).

181. Jacobs, J. *et al.* Removing high-frequency oscillations: A prospective multicenter study on seizure outcome. *Neurology* **91,** e1040–e1052 (2018).

182. Lee, C., Kim, J. S., Jeong, W. & Chung, C. K. Usefulness of interictal spike source localization in temporal lobe epilepsy: Electrocorticographic study. *Epilepsy Res.* **108,** 448–458 (2014).

183. Marsh, E. D. *et al.* Interictal EEG spikes identify the region of electrographic seizure onset in some, but not all, pediatric epilepsy patients. *Epilepsia* **51,** 592–601 (2010).

184. Alkawadri, R. *et al.* The spatial and signal characteristics of physiologic high frequency oscillations. *Epilepsia* **55,** 1986–1995 (2014).

185. Noh, B. H., Berg, A. T. & Nordli, D. R. Concordance of MRI lesions and EEG focal slowing in children with nonsyndromic epilepsy. *Epilepsia* **54,** 455–460 (2013).

186. Sinha, N. *et al.* Structural brain network abnormalities and the probability of seizure recurrence after epilepsy surgery. *Neurology* **96,** 10.1212/WNL.0000000000011315 (2020).

187. Bernabei, J. M. *et al.* Electrocorticography and stereo EEG provide distinct

measures of brain connectivity: Implications for network models. *medRxiv* 2020.12.02.20242669 (2020). doi:10.1101/2020.12.02.20242669

188. Taylor, P. N. *et al.* Normative brain mapping of interictal intracranial EEG to localise epileptogenic tissue. (2021).

189. Welch, P. D. The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms. *IEEE Trans. Audio Electroacoust.* **15,** 70–73 (1967).

190. Rosso, O. A. *et al.* Wavelet entropy: A new tool for analysis of short duration brain electrical signals. *J. Neurosci. Methods* **105,** 65–75 (2001).

191. Zaveri, H. P. *et al.* Controversies on the network theory of epilepsy: Debates held during the ICTALS 2019 conference. in *Seizure* **78,** 78–85 (W.B. Saunders Ltd, 2020).

192. Warren, C. P. *et al.* Synchrony in normal and focal epileptic brain: the seizure onset zone is functionally disconnected. *J. Neurophysiol.* **104,** 3530–9 (2010).

193. Shah, P. *et al.* Characterizing the role of the structural connectome in seizure dynamics. *Brain* **142,** 1955–1972 (2019).