Abstract

Assessing the residual cognition of behaviourally unresponsive patients with critical brain injuries continues to be a challenge. There is a need for tools that can predict patient outcomes to better inform decisions being made in the intensive care unit. In this study, we examine brain complexity through various measures that quantify the intricate patterns and dynamics of electrical signals recorded through high-density electroencephalography (EEG). We compare responses to a naturalistic auditory stimulation task with those from the scrambled versions of the same stimuli to determine if these differences can predict the survival outcomes of patients with acute brain injury. I assess 52 acutely brain-injured patients and 18 healthy controls. Results are presented in three parts. Part One showed that while complexity measures could distinguish between conditions in healthy controls, they were less successful in predicting patient outcomes. However, Part Two found a significant difference between patients with favourable and unfavourable outcomes when complexity was examined independent of Task differences. In Part Three, I found that various complexity measures can be used to predict outcomes at an individual level with machine learning. These findings suggest that EEG complexity measures have the potential as prognostic tools for individual patients, with the quality of complex brain activity being more informative for prognosis than task-based differences in acute brain-injury patients.

Keywords: complexity, EEG, naturalistic narrative stimuli, machine learning, coma, ICU

Summary for Lay Audience

My study explored whether brain activity patterns could help predict recovery in patients with severe brain injuries. I looked at brain activity (using EEG) of patients in intensive care with serious brain injuries and compared it to healthy individuals. Participants listened to a movie audio clip and a scrambled version of it while we recorded their brain activity. I used complex mathematical analysis (called "complexity measures") to analyze the brain activity patterns, with the goal of seeing if these patterns could tell us which patients might recover better.

I found that for injured patients, I could not reliably use this audio task to predict their recovery. Despite this, I did find that patients who survived had different overall brain activity patterns compared to those who did not recover.

While my specific audio task was not the best for predicting recovery, my study shows promise in using the overall complexity of brain signals to understand the chances of recovery in severe brain injury. This research is a step towards developing better tools to help doctors make informed decisions about patient care and recovery potential. More research is needed to refine these methods and make them useful in real-world medical settings. This study contributes to our understanding of brain activity in injured patients and opens doors for future improvements in predicting recovery.

Acknowledgements

I would like to express my gratitude to my supervisor, Dr. Derek Debicki, for his trust and support as I navigated my research project. I am deeply thankful for his unwavering support, guidance, and encouragement throughout this research journey. The expertise, insightful feedback, and mentorship have been instrumental in shaping this thesis and my growth as a researcher.

I would like to further extend my heartfelt appreciation to the entire Owen lab for creating a stimulating and collaborative research environment. The collective knowledge, support, and camaraderie within the lab have been invaluable in shaping my research skills. I would like to give a special thanks to Dawn Pavich, who has been instrumental in keeping both the lab and myself organized and running smoothly. I would also like to thank Dr. Bruce Morton, my advisory committee member, for his insightful feedback and guidance throughout this research endeavour.

Most of all I would like to express my deepest gratitude to my girlfriend, Nicole Ilayda Wang, who has been an unwavering source of support, encouragement, and inspiration throughout this entire process. Her patience, understanding, and constant belief in me have been invaluable in helping me navigate the challenges of this research journey. Her ability to provide perspective, offer comfort, and celebrate every small victory has been a cornerstone of my perseverance. I am profoundly grateful for her presence in my life and for her significant contribution to the completion of this thesis. Without her, this achievement would not have been possible, and I am forever indebted to her for her love and support.

Table of Contents

List of Figures

Chapter 1: Introduction

Brain Injury & The Need for New Tools to Aid Prognosis

The study of consciousness is a perplexing one because we all experience it, yet it remains very difficult to measure and define. From a clinical perspective, a practical definition of consciousness is the state of being awake and aware since both are needed to be conscious (Fernández-Espejo & Owen, 2013; Koch et al., 2016; Laureys et al., 2005). Wakefulness, a key component of consciousness, refers to the level of consciousness that describes the state of arousal or the potential to experience awareness. Awareness, on the other hand, is the content of consciousness, which encompasses our internal sense of self and our perception of the external environment. These two elements, wakefulness and awareness, form the basis of our conscious experiences (Laureys, 2005).

Even more challenging in our understanding of consciousness is trying to understand how physical neural activity can manifest into subjective phenomenal experiences. All we know for certain is that the brain is responsible for giving rise to consciousness, since damage to the brain, caused by traumatic, vascular, anoxic, infective, toxic, or other etiology (Formisano et al., 2022), can lead to Disorders of Consciousness (DoC), a serious medical condition marked by disruptions to consciousness (Giacino et al., 2014). Following an acquired brain injury, patients are often in a temporary state of unconsciousness called a coma, characterized by a lack of responsiveness, closed eyes, impaired brainstem reflexes, and absence of voluntary movements (Posner et al., 2019; Young, 2009). A coma is a type of acute DoC that often precedes chronic DoCs (Giacino et al., 2014). Damage to the ascending reticular activating system (ARAS), which is located in the pons and projects to the intralaminar nuclei of the thalamus, is one path that will lead to a coma (Morgan, 2023). On the other hand, individuals who suffer from diffuse damage to the cerebral

cortex may not fully recover consciousness, leading to chronic DoCs with partial or full disruptions to awareness (Laureys et al., 2005), such as unresponsive wakefulness syndrome (UWS) or minimally conscious state (MCS). Patients with UWS, previously known as vegetative state, experience normal circadian rhythms but can only move reflexively and are completely unaware of the environment around them (Fernández-Espejo & Owen, 2013). Patients who show low levels of awareness with some purposeful movements are in an MCS (Giacino et al., 2002; Vanhaudenhuyse et al., 2008).

A coma is a severe type of DoC, which necessitates immediate medical attention that requires direct admission to the ICU for life-sustaining therapies such as airway management and vascular support. Advances in life-saving medical technology dramatically increased survival rates for comatose patients after severe brain injury. This necessitates decisions about the cessation of life-sustaining therapies (Mayer & Kossoff, 1999). In determining whether to withdraw life support for comatose patients, medical professionals assess a range of factors, including the patient's medical condition and expected functional recovery outcome. The changes, degree, and speed of functional recovery can vary significantly, depending on the individual patient, the cause and severity of the coma, and the effectiveness of the treatment and rehabilitation (Edlow et al., 2021). In numerous instances, decisions about withholding or withdrawing life support heavily depend on the expected outcome, with a focus on minimizing patients' suffering (Turgeon et al., 2011, 2013). Additionally, the prognosis also influences the selection of early interventions and treatments and subsequent rehabilitative measures, affecting the outcome and quality of life after recovery. Because of these reasons, accurately predicting the prognosis for patients with acute brain injury is vital (Weijer et al., 2016).

Neurological measures used to assess prognosis rely on a patient's observable behaviour. Routinely, all behaviorally unresponsive patients in the ICU undergo neurological assessments. The basic neurological observation includes assessment of the patient's pupillary response, limb movement, vital signs, and level of consciousness (LOC) using the Glasgow Coma Scale (GCS). The GCS is a behavioural scale with higher scores reflecting better neurological function (Teasdale & Jennett, 1974). Coma would be indicated by a GCS score of eight. Another measure, the Glasgow Outcome Scale Extended (GOSE), is designed to measure the level of functional recovery following severe brain injury (Jennett & Bond, 1975; Jennett & Plum, 1972). This scale aims to gauge an individual's ability to perform daily activities such as self-care, mobility, communication, and social interaction. The GOSE is commonly employed in clinical trials to assess outcomes and by medical professionals to guide patient care (Wilson et al., 1998).

The reliability of current behavioural assessments is not sufficient to ensure the high level of precision needed for important decisions (Gill-Thwaites, 2006; S. M. Green, 2011; Stevens & Sutter, 2013). Patient sensorimotor dysfunctions, fluctuating states, and subjective examiner evaluations may restrict the effectiveness of behavioural assessments, especially in patients with inconsistent or unclear behavioural responses (Gill-Thwaites, 2006; S. M. Green, 2011; Stevens & Sutter, 2013). Behavioural examinations are inherently subjective and result in poor inter-rater reliability (S. M. Green, 2011; Teasdale et al., 2014). Additionally, distinguishing comas from other conditions with similar symptoms classified under cognitive motor dissociation can be challenging (Schiff, 2015).

Patients with cognitive motor dissociation have complete covert awareness but are unable to make purposeful movements (Schiff, 2015) these patients are frequently misdiagnosed and, in some cases, left for years without anyone knowing they are conscious (Laureys et al., 2005).

Cognitive motor dissociation can occur from Locked-in syndrome, caused by damage to the brain stem, or from severe Guillain-Barré syndrome (GBS), an autoimmune condition that affects the peripheral nervous system, sometimes leading to total paralysis, rendering them behaviorally unresponsive (Bauer et al., 1979). In either case, patients may lack the motor functions necessary to demonstrate 'internal awareness' (Giacino & Zasler, 1995). Conversely, patients with cognitive motor dissociation may appear behaviorally unresponsive, which leads to misdiagnosis when assessed using behavioural tests (Laureys et al., 2005). In such cases, they cannot predict good functional recovery, possibly leading to mistreatment or premature withdrawal of life support (Turgeon et al., 2011, 2013; Weijer et al., 2016; Young, 2009). This is concerning given that withdrawal from life-sustaining treatment is the most common event preceding in-hospital death (Callaway et al., 2014).

In the ever-evolving field of medicine, advancements in neuroimaging techniques hold great promise for providing new insights into the assessment of consciousness in patients suffering from acute brain injury. Recently, functional neuroimaging has been recognized as a valuable tool for identifying positive prognostic indicators of recovery for patients in the early stages of brain injury in the ICU. Functional neuroimaging can be used alongside behavioural measures to assist in neurological assessment, ultimately leading to better-informed decisions in the ICU regarding the discontinuation of life-sustaining therapy.

In the next section, we will discuss complex systems as a method for analyzing brain signals and how they can be used to broadly study pathological diseases. The following section will focus on the development of neuroimaging techniques that are specifically designed to assess the prognosis of patients with acute disorders of consciousness (DoC). Finally, we will introduce the specific approach used in this paper.

Complex Systems

A system or process is composed of various interconnected elements that can interact in multiple ways. Managing the numerous interactions between these elements can be quite complex. Complexity refers to the level of intricacy and diversity among the elements required for the system to function effectively. Higher complexity typically indicates a greater number of variables and interdependencies necessary for the system to operate successfully.

The study of complex systems focuses on systems characterized by several key features. First, the non-linear interactions among their individual components result in meaningful yet irregular and unpredictable connections that enhance information processing. Second, these numerous components can self-organize without a central authority/controller, allowing them to operate collectively through local interactions. Additionally, the non-linear interactions can lead to emergent properties, which are vital in understanding complex systems (Mitchell, 2009). Emergent properties occur when individual components of a system interact in ways that produce behaviors or characteristics not present in the components themselves (Johnson, 2006). In other words, the whole becomes more than the sum of its parts.

These characteristics can be observed in complex systems like the brain. In this context, the individual elements are neurons. The brain's interconnected structure and nonlinear interactions can lead to emergent properties such as complex behaviors, cognition, and consciousness. Despite the absence of central control, the vast network of neurons and their connections work together cohesively, enabling us to adapt, evolve, and process information.

The association between the brain’s inherent complexity and its capacity for information processing (Garrett et al., 2013) is the basis for why complexity is used for interpreting neuroimaging signals in the study of pathological and psychological conditions. The brain operates

as a nonlinear dynamic system with constantly changing interactions between neurons and neural networks that create a lot of variability in the brain signal. Signal variability can reflect a greater dynamic range of responses from perturbation (i.e., the range of possible responses to incoming stimuli). Greater dynamic range is generally beneficial to the adaptability and efficiency of neural systems because it permits a greater range of response to a greater range of stimuli (Garrett et al., 2013). These interacting neural circuits that give rise to brain signal variability are organized to maximize the brain's information capacity (Keshmiri, 2020). Given the direct relationship between variance and the amount of information the brain can process, a decrease in complexity has been proposed as a common feature of pathological conditions (Costa et al., 2002; Goldberger, Amaral, et al., 2002, 2002; Grassberger, 1991). This approach has yielded promising results in evaluating changes in consciousness, brain aging, and quantifying the information processing of brain networks (Keshmiri, 2020).

Quantifying Complexity

Complexity Calculation Methods. There are many ways to calculate complexity, and there is no single agreed-upon way in which it should be measured. Depending on the field and the purpose, different measures may prioritize different aspects of complexity. In the field of neuroscience, complexity is used to characterize and understand non-linear brain signal dynamics. Concepts such as entropy, complexity, and chaoticity have been used to explore the temporal dynamics of brain signals. These concepts are derived from information theory, dynamic systems theory, and nonlinear dynamics, which provide a broad framework for understanding complex systems.

The similarities between different measures of Complexity allow us to broadly categorize calculation methods into two main groups: measures that calculate predictability and measures that calculate regularity (Lau et al., 2022; Makowski et al., 2022). Predictability measures look at how a system changes over time or space. They assess complexity by quantifying how difficult it is to predict future states of the system. A highly complex system would be less predictable, as its future states are not easily determined from its current state. Regularity measures examine repeating patterns to determine whether the system's behavior is consistent or varied. Patterns can be observed at a specific level of detail (single-scale measures) or across multiple levels of detail and time scales (multi-scale) to capture both fine-grained and broader patterns of regularity. A system with high complexity would typically show a balance between complete randomness (no repeated patterns) and perfect order (constant repetition; Lau et al., 2022).

Thus, regularity and predictability provides insights into complexity by quantifying how information is structured and how it unfolds over time. The characteristics of complex systems—adaptability, emergence, self-organization, and non-linearity—contribute to their irregular and unpredictable nature. As a result, measuring regularity and predictability are key correlates of system complexity (Lloyd, 2001). However, the issue with this approach is that both complex systems and random systems can generate outputs that appear unpredictable and irregular, which we discuss in the next section.

Distinguishing Entropy and Complexity. Entropy measures have been widely used to quantify complex systems. However, it is important to note that entropy, which refers to randomness or chaos, does not necessarily indicate ‘true’ complexity (Goldberger, Peng, et al., 2002). A high entropy value does not always mean that a system is complex. It has been

demonstrated that traditional entropy measures such as ApEn and SampEn yield high values with random data, despite the data not possessing intrinsic physiological complexity (Costa et al., 2002, 2005). This is because traditional entropy measurements typically focus on the level of disorder or predictability in a system. On the other hand, the concept of complexity also considers the amount of information conveyed in a system. Assessing the true complexity of a system can be challenging because both complex and random systems can generate unpredictable and irregular outputs. Therefore, any measure of 'true' complexity must capture the meaningful information maintained in a complex system but lacking in a random system.

Thus, a complex system does not simply possess high entropy; the rich amount of information it contains maintains higher-level patterns, positioning it between order and disorder. Accordingly, it has been proposed that complexity first increases and then decreases as entropy rises (Figure 1; Gell-Mann, 1994; Grassberger, 1989; Huberman & Hogg, 1986). Howeve, our limited knowledge of the physical significance of various complexity measures hinders our ability to clearly distinguish between entropy/randomness and True complexity.

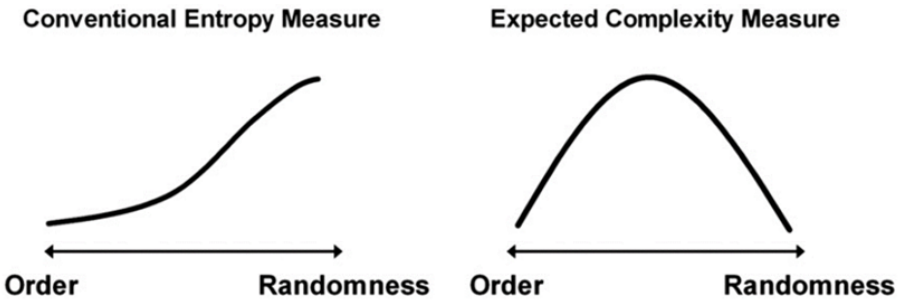


Figure 1. **The difference between conventional entropy and expected complexity measures. Adapted from Yang and Tsai (2013).**

While entropy alone does not capture the full complexity of a system, it remains a valuable tool due to its broad applicability; entropy measures can be applied across various domains from physics to biology and neuroscience (Shannon, 1948). Moreover, entropy measures are relatively simple to calculate, which makes them accessible and practical for researchers who need a straightforward method to assess complexity (Pincus, 1991). Furthermore, entropy can serve as the basis for more advanced and refined metrics of complexity. Techniques such as multiscale entropy (MSE) build on traditional entropy measures to capture more nuanced aspects of complexity across different scales (Costa et al., 2002). For example, with MSE, randomness can be identified when high entropy at short time scales decreases at larger time scales (Ma et al., 2018), since complex systems should display similar properties across different time scales (Alej et al., 2017).

Complexity Analysis to Study Pathological Conditions. Complexity measures show promise for analyzing EEG signals from patients in the ICU. This is because complexity theoretically correlates with the underlying brain activity, without being limited by modeling parameters or the templates of specific paradigms—two common drawbacks found in other analysis techniques. Also, complexity may be related to pathological conditions or outcomes after a brain injury as reduced complexity in neural dynamics is thought to be driven by factors such as diminished neural connectivity and lower neural network activation, which are common in neurodegenerative diseases (Jeong, 2002).

Additionally, complexity measures a system's capacity to adapt to ever-changing environments (Goldberger, Amaral, et al., 2002). Such adaptive capability is frequently compromised in individuals with brain injuries, leading to either overly ordered or chaotic brain

patterns. Consequently, a decrease in brain complexity can indicate a loss of consciousness. This could be a reason why many studies report higher complexity in healthy states and lower complexity in DoC (Casali et al., 2013; Casarotto et al., 2016; D. A. Engemann et al., 2018; Lei et al., 2022; Luppi et al., 2019; Sitt et al., 2014).

Development of Neuroimaging for Assessing Prognosis of Acute DoC Patients

Patients with DoC are not able to consistently and reproducibly participate in experimental tasks that require behavioural responses. Therefore, neuroimaging techniques help clinicians and researchers examine the functionality of the brain by studying the neural correlates of these functions. This information is then used to make judgments about diagnosis and prognosis.

Numerous neuroimaging methods have been developed to assess DoCs (Kondziella et al., 2016). These methods differ significantly, encompassing different neuroimaging modalities and setups, various designs and paradigms, and many analysis techniques. Neuroimaging paradigms observe brain activity at rest (resting-state studies) as well as brain activity in response to external stimuli (task-based studies). Task-based studies can further be classified into passive and active tasks.

Over the past two decades, the various methods originally developed for chronic DoC have proven to be valuable measures of awareness, residual cognition, and/or general indices of healthy brain function (Bai et al., 2021; Song et al., 2018). Consequently, adapting these methods for application in the ICU holds promise for assessing acute brain-injured patients. Specifically, these approaches have the potential to index positive prognosis in this critical population. As a result, research has shifted focus from chronic to acute cases and is incorporating both new and established techniques to satisfy this endeavour. Here, we summarize the work that has been done,

beginning with the early solutions that utilize functional magnetic resonance imaging (fMRI) to measure covert consciousness.

Command Following and Active Tasks

Functional neuroimaging serves as a valuable tool for assessing awareness in patients with DoC. This section provides background on functional neuroimaging research in DoC and how it was adapted for other imaging modalities to aid in the diagnosis and prognosis of acute DoC patients.

The early development of functional neuroimaging research in DoC began with the detection of cognitive motor dissociation using fMRI. fMRI measures changes in blood flow in response to external stimuli. In a landmark study by Owen and colleagues (2006), fMRI was used to examine a behaviorally unresponsive patient. The researchers developed a method employing two mental imagery tasks: "Imagine playing tennis" and "Imagine visiting all the rooms in your house." This approach revealed that the patient was covertly aware despite showing no outward behavioural responses. The patient could follow commands by modulating their brain activity, as evidenced by activity in the supplementary motor area during motor imagery and in the parahippocampal gyrus during spatial navigation imagery. This command-following paradigm requires patients to cooperate by intentionally modulating their brain activity in response to commands.

This paradigm was later replicated in healthy subjects, confirming that imagery tasks can reliably identify volitional brain activation at the individual level (Boly et al., 2007). A subsequent study testing 24 patients diagnosed with UWS found that 17% showed signs of covert awareness,

and even adapted the paradigm for communication through yes-no questions for one patient (Monti et al., 2010).

Overall, neuroimaging acquired during command-following tasks has improved patient diagnoses by detecting covert awareness. A review found that 14% of patients could modulate their brain activity in response to verbal commands (Kondziella et al., 2016). Consequently, the European Academy of Neurology (Kondziella et al., 2020) and the American Academy of Neurology (Giacino et al., 2018) now advocate for integrating functional neuroimaging to enhance current diagnostic procedures. Adapting the motor imagery task for use in the ICU is crucial for prognostic purposes in acute severe brain injury, as the ability to follow commands is vital for decisions regarding the cessation of life-sustaining therapy. Most research has focused on assessing brain function in chronic DoC using fMRI, partly due to the practical challenges of testing patients under intensive critical care in medically unstable and extremely vulnerable conditions. For neuroimaging paradigms to be clinically viable in the ICU, they should be moved to the bedside to reduce the physical toll on patients during fMRI testing (Cruse et al., 2011).

Portable and low-risk modalities that overcome the limitations of fMRI, such as electroencephalography (EEG) and functional near-infrared spectroscopy (fNIRS), are promising neuroimaging techniques for assessing cortical function at a patient's bedside. Research supports the use of fNIRS for providing objective measures of cortical activity in this patient population (Abdalmalak et al., 2021). However, fNIRS is relatively new, and research is still in its infancy. On the other hand, EEG has been extensively researched and is commonly used in ICU settings. Moreover, applying motor imagery tasks for diagnosis has also been shown to be viable with EEG (Cruse et al., 2011). Thus, EEG has great potential to aid in the prognosis of brain-injured patients in the ICU.

Regarding the prognosis of patients in the ICU, EEG examinations using a motor command paradigm, conducted about one week after acute brain injury, can predict clinical improvement (Claassen et al., 2019). The paradigm used in that study was similar to the tennis task, but examined the brain activity of 104 behaviorally unresponsive patients in response to instructions to move their hands. Machine learning algorithms detected that 15% of patients (n = 16) showed the expected pattern of brain activity in response to commands, and those patients had a 50% probability of regaining the ability to be independently active one year following the injury (Claassen et al., 2019). A similar study (Endlow et al., 2017) assessed the prognosis of acute DoC patients in the ICU using fMRI and EEG aquired during an active motor imagery task, previously established in chronic DoC patients (Cruse et al., 2011, 2012). In that sample, four patients responded when asked to squeeze their hand, three of whom were clinically diagnosed with UWS. However, this was not related to their 6-month outcomes.

It is possible that these tasks are too challenging for patients due to cognitive impairments that may hinder their ability to perform active tasks like motor imagery, which require sustained periods of vigilance (Monti et al., 2010; Naci & Owen, 2013). Evidence indicates that active task-based fMRI or EEG studies have a high rate of false-negative findings, which is important to consider when interpreting the results. For example, some ICU patients who show behavioural evidence of command-following do not show expected patterns of neural activity during a motor imagery task (Bodien et al., 2017; Edlow et al., 2017).

If a patient is unable to respond to commands, we need to assess their capabilities. Many cognitive faculties, such as sensory and perceptual abilities, are prerequisites to command-following. Thus, we need assessments that reflect lower and higher-level residual cognitive abilities in the absence of command-following. Perhaps resting-state or passive experimental

paradigms, rather than active ones, can identify residual cognition and provide valuable information when responses to active tasks are not evident (Boly & Seth, 2012).

Residual Cognition

Resting-State Paradigm for Prognosis of Acute DoC Patients. Resting-state connectivity offers a valuable approach for prognostic assessment, particularly in evaluating functional connectivity in the brain (Soddu et al., 2011). This method is especially useful for studying patients with DoC who cannot communicate behaviorally, as it doesn't require active task participation or stimulus response. Resting-state connectivity analyzes the brain's spontaneous low-frequency fluctuations from synchronized activity across functionally related regions, enabling the examination of large-scale cerebral networks (Damoiseaux et al., 2006). This technique might provide additional information about preserved neural networks that could support cognitive and functional recovery.

Researchers have extensively studied resting-state fMRI for its potential to aid in prognosis for acute brain-injured patients (Kazazian et al., 2020; Koenig et al., 2014; Norton et al., 2012; Pugin et al., 2020; Sair et al., 2018; Silva et al., 2015). Norton et al. (2012) found a link between intact default mode network (DMN) connectivity and positive outcomes in DoC patients. Koenig et al. (2014) discovered that higher neural connectivity within the DMN correlated with better functional outcomes in patients with an indeterminate prognosis after cardiac arrest. In a case study, Kazazian et al. (2020) scanned an acute DoC patient who recovered, using fMRI in the ICU and again 9 months later. Their findings showed increased functional connectivity across multiple resting-state networks at recovery.

While resting-state paradigms provide valuable insights, relying solely on them can be limiting and is generally recommended only in specific circumstances (Finn, 2021). Researchers should carefully consider their assumptions when selecting acquisition states, even for coma patients with limited capacity for task-based activities.

Passive Tasks for Prognosis of Acute DoC Patients. Employing passive stimuli may reveal time-locked brain activity in unresponsive individuals. This approach could potentially offer deeper insights into residual cognition compared to resting-state analysis alone. Therefore, various studies have used passive tasks to assess DoC patients. For example, Gofton et al. (2009) reported that greater somatosensory cortex activation correlated with the recovery of consciousness. Moritz et al. (2001) found intact sensory processing in a comatose patient, which predicted cognitive recovery. Cruse et al. (2014) also revealed positive prognostication from median-nerve somatosensory evoked cortical potentials. Furthermore, Formisano et al. (2019) noted that in the ICU, neuroimaging exams with EEG neurophysiological markers such as sensory evoked potentials and event-related potentials were more frequently chosen as prognostic indices in the acute phase.

Early work using EEG has shown some promise in predicting outcomes in the ICU. In particular, the presence of the mismatch negativity event-related brain potential (ERP) has demonstrated a positive correlation with awakening from a state of coma (Morlet & Fischer, 2014). While the presence vs. absence of this brain response has high specificity (>90%), it has low sensitivity (<30%) and, thus may not be useful for clinical practice in its current form. Other long-latency event-related brain potentials such as the P300 and N400 have also been tested for prognostic purposes but have shown poor sensitivity and specificity (Wang & Ding, 2011). At

present, the clinical utility of EEG in the context of prognosis remains limited and must be more fully developed.

Passive Tasks for Auditory and Language Processing for Prognosis of Acute DoC Patients. Coleman et al. (2007) developed an auditory perception paradigm to determine if coma patients can process auditory stimuli. Specifically, they investigated whether patients maintain sound perception, speech perception, and language comprehension by presenting short and simple sentences, between 1.2 and 4.3 seconds in duration. Differentiating lower and higher-level cognitive processing can inform the amount of residual cognition available for a given patient and reflect their potential capacity for awareness. Sound perception can occur without awareness, as lower-level perceptual responses are known to occur under sedation (Dueck et al., 2005; Eagleman & MacIver, 2018; Kerssens et al., 2005). However, higher-level semantic processes that support language comprehension are abolished under sedation along with decreased levels of awareness (Davis et al., 2007; Heinke et al., 2004; Plourde et al., 2006). Therefore, the level of auditory capabilities is related to awareness. Indeed, Coleman et al. (2007) were able to distinguish between the various levels of auditory processing and how they relate to awareness and DoC.

Coleman et al. (2009) later adapted the passive auditory paradigm to assess comatose patients and found that those who recovered consciousness had greater activation in the high-order language processing region than patients who had a poor outcome. Thus, having this area intact is predictive of positive outcomes, demonstrating the prognostic utility of passive auditory stimuli. Importantly, this is applicable for acute DoC patients in the ICU, as auditory processing is positively associated with recovery (Norton, 2017).

Utilizing EEG, Sokoliuk et al. (2021) evaluated 28 patients who were clinically unresponsive following acute traumatic brain injury (TBI). The study involved presenting sequences of monosyllabic words that formed coherent phrases and sentences. The findings indicated that patients who exhibited a stronger response to speech comprehension, as measured by inter-trial phase coherence, had more favourable outcomes on the GOSE six months post-injury. Additionally, their linear regression analysis demonstrated that the intensity of the comprehension response significantly improved the predictive accuracy of outcomes, surpassing predictions made using only clinical measures such as the GCS and anatomical neuroimaging.

On the other hand, not all research has established a link between functional neuroimaging findings and recovery outcomes. Edlow et al. (2017), used fMRI and EEG with passive music and language paradigms to detect activity in higher-order auditory regions in acute DoC patients. They found low sensitivity to predicting positive six-month outcomes using their methods. However, their study used simple language stimuli that may have precluded the assessment of more complex and continuous information-processing capabilities. Simple language stimuli can demonstrate the preservation of thalamocortical circuits and cortical responses to spoken language, which may provide prognostic insights. However, complex language stimuli that activate neural networks that mediate attention, emotion, auditory processing, and plot following, may be more effective for indexing prognosis. This perspective brings attention to the potential benefits of using naturalistic narrative stimuli, which we discuss in the following section.

A New Direction

Naturalistic Movie Narrative Stimuli. One of the limitations of implementing neuroimaging in the ICU is developing a task that is easy for a patient to engage with while

remaining informative about the residual cognitive capabilities of the patient. Many of the neuroimaging tasks previously discussed are very simple and can only tell us about processing simple stimuli or, at most, language abilities (which do not engage emotions or a plot-following ability). On the other hand, some tasks are too difficult for patients due to cognitive impairments that may impede their ability to perform active tasks like the tennis task.

To better assess patients in the ICU with acute brain injury, we can record their brain activity while they receive passive stimuli. An alternative neuroimaging paradigm was developed which examined the neural activity in more natural conditions using engaging, real-world stimuli like movies (Hasson et al., 2004, 2008, 2010). These naturalistic paradigms are better suited for behaviorally unresponsive brain-injured patients because they have fewer constraints and are less demanding for the patient as they do not require formal responses (Naci et al., 2014, 2017). Moreover, employing a naturalistic narrative paradigm could be an effective method for measuring higher and lower-order residual cognitive function, as the movie can engage various levels of brain function, from basic sensory (such as audio/visual stimuli) to higher-order cognitive functions like following the plot.

Techniques for analyzing naturalistic narrative stimuli are an important topic of discussion (Bartels & Zeki, 2004, 2005; Hasson et al., 2004, 2010; Naci et al., 2017; Spiers & Maguire, 2007), to develop new ways of interpreting brain signals in naturalistic settings. Traditional neuroimaging techniques used to delineate the meaning of brain signals are not suited for naturalistic paradigms. The challenge with analyzing EEG signals in a naturalistic stimulation paradigm is that it is difficult to model predictors without specific set events. EEG data analysis requires a structured experimental paradigm in that there needs to be predefined stimulation times with a specific duration set (Friston et al., 1994). Set events allow for an average over these events to calculate

the event-related potential, a waveform that models brain activity in response to a repeating stimulus. However, this is not feasible in a naturalistic task with continuous stimuli. Moreover, traditional methods of EEG analysis require a deep understanding of neurophysiology and are prone to error (Donoghue et al., 2020). Recordings of complex brain signals can be challenging to model and interpret because brains are very complicated in that they are dynamic systems with many linear and non-linear factors that directly or indirectly affect how the system operates.

Inter-subject synchrony. An alternative approach to directly modelling brain activity is based on the principle that participants who listen to engaging, story-driven narratives are likely to exhibit similar patterns of brain activity, as the dynamic narratives engage them in similar ways. Therefore, analysis techniques use inter-subject synchrony to quanitfy engagement and the ability to follow the plot with the assumption that as participants experience the events of the same story, their brain activity should synchronize as the events unfold over the same time span (Hasson et al., 2004, 2010; Naci et al., 2017). One study presented a brief suspenseful movie to a behaviorally unresponsive patient and successfully showed that the patient's brain activity was similar to that of healthy individuals, which suggests that the patient is covertly conscious (Naci et al., 2014). Next naturalistic audio-only stimuli, from the movie ‘Taken’ could produce comparable results (Naci et al., 2017). Auditory-only movie stimuli can be particularly effective in assessing cognitive capabilities in acute DoC patients, since even within an unconscious state, auditory perceptual abilities are preserved (Dueck et al., 2005; Eagleman & MacIver, 2018; Kerssens et al., 2005), making a neural assessment based on auditory processing more suitable for patients with eyes closed. Thus the Taken paradigm was adapted for use with EEG for bedside assessment of acute DoC patients (Laforge, 2017, 2022). Moreover, brain activity similar to healthy controls while

listening to auditory narratives has been proposed as a promising method for assessing a patient's residual cognitive capacity and predicting early recovery of consciousness (Laforge, 2022).

Tasks that can detect consciousness and higher-level brain function in DoC patients can forecast significant recovery of neurological function (Coleman et al., 2009; Edlow et al., 2017; Giacino & Kalmar, 1997; Stender et al., 2014; Whyte et al., 2001). The inter-subject neural synchrony is also extended to detect residual cognition in patients who remain behaviorally non-responsive (Laforge, 2022; Laforge et al., 2020; Naci et al. 2014, 2017). It follows that the Taken task may support prognosis and prevent premature cessation of care in patients who still have enough cortical function to support recovery.

Rationale for Using EEG Complexity for Taken Analysis. Techniques like inter-subject synchrony are useful for analyzing patterns of neural activity elicited in response to naturalistic stimuli. However, our ability to assess an individual's neural integrity is restricted to the degree to which it matches a template set by other participants who listen to the task. Thus, the absence of synchrony between a patient's brain activity and that collectied from a normative sample cannot give us any more information about the patient. We need a measure that can inform higher-level cognitive abilities by contrasting intact and scrambled audio clips, while also telling us something about the inherent structure of the brain signals.

A promising alternative is to use measures of complexity for interpreting brain signals during naturalistic narrative stimuli. Complexity analysis techniques have gained attention in EEG data analysis (Sarasso et al., 2021). Unlike ERP analysis and inter-subject synchrony, complexity measures do not require the modelling of complicated brain signals, nor do they rely on the similarity of activation from other participants. Instead, they quantify how complex the system is,

and researchers can determine if this complexity is related to a person's cognitive state or consciousness level. Measures of complexity can be applied to functional neuroimaging data to quantify how complex the signal is.

The examination of neuroimaging data is naturally complex, involving many voxels/electrodes/sensors and time points, at each of which a brain signal can be measured. The complexity of these data forces allows researchers to collapse them in some meaningful way. The concept of quantifying the "complexity" of physiological signals has potentially important implications in evaluating both dynamical models of biological systems and bedside diagnostics (Costa et al., 2002; Goldberger, Amaral et al., 2002; Goldberger, Peng et al., 2002; Grassberger, 1991).

The following section begins with relevant background on using complexity for diagnosis and prognosis of DoC. Then, we discuss the potential of using complexity to analyze naturalistic narrative stimuli for the prognosis of acute brain-injured patients.

Using Complexity in DoC

DoC - State-Based Differences in Complexity

Extensive research has demonstrated the effectiveness of complexity measures in analyzing EEG data to understand the various states of consciousness. These studies have been conducted in a variety of contexts, including sleep patterns (Tosun et al., 2017), psychedelic experiences (M. M. Schartner et al., 2017), and the differential diagnosis of DoC (Casali et al., 2013).

Li et al. reported that MCS patients showed higher entropy and EEG complexity than UWS/VS patients (2014). Moreover, Lei et al. used approximate entropy (ApEn), sample entropy

(SampEn), and Lempel-Ziv complexity (LZc) to investigate the dynamic changes of EEG signals in DoC patients during consciousness recovery (Lei et al., 2022). Sitt et al. (2014) found that the complexity of cortical activity indexes consciousness and that EEG frequency power and complexity were the most reliable parameters to differentiate between UWS, MCS, or the conscious state by conducting a large study involving EEG recordings of 113 patients (Sitt et al., 2014). Furthermore, Rosanova et al. (2018) reported complexity recovery using transcranial magnetic stimulation (TMS)-EEG measures in patients recovering from UWS/VS to MCS and beyond MCS.

Research using TMS-EEG, in particular, is very promising. Combining TMS-EEG allows us to monitor cortico-cortical interactions (Massimini et al., 2009). The cortical responses to TMS calculated with complexity have been found to be significantly different between conscious and non-conscious states (e.g. healthy conditions, anesthetized subjects, and patients with DoCs) (Ferrarelli et al., 2010; Massimini et al., 2005, 2010; Rosanova et al., 2012; Sarasso et al., 2015). Thus, the cortical response to TMS has been interpreted as an "index" of the level of consciousness, a value called the perturbation complexity index (PCI; Casali et al., 2013). The clinical validation of this threshold revealed its ability to discriminate between different DoCs (namely UWS and MCS), showing higher sensitivity and specificity than behavioural diagnostic tools (Casarotto et al., 2016).

While the application of PCI is great for chronic DoC patients, there is a lack of research using this method in acute DoC patients. Perhaps this is due to the limitation of using TMS because it is invasive and less portable making it less practical for use in the ICU. These limits (invasive, portable) are especially detrimental for bedside testing, which is important for acute patients that

we want to test. Could a different type of perturbation be used that was non-invasive and easy to control?

Using Complexity and Taken Stimuli

Perhaps task-based stimuli like the Taken paradigm can be used to perturb brain activity. The rationale for choosing naturalistic auditory stimuli was based on the premise that such stimuli would require the preservation of higher-order cognitive functions, such as language processing and plot following (Naci et al., 2017). The idea is that these stimuli would perturb brain regions associated with more complex cognitive functions, which could be captured through complexity measures. Capturing such a change was inspired by previous studies that demonstrated the sensitivity of complexity measures to subtle changes in cognitive states and task engagement. For example, studies have shown that complexity varies depending on the task we are engaged in, such as differentiating between internally and externally generated perceptions (Ibáñez-Molina & Iglesias-Parro, 2014), attentional states (Bornas et al., 2013), hypnosis (Solhjoo et al., 2005), language processing (Omidvarnia et al., 2022), motor imagery (Stam et al., 2002), working memory (Sauseng et al., 2010), and studies investigating vigilance during mental tasks (L.-C. Shi et al., 2013). Most of these studies utilize a resting state as the baseline measure for measuring change in complexity induced by a task. For example, some studies have shown that changes in complexity from a resting state can be induced just by changing the way people think using a creative thinking task (Mölle et al., 1999) or with various types of meditation methods (Aftanas & Golocheikine, 2002).

Moreover, research has shown that the difference in complexity between task and rest conditions can vary depending on factors such as age (Takahashi et al., 2009) and attention deficit

hyperactivity disorder (Gu et al., 2022). These findings suggested that task-based changes in complexity could potentially differentiate between patient populations and inform us about underlying deficits.

A previous study investigated the differentiation of brain activity patterns with complexity, in response to different stimuli, such as a movie, a scrambled movie, or 'TV noise' (Boly et al., 2015). The paper reported that the differentiation of complexity was correlated with the meaningfulness of the stimulus set, being highest in the movie condition, intermediate in the scrambled movie condition, and minimal for 'TV noise.' This demonstrated that movie stimuli are also an effective task for examining complexity change.

These findings provide a strong rationale for using complexity differences between scrambled and intact movie stimuli in our study. By comparing the complexity of EEG signals during intact and scrambled movie conditions, we can potentially capture the brain's ability to engage with and process meaningful narrative content. Therefore, observing task-based changes in complexity can help identify underlying deficits and differentiate between patient populations. This principle can be applied to studying acute brain-injured patients in the ICU.

Choosing Measures

There are many measures to represent complexity mathematically (Shalizi, 2006). Moreover, it has been suggested that combinations of features synergistically outperform individual measures (Sitt et al., 2014). Therefore, for this study I adopted a comprehensive approach, assessing ten widely used complexity measures to gain a broader understanding of their effectiveness in predicting prognosis for brain-injured patients. These measures include Singular Value Decomposition (SVD), Approximation (Approx), Sample Entropy (Sample), Detrended

Fluctuation Analysis (Detrended), Katz Fractal Dimension (Katz), Line Length Index (LL), Multiscale Wavelet Permutation Entropy (MSWPEn), Bubble Entropy (bubbEn), Condition-Weighted Permutation Entropy (CWPEn), and Lempel-Ziv Complexity (LZc). Table 1 provides a detailed description of each measure.

Table 1. Description of each complexity measure used in this study.

Measure	Calculation Method	Category	References
Line Length Index (LL)	LL calculates the sum of absolute differences between consecutive data points in a time series, normalized by the series length. It provides a measure of signal variability and complexity, originally developed for efficient seizure detection.	Predictability	(Esteller et al., 2001)
Detrended Fluctuation Analysis (Detrended)	Detrended quantifies long-range correlations in non-stationary time series. It computes the average fluctuation of the integrated and detrended time series at different time scales, with the scaling exponent indicating the nature of correlations.	Predictability	(Hardstone et al., 2012; Peng et al., 1994)
Singular Value Decomposition (SVD)	SVD Entropy measures the entropy of normalized singular values obtained from the SVD of a time series matrix. It quantifies the number of significant singular values, reflecting the signal's dimensionality and complexity.	Predictability (can be considered multi-scale in the spatial domain)	(Roberts et al., 1999)
Approximation (Approx)	Approx Entropy quantifies the predictability of subsequent amplitude values in a time series based on knowledge of the previous values. It measures the logarithmic likelihood that similar patterns of observations will remain similar in the next incremental comparisons.	Regularity	(Pincus, 1991)
Sample Entropy (Sample)	Sample Entropy improves upon ApEn by excluding self-matches and being less dependent on time series length. It calculates the negative natural logarithm of the conditional probability that two sequences similar for m points remain similar at the next point.	Regularity	(Richman & Moorman, 2000)

Bubble Entropy (bubbEn)	BubbEn applies the bubble sort algorithm to ordinal patterns in a time series. It counts the number of swaps needed to sort each pattern, providing a measure of the signal's complexity that is robust to outliers.	Predictability	(Manis et al., 2017)
Katz Fractal Dimension (Katz)	Katz estimates the fractal dimension of a time series by comparing the total length of the signal to the maximum distance between the first point and any other point. It quantifies the complexity of the signal's path through space.	Predictability (can be considered multi-scale in the spatial domain)	(Katz, 1988)
Condition-Weighted Permutation Entropy (CWPEn)	CWPEn enhances traditional permutation entropy by incorporating amplitude information and considering the relationships between successive ordinal patterns. It calculates the difference in weighted entropy between embedding dimensions m and m+1.	Multi-scale Regularity	(Unakafov & Keller, 2014)
Lempel-Ziv Complexity (LZc)	LZc measures the number of distinct substrings and their rate of occurrence along a time series. It quantifies the gradual buildup of new patterns in the sequence, with higher values indicating more complexity or randomness.	Regularity (Temporal & Spatial)	(Lempel & Ziv, 1976)
Multiscale Wavelet Permutation Entropy (MSWPEn)	MSWPEn applies permutation entropy to multiple scales of a signal obtained through wavelet decomposition. It captures the complexity of a time series across different frequency bands and time scales.	Multi-scale Regularity	(Dávalos et al., 2019)

These measures were selected based on criteria such as their previous success in the literature (Lau et al., 2022; Makowski et al., 2022), and a desire to include a broad representation of approaches and analytical techniques. The aim was to ensure that the selected measures would not only provide comprehensive coverage of the various dimensions of complexity but also be relevant to the field of brain injury prognosis and research on consciousness.

The first consideration in selecting these measures was the inclusion of different computational approaches to quantifying complexity, to ensure broad methodological diversity. That is why out of the ten measures, five are regularity measures and five are predictability

measures. Among these, eight are temporal measures (LL, Detrended, Approx, Sample, bubbEn, CWPEn, LZc, MSWPEn) and two are spatial (SVD, Katz). Temporal measures better complement the characteristics of the EEG signal, as EEG has high temporal resolution but poor spatial resolution. In terms of scale, six measures are single-scale (LL, Approx, Sample, bubbEn, LZc) while four are multi-scale (CWPEn, MSWPEn, SVD, and Katz), offering a more complex depiction of brain signal variability across different time scales. By including these diverse categories, I aimed to ensure a well-rounded assessment of complexity that could capture the multifaceted nature of brain signals in the context of prognosis.

A second guiding factor in selecting these measures was their established success in the neuroscience literature, particularly in EEG-based studies assessing states of consciousness, brain injury, and related areas. For example, Lempel-Ziv Complexity (LZc) has been a robust measure in studies of brain activity and consciousness states, notably used by Schartner et al. (2015, 2017) in sedation studies and Casali et al. (2013) to assess consciousness levels. These works have demonstrated that LZc can reflect brain signal variability and complexity in altered consciousness states, suggesting its relevance for assessing prognosis in brain-injured patients. Similarly, measures like Detrended Fluctuation Analysis (DFA) and Katz Fractal Dimension (Katz), which analyze long-range correlations and spatial complexity, respectively, have also shown promise in capturing meaningful characteristics of EEG data in these contexts (Katz, 1988; Peng et al., 1994). Including these well-documented measures allows for comparisons with prior studies and enhances the clinical relevance of the results.

Finally, I drew on recent data-driven research by Lau et al. (2022), which explored various complexity measures to understand which metrics accounted for the most variance in EEG brain signal data. Based on their findings, I selected four measures—Line Length Index (LL), MSWPEn,

CWPEn, and Bubble Entropy (BubbEn)—for their significant explanatory power (Lau et al., 2022). Additionally, adding other widely used measures like Approximate Entropy and Sample Entropy enabled the inclusion of simpler yet proven metrics that have long been validated in brain activity research (Pincus, 1991; Richman & Moorman, 2000).

Research Question

A non-invasive perturbation method with EEG can be easily implemented in the ICU to assess brain complexity and potentially improve diagnosis and prognosis for acute brain-injured patients. Therefore, I propose using a naturalistic auditory stimulus - a clip from the movie "Taken" - as a non-invasive means of perturbing brain activity.

Study Aim. My primary aim is to investigate whether differences in EEG complexity between auditory naturalistic stimulation and its scrambled version can serve as early brain-based markers of recovery from serious brain injury.

Experimental Approach & Hypotheses. This study was completed in three parts. I employed two different approaches (Part I and II) to explore our research question, followed by a third exploratory analysis (Part III). My first approach (Part I) was to have a comprehensive first look at several complexity measures for their utility in providing prognostic information. Specifically, I employed a data-driven machine learning approach that utilized intra-subject level binary classification of tasks from each participant's data to see how well models performed with data from patients with favourable outcomes vs patients with unfavourable outcomes. To this end, I used healthy control data to first identify the top-performing complexity measures and machine learning algorithms for intra-subject level classification between intact and scrambled audio clips.

Then, I applied these top models and measures to patient data and obtained accuracy scores of for each measure and each patient. Accuracy scores provide a metric of how well the machine learning models perform at discrimination between Scrambled and Intact. Finally, I split patients based on their outcome (favourable vs unfavourable) to compare their accuracy scores at a group level. I hypothesized that there would be a difference in accuracy scores between patients with favourable and unfavourable outcomes. Specifically, I expected that participants with favourable outcomes would have greater accuracy scores due to the observed change in their complexity scores between intact and scrambled clips. This may be due to the change in brain dynamics perturbed by the higher-order cognitive process introduced in the Intact condition. The degree to which a patient's brain activity reflects this change is operationalized as patient accuracy scores for correct classification of Intact vs Scrambled.

In my second approach (Part II), I used traditional statistical methods to explore group-level differences rather than individual machine-learning models for each patient, as in Part I. While my first approach was highly comprehensive, this second approach allowed for a more nuanced exploration of the within-subject task-based differences and between-subject outcome variables separately, as well as their interaction. In this approach, patients were analyzed in a two-way mixed-factorial design with ‘Outcome’ as the between-subjects factor and ‘Task’ as the within-subjects factor that represents the different conditions each subject undergoes. I hypothesize a main effect of Task, a main effect of Outcome, and an interaction effect such that the difference will be observed between Scrambled vs Intact only for patients with favourable outcomes.

Lastly, I also employed an additional third exploratory approach (Part III) based on the findings from my first two approaches, in which I speculated that absolute values of complexity may be more informative for predicting outcomes than task-based changes in complexity.

Therefore, this approach focused mainly on patient prognosis using absolute complexity values rather than differences between tasks. Here, I trained machine learning models for the binary classification of outcomes to assess their accuracy for predicting prognosis at an individual level. I hypothesized that the machine learning algorithms will be able to predict outcomes at an individual level.

Chapter 2: Methods

Patient Selection, Recruitment, and Study Sites

Participants

52 unresponsive brain-injured patients were recruited from the London Health Sciences Centre (LHSC; London, Ontario, Canada). Patient demographics and data can be found in Table 2. All patients were admitted to the ICU suffering from a severe brain injury, rendering them unresponsive, and were receiving life-sustaining therapies. Written informed consent was obtained from the substitute decision-makers. Ethical approval for the research study was obtained from the Health Sciences Research Ethics Board of Western University.

18 right-handed healthy participants also participated in the study (11 female, 7 male; 22-40 years old). Controls were recruited from our group as part of a past study (Laforge, 2017). All volunteers had no known neurological or psychiatric disease and provided their written informed consent.

Inclusion and Exclusion Criteria

Inclusion Criteria for the patients included:

- Having suffered a brain injury that has rendered the patient unresponsive (e.g. traumatic brain injury, anoxic brain injury, stroke, subarachnoid hemorrhage)
- Being acutely ill with required hospitalization in the ICU
- Between 4 and 75 years of age
- Normal cognition prior to ICU admission
- Healthy Controls: No prior history of neurological disease or cognitive impairment

Exclusion criteria included:

- Anyone who is deemed medically unsuitable for this study as determined by a physician involved with the study
- Presence of status epilepticus
- Head injury that would impede electrode/probe placement on the scalp leading to poor data quality

Patient Clinical Health Information

When available, various pieces of clinical patient information were collected during this study including:

1) patient demographics: age, sex, level of education, ethnicity, languages spoken, handedness, pre-morbid medical history (including hearing, neurocognitive and psychiatric disorders)
2) Acute critical illness characteristics: mechanism of injury, GCS score, results from general neurologic examination, sedation/analgesia drug dose equivalents if any, physiological parameters such as heart rate and CO_2 levels, delirium duration, ICU stay duration, and hospital stay duration. Note that the GCS assesses one's level of consciousness, which ranges from three ("deep unconsciousness") to fifteen ("fully alert and oriented"). The GCS is a behavioural assessment that gives three separate scores to each category: motor, verbal, and eye-opening responses (Teasdale & Jennett, 1974). Our patient population was in a coma, which is commonly defined as a score of eight or lower. A GCS score of thirteen or higher suggests a mild brain injury to completely healthy, a score of nine to twelve suggests a moderate brain injury and a score of eight or less suggests a severe brain injury. The patient's

best responses in each category (verbal, motor, and eye-opening responses) were added together to give a total score (refer to Table 2).

3) Functional outcomes were based on the GOSE. In a 3, 6, and 12-month follow-up assessment, patients were assessed using GOSE to determine the level of recovery (1 = death; 2 = persistent vegetative state; 3 = severe disability; 4 = moderate disability; 5 = good recovery). Final outcome scores were binarized as Favourable or Unfavourable. Favourable outcomes were defined as a GOSE score of >1. The patients that 'survived' but were not followed up at the regular time points (3, 6, and 12 months) were also labelled Favourable if they could follow commands in the weeks following their EEG recording (see Timeline) and prior to their discharge from the hospital.

Table 2 Patient demographic information.

Patient ID	Age	Sex	GCS[#]	Best GOSE	Outcome	Sedation
01_1	63	M	3T	-	Favourable	Propofol + Hydromorphone
01_2	63	M	3T	-	Favourable	Propofol + Hydromorphone
02	86	M	3T	5	Favourable	Propofol + Hydromorphone
03_01	62	F	3T	5	Favourable	-
03_02	62	F	15	8	Favourable	-
04	26	M	3T	-	Favourable	Propofol + Hydromorphone
05	69	M	5T	-	Unfavourable	-
06	60	F	3T	-	Unfavourable	-
07_1	63	F	5T	-	Unfavourable	-
07_2	63	F	4T	-	Unfavourable	-
08	25	M	4T	-	Unfavourable	Hydromorphone
09	67	M	3T	-	Unfavourable	-
10	78	M	6T	-	Unfavourable	Propofol + Hydromorphone
11	60	F	3T	-	Unfavourable	-
12	67	M	7T	-	Unfavourable	-
13	67	M	8T	-	Favourable	-
14	75	M	9T	-	Unfavourable	-
15	67	M	4T	7	Favourable	-
16	63	F	7T	7	Favourable	Propofol
17	61	M	6T	14	Favourable	-

18	54	F	9T	-	Favourable	-
19	59	F	5T	-	Favourable	-
20	19	M	3T	-	Unfavourable	Propofol + Hydromorphone
21	58	M	5T	-	Unfavourable	-
22	66	F	6T	-	Favourable	-
23	59	M	9T	-	Unfavourable	-
24	82	F	6T	-	Favourable	-
25	48	F	6T	-	Favourable	-
26	68	F	3T	7	Favourable	Hydromorphone
27	40	F	10T	1	Unfavourable	Propofol
28	61	F	6T	1	Unfavourable	Propofol
29	58	F	3T	-	Favourable	Propofol
30	55	M	3T	-	Favourable	Hydromorphone
31	45	F	6T	1	Unfavourable	Hydromorphone
32	56	M	9T	3	Favourable	Hydromorphone
33	51	M	4T	7	Favourable	Propofol
34	55	M	4T	1	Unfavourable	Propofol
35	30	M	3T	-	Unfavourable	Propofol
36	64	M	10T	8	Favourable	Propofol
37	60	M	11T	6	Favourable	Propofol
38	54	F	7T	3	Favourable	Propofol
39	66	M	11T	5	Favourable	Propofol
40	80	M	3T	1	Unfavourable	Propofol
41	37	M	3T	1	Unfavourable	Propofol
42	68	F	8T	1	Unfavourable	Propofol
43	68	M	3T	-	Favourable	Propofol
44	76	M	3T	1	Unfavourable	Propofol
45	63	M	11T	-	Favourable	Propofol
46	34	F	6T	1	Unfavourable	Propofol
47	55	M	10T	-	Favourable	Propofol
48	55	F	5T	-	Unfavourable	Propofol
49	58	M	-	-	Unfavourable	Isoflurane
50	62	M	9T	-	Unfavourable	Propofol
51	47	M	3T	-	Unfavourable	Propofol
52	63	M	8T	-	Favourable	Propofol

[#]In the GCS column, the 'T' means the patient was intubated during the behavioural assessment.

GBS Patient:

One of the patients (03_01) was followed 12 months later for a second scan (03_02). She was admitted to the ICU due to the rapid onset of sensory changes and weakness caused by Guillain-Barré Syndrome (GBS). The patient's condition deteriorated quickly, leading to loss of brainstem and motor responses, requiring intubation. Further tests revealed significant nerve

damage and paralysis. Despite being fully capable of consciousness, the patient's motor and sensory issues resulted in a very low Glasgow Coma Scale (GCS) score of 3. However, the patient showed signs of improvement over time and was eventually transferred to a rehabilitation facility for further recovery. On day 22, she could respond and communicate behaviorally. On day 73, she was moved to an in-patient rehabilitation facility, where her recovery continued.

Study Procedures

Timeline

This study is part of a larger project in which protocol details can be found at Kazazian et al. (2021). Briefly, pre-screening and documenting eligibility occurred daily in the ICU to recruit eligible participants. A research team member presented the study to the substitute decision-maker (SDM) so that they could sign the informed consent document. Study procedures occurred on days 2–6 after ICU admission. Patients underwent Neurological Examination and Clinical Rating Scales directly before recording brain activity. EEG testing occurred at some point between days 2-6 of ICU admission or when medically stable. On some occasions, EEG was also repeated between days 7–10.

Functional outcomes were also assessed through phone or in-person interviews at 3 months, 6 months, and 12 months post-injury. Follow-up EEG testing was done for one patient (Patient 03) approximately 12 months post-injury to assess for changes in brain responses from acute injury to recovery. This was done for only one patient who presented with GBS in their first scan. The EEG procedure and paradigm at the follow-up scan were the same as those at the acute injury scan.

EEG Procedures

Patients were fitted with an electrode cap, known as the Electrical Geodesics Inc. (EGI) EEG system, which is a high-density EEG containing 128 electrodes (Figure 2). The electrode impedance threshold was set to below 50 Kohms. Testing occurred at the patient's bedside. After the cap was fitted, the resting state was recorded for 5 minutes, followed by 10 minutes of the auditory stimuli paradigm. Audio was presented over Etymotics ER-1 in-ear headphones at a comfortable volume level. The total time of the testing was approximately 15 minutes.

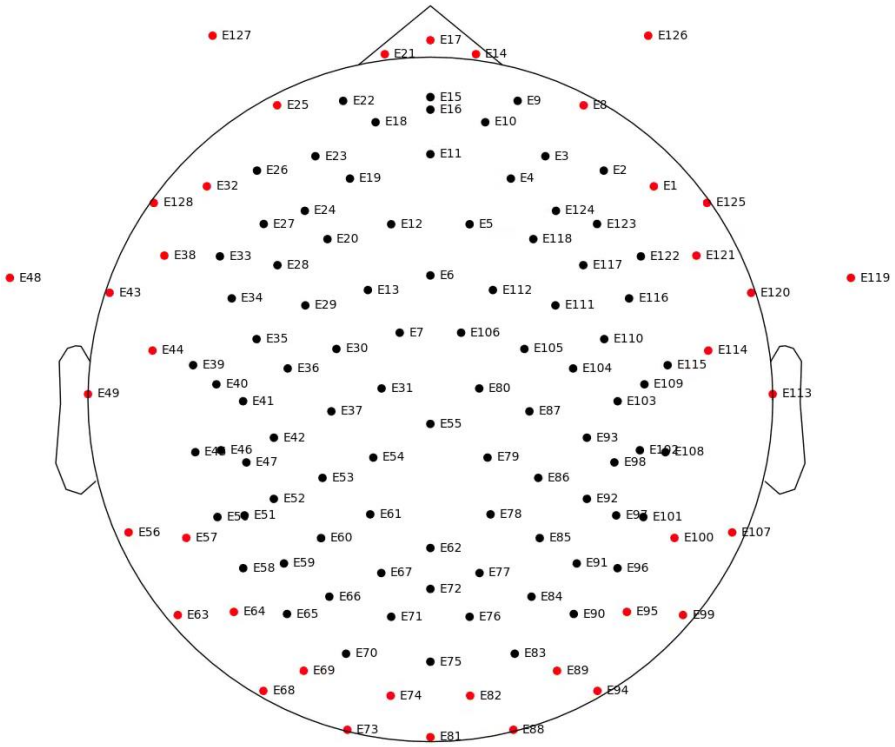


Figure 2. EEG montage. 128-channel Hydrocel Geodesic Sensor Net. Red dots indicate non-brain electrodes.

Stimulus Paradigms

A resting state scan, absent of any patient stimulation, was acquired and took approximately 5 minutes. Additionally, an audio clip from the movie “Taken” was presented to participants in a manner similar to that outlined in (Naci et al., 2014).

In the 5-minute audio clip from the movie "Taken," the listener hears a phone conversation between a father and his daughter, who is away on vacation. The conversation quickly changes tone as she becomes aware of the kidnappers in her hotel suite. The kidnappers eventually discover where she is hiding and take her away, all of which can be heard over the father's end of the call. The movie clip was selected for its immersive elements and intense nature to better engage the audience and was highly acclaimed for its suspenseful nature (Naci et al., 2017). Additionally, the scene unfolds during a cellphone conversation, making it well-suited as an auditory-only stimulus because no contextual information is lost when the video is removed.

In addition to the “Intact” audio clip participants also listened to a “Scrambled” version of the same length. I used a scrambled control stimulus to separate the complexity of neural responses elicited by the sensory properties of listening to the movie from those involved in following the plot. In order to produce the scrambled version of "Taken," the movie clip is distorted by spectrally rotating the audio frequencies (Naci et al., 2017). This retained most of its sound characteristics but made the speech impossible to understand (T. Green et al., 2013).

To prevent carry over-effects of the narrative from any prior knowledge of the storyline influencing the scrambled versions, the intact stimuli always appeared after their scrambled version. The movie paradigm lasted approximately 10 minutes.

Analysis

Preprocessing

The analysis workflow is summarized in Figure 3. First, the data were down-sampled from 1000Hz to 250Hz, and a 0.1-50Hz bandpass filter was applied. Butterworth notch filtering was used to remove 50Hz and 100Hz frequency noises. Following the removal of the non-brain electrode (see Figure 2), the remaining electrodes (91) were referenced to the average. Finally, linear de-trending and baseline subtraction were performed for each channel of each segment. Analyses were performed using non-overlapping 10-second segments, for an average of 60 segments of EEG recording per subject and condition.

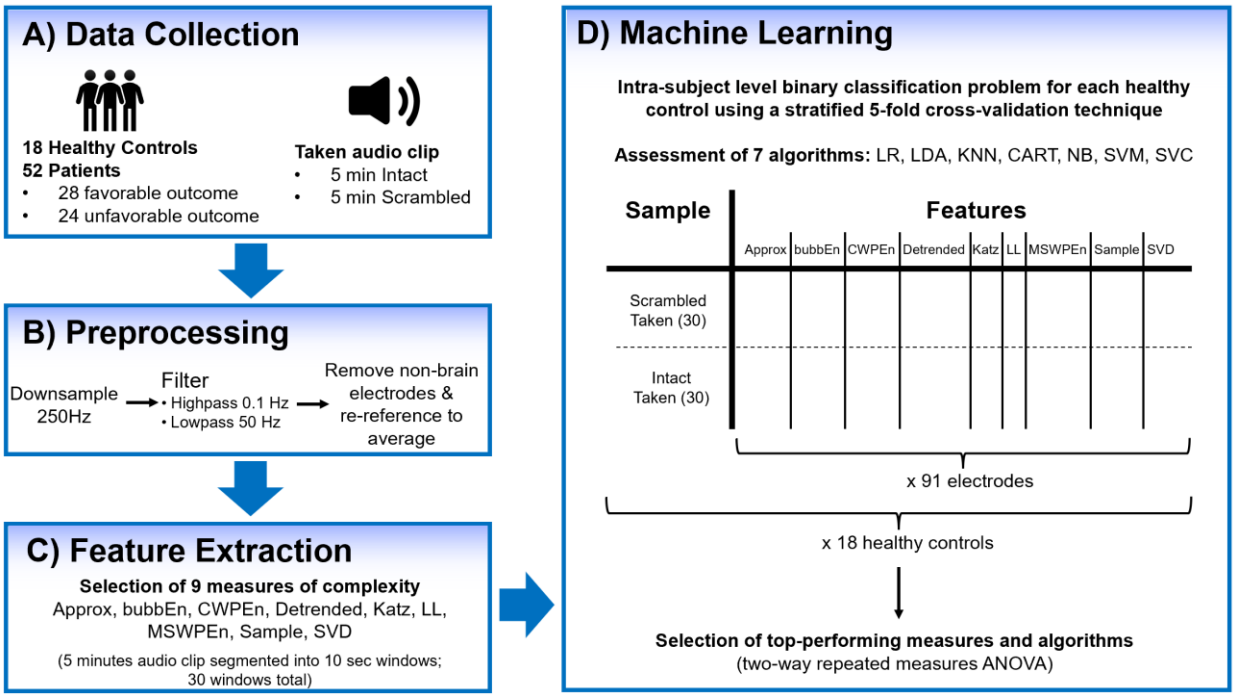
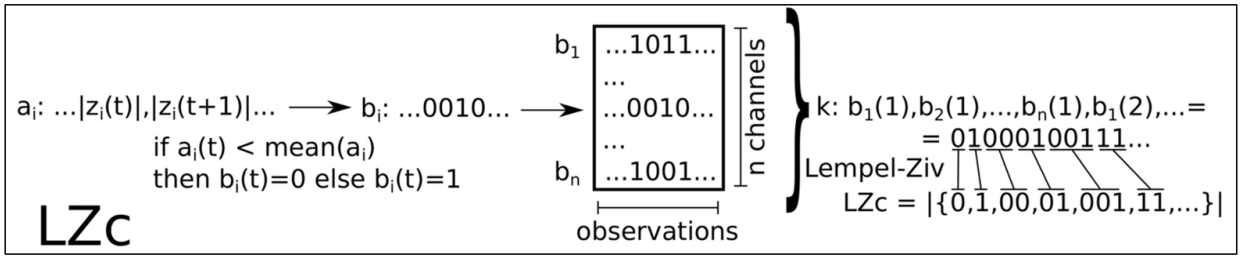


Figure 3. Intra-subject level binary classification problem for participants. An outline of each step done to prepare data before Part I of the main analysis. The figure is a broad representation each step beginning with (A) stimulation presentation and data collection of healthy and patient populations, (B) steps for basic EEG preprocessing, (C) segmenting

continuous time-series data for subsequent feature extraction with various complexity measures using Neurokit2 python toolbox, and finally (D) a visual representation of how each participant's data is organized into samples and features within the Pandas DataFrame in preparation for (Part I) intra-subject binary classification of task.

Feature Extraction

Following preprocessing, various complexity and entropy measures were extracted from the signals of each electrode (91 electrodes) as features. Ten measures were used to calculate the feature sets (SVD, Approx, Sample, Detrended, Katz, LL, MSWPEn, bubbEn, CWPEn, LZc); details of how these measures are calculated and what they measure are well-described in Table 1. These measures were calculated using functions from the Python library called Neurokit2. However, LZc was the only measure calculated using a different method to insure consistency with previous research. LZc was calculated using adapted Python code provided by Schartner et al. (2017) (see Figure 4).



(Schartner et al., 2017)

Figure 4. Lempel-Ziv Complexity (LZc) calculation. Schematic of altered analysis workflow to calculate spatial-temporal variant of LZc using code adapted from Schartner et al. (2017). The data is binarized first, by calculating the Hilbert transform to obtain the instantaneous amplitude. Binary values replace the time series based on the amplitude being each point being higher or lower than a threshold set using the mean over respective source channels. This gives a matrix of

ones and zeros with rows for each channel and columns for each observation within a segment of 10 seconds. The matrix is transformed into a single string of binary digits by concatenating each column over one another, so that subsequent analysis will have spatial and temporal information (unlike previous measures which calculate complexity over time for each spatial location independently). Finally, the binary string is encoded with the Lempel-Ziv compression algorithm (Lempel & Ziv, 1976).

Statistical Analysis – Part I

Given the high-dimensional nature of these measures, I initially used a data-driven approach that included many measures. Data-driven machine-learning approaches are common and can address the challenge of processing large amounts of high-dimensional EEG data (D. Engemann et al., 2015).

The chosen complexity measures I implemented resulted in high-dimensional data. They calculate the temporal complexity of short segments of the data, resulting in multiple values for a single electrode timeline. Added to this is the fact that I utilize a high-density EEG cap and, therefore, end up with many values per participant. Averaging across non-linear measures will result in a loss of information, as one study revealing a distinctive subset of age-relevant features would not have been discovered with mean-based measures (Garrett et al., 2010). Thus, I utilize a machine learning approach that can account for the high dimensional data while also being able to answer my research question.

To avoid bias and overfitting while answering my research question, I used an intra-subject level binary classification of Tasks for each participant. This analysis involves comparing the performance of various intra-subject machine learning models across several measures and models. The models analyzed include Support Vector Classification (SVC), Linear Discriminant Analysis (LDA), K-Nearest Neighbors (KNN), Logistic Regression (LR), Classification and

Regression Trees (CART), Naive Bayes (NB), and Support Vector Machine (SVM). The measures used for comparison include SVD, Approx, Sample, Detrended, Katz, LL, MSWPEn, bubbEn, CWPEn, LZc. By including a wide range of measures and models, I aimed to thoroughly test my hypothesis. Although the selection process can be challenging due to the abundance of measures and models available, my approach allows for comprehensive testing and helps avoid issues of non-independence. This is achieved by narrowing down the top performing measures and models using data from healthy control participants, and then restricting further analysis to include only those best models patient-derived data. This ensures transparency in my results.

First, the various classifiers were used in a stratified 5-fold cross-validation technique to measure their success in classifying between Intact and Scrambled conditions. Accuracy scores for discriminating between intact and scrambled clips were then calculated for each participant (i.e., 5 folds for each of the intra-subject models, meaning each participant has five accuracy scores which are averaged into one). Next, I compared individual subject accuracy scores for each of the model's abilities to correctly predict the scrambled and intact labels of each sample in each participant (60 samples per subject; see Fig.3D). These steps were first completed on healthy controls to find the measures and algorithms that can best discriminate between intact and scrambled audio clips. The top two performing measures and models among the healthy participants were *chosen* for the subsequent step which involves patients this time.

Using only patient data this time, I repeat the previous steps (from the previous passage), but only with the two *chosen* classifiers and measures. This allows me to narrow down the relevant models for the purposes of classifying between tasks while remaining unbiased and avoiding double dipping of the data. Finally, after obtaining accuracy scores for each patient, with the respective chosen models and measures, I assessed whether those results were related to clinical

outcomes. I did this by dividing the patient's accuracy scores into two groups corresponding to the patient's respective outcome: Favourable or Unfavourable.

For each model and measure, individual accuracy scores were used as the dependent variable in a Mann-Whitney U rank test on two independent samples to compare the mean accuracy scores of patients with favourable and unfavourable outcomes. Mann-Whitney test was used because it does not assume a normal distribution of the data and it is appropriate for comparing independent samples. A one-tailed test was conducted to test the alternative hypothesis that the mean scores of the Favourable group are greater than the mean scores of the Unfavourable group. To control for the family-wise error rate due to multiple comparisons, a Bonferroni correction was applied. The significance level (alpha) was set at 0.05 and given that there are 10 comparisons for each measure, the adjusted alpha level was alpha = 0.005.

Bar graphs were generated to visualize the individual patient accuracy scores, incorporating error bars to represent the confidence intervals (Figure 6). Specifically, I used a 95% confidence interval and performed 1000 bootstrap samples to compute these intervals. This approach ensures that the error bars accurately reflect the variability and reliability of the data. Violin plots were used to visualize the group-level differences.

Statistical Analysis – Part II

The second part of the analysis used a different approach. Unlike the machine learning approach from Part I, this statistical approach is limited to investigating one or a few dependent variables at a time. The feature extraction method results in highly dimensional data, making classical statistical methods difficult without major data reduction. To avoid the loss of information

from averaging across non-linear measures and rendering our measures less valuable (Garrett et al., 2010), I used only one measure: spatial-temporal LZc.

The dependent variable 'LZc' is the measurement of each subject under each condition. Importantly, LZc is a measure of spatial-temporal complexity, which is less dimensional compared to the commonly used temporal complexity measures. Consequently, LZc is resistant to noise and is frequently used as a measure of the depth of anesthesia due to its reliability in assessing short, noisy data (Ferenets et al., 2006; Zhang et al., 2001).

Patients were analyzed in a two-way mixed-factorial design with 'Outcome' as the between-subjects factor that divides the sample into discrete subgroups (e.g., 'Favourable', 'Unfavourable', and 'Healthy Control') and 'Task' as the within-subjects factor that represents the different conditions presented to each subject (e.g., 'Intact' and 'Scrambled' or 'Rest'). For the between subjects factor we used 'Favourable' and 'Unfavourable' as individual subgroups, and then repeated the analysis with 'Healthy Control' as an additional subgroup. The control group provides a baseline against which the patient groups can be compared to provide context for interpreting patient data, ensuring that any observed improvements or deteriorations are meaningfully related to complexity trajectories. This can inform prognosis by identifying patterns associated with better or worse outcomes.

We perform a complete case analysis to handle missing values in repeated measures. Using a strict listwise approach, participants with a missing Task condition (due to complications during scanning procedure) are automatically removed. The listwise approach is quite conservative as it completely removes any subject with one or more missing value(s), thereby reducing statistical power. However, it is more appropriate for post-hoc analysis following an ANOVA. 17 patients

with favourable outcomes, 23 patients with unfavourable outcomes, and 18 healthy participants were included in the analysis.

To correct for violations of the assumption of sphericity, all analyses in the study with repeated measures were subjected to the Greenhouse-Geisser adjustment to the degrees of freedom in the ANOVA to provide more accurate p-values.

Post-hoc tests were used after a significant ANOVA result to make pairwise comparisons (t-tests) between group means. To correct for multiple comparisons, the Benjamini/Hochberg FDR correction method was used to adjust p-values. All pairwise comparisons were conducted using the Python Pingouin library. This package provided post hoc-tests for the pairwise multiple comparisons that I performed.

Analysis of Different Frequency Bands. Furthermore, to identify whether LZc differences were driven by a specific frequency band, I repeated the analysis for each frequency band. The EEG signal was filtered into five frequency bands: Delta 1-4Hz, Theta 4-8Hz, Alpha 8-12Hz, Beta 12-30Hz, and Gamma 30-40Hz. Next LZc was extracted from each signal. Finally, statistical analysis was completed for each individual frequency following the same methodology as described above.

Other Measures. Different variations of this analysis were done with different measures and conditions. I investigated other measures of complexity to make sure that results from Part II using spatial-temporal LZc are consistent across other measures and that the lack of task-based differences is not unique to differences in calculation. Therefore, we conducted a similar analysis for all other temporal complexity measures as was done with spatial-temporal LZc (Appendix G).

The temporal complexity measures were averaged across time and electrodes to get a single value for each participant so that normal statistical methods could be applied. However, as previously stated, there is a potential loss of information when averaging high-dimensional data. Even with the loss of some information, it may be informative to investigate other measures to examine the data further. To this end, precaution is taken when interpreting findings, in that results from the various measures with averaged scores are not considered independently; rather, results are used to supplement the main results (Figure 8; spatial-temporal LZc) by observing if similar patterns are present across each measure.

Statistical Analysis – Part III

In this analysis, I used the K-Nearest Neighbours (KNN) algorithm to do a binary classification of favourable and unfavourable outcomes. The KNN algorithm requires numerical input. Therefore, the categorical outcome variable was encoded into numerical values: 1 for 'Favourable' outcomes and 0 for 'Unfavourable' outcomes.

The data set was stratified and split into training and test sets while maintaining the same proportion of samples with different 'outcome' labels in each group. This was done using the train_test_split function from sklearn.model_selection. Furthermore, the train/test split was done while ensuring all samples from the same participant stick together, ensuring that I avoid "double dipping" in that participants with multiple samples are not separated. This is important because training and testing a model is biased if the algorithm is being tested on a participant who has already been seen by the algorithm by way of a different sample in the training set.

Before machine learning, I also normalized the data for numerical stability. Note that I normalized after splitting the data. It is good practice to apply any data transformations to training and testing data separately to prevent data leakage. Data leakage occurs when information from

the training dataset is spilled into the testing dataset, which can lead to overly optimistic performance metrics and poor generalizability.

If the model successfully predicts the patients as Favourable, this case is called True Positive (TP). If the model successfully predicts patients as having an unfavourable outcome, this is called True Negative (TN). The binary classifier may make an incorrect prediction for some patients as well. If a patient with an unfavourable outcome is classified as Favourable by a negative test result, this error is called False Negative (FN). Similarly, if a patient with a favourable outcome is classified as Unfavourable by our algorithm, this error is called a False Positive (FP). I generated confusion matrices to represent this, and I calculated balanced accuracy scores as (TP + TN) / (TP + FP + TN + FN).

Receiver Operating Characteristic (ROC) and Area Under the Curve (AUC) analyses were performed to evaluate the model's ability to differentiate patients with favourable outcomes from those with unfavourable outcomes. The ROC curve was generated by plotting the true positive rate (sensitivity) against the false positive rate (1-specificity) at various threshold settings. The AUC was calculated to quantify the overall performance of the classifier, with higher AUC values indicating better discriminative ability.

The analysis was conducted as follows: The classifier's predicted probabilities for each patient were obtained. Various threshold values were applied to these probabilities to calculate the true positive rate and the false positive rate at each threshold. The ROC curve was plotted using these true positive and false positive rates. The AUC values range from 0 to 1, with 0.5 representing chance level performance and values closer to 1 indicating better classification performance.

Chapter 3: Results

Part I: Comprehensive Evaluation of the Prognostic Utility of Complexity Scores using Intact VS Scrambled Audio Clips

The following results for Part I were completed to test the hypothesis that changes in complexity induced by auditory naturalistic stimulation will index prognosis for acute brain-injured patients. First, I captured the best-performing measures for task-based discrimination in healthy controls (Figure 5). I found that the Support Vector Classifier (SVC; $M = 0.86$, $SD = 0.12$) and Linear Discriminant Analysis (LDA; $M = 0.84$, $SD = 0.12$) were the top-performing machine-learning algorithms among healthy controls (Figure 5). Further, Fractal Line Length Index (LL; $M = 0.87$, $SD = 0.15$) and Conditional Weighted Permutation Entropy (CWPEn; $M = 0.86$, $SD = 0.12$) were the top-performing complexity measures for discriminating between intact and scrambled clips among healthy controls (Figure 5).

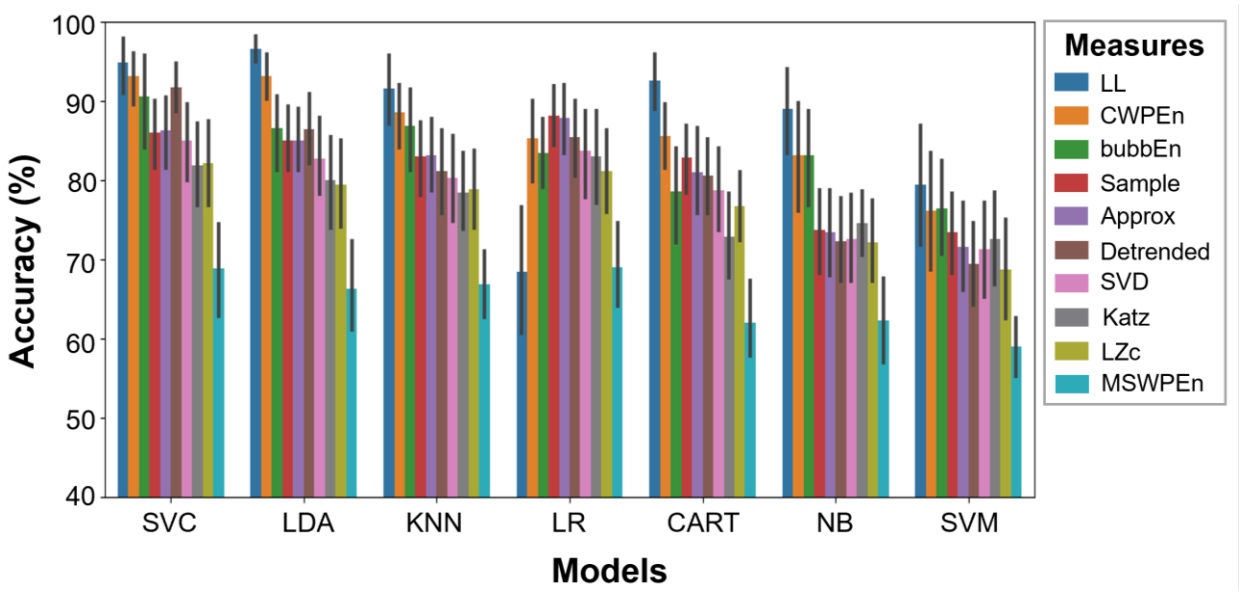


Figure 5. ***Performance of various models and measures for healthy control intra-subject level binary classification of Intact VS Scrambled.*** Mean scores of accuracy from individual healthy control subjects' classification of Intact vs Scrambled for 10 complexity measures assessed using seven different model algorithms. SVC and LDA model algorithms and LL and CWPEn complexity measures were top-performing and selected for downstream analysis.

The purpose of this analysis was to find a metric that I can use to infer a patient's 'ability' to discriminate between the two tasks. Thus, I tested patients using the top models (SVC and LDA) and measures (LL and CWPEn) that were found to be successful in the healthy controls (Figure 6). Since the models and features I used on patients were ones that previously worked on healthy controls, I made the assumption that the accuracy scores represent the metric needed to infer a patient's 'ability' to discriminate between the two tasks.

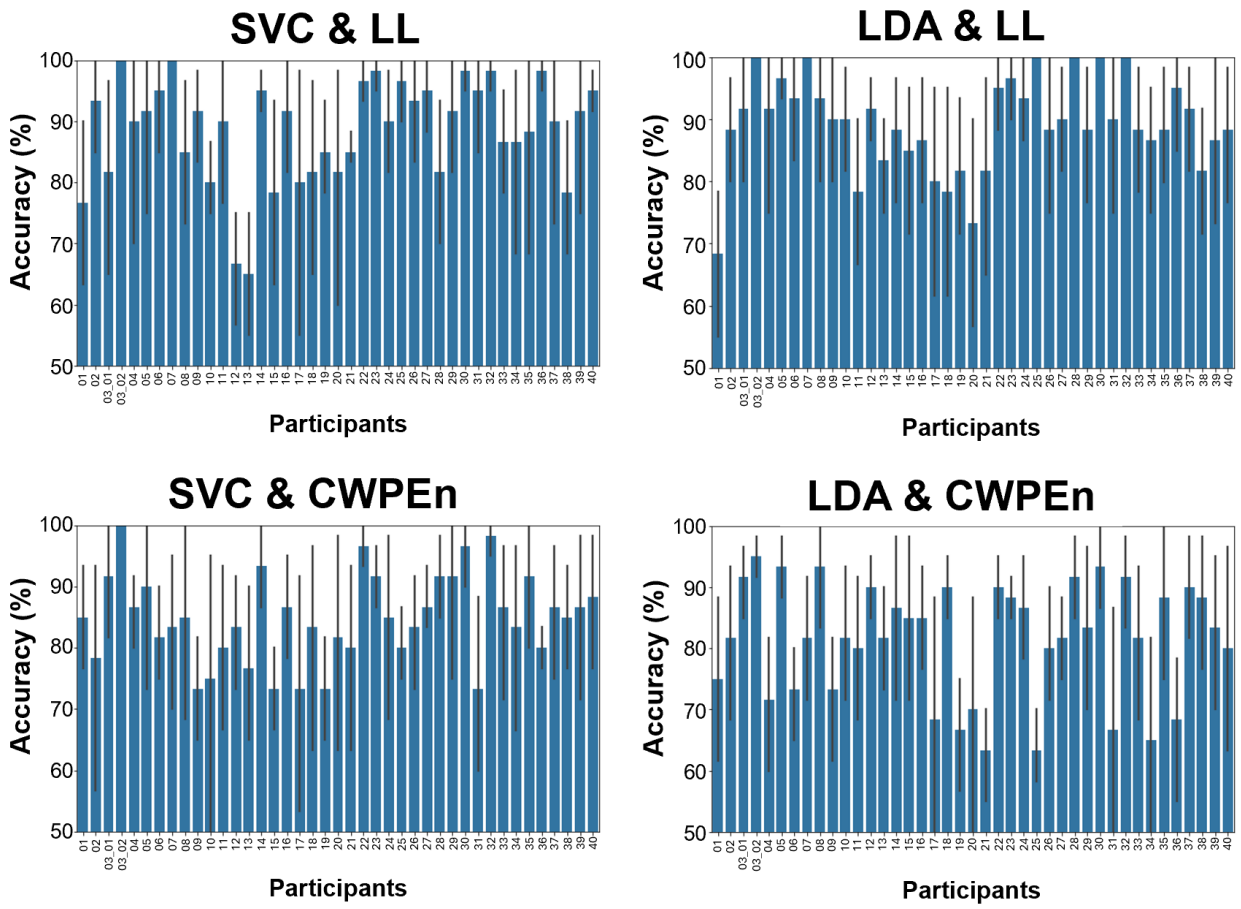


Figure 6. Evaluating top models and measures for classifying brain-injured patients in the ICU. Individual patient accuracy scores of CWPEn and LL measures using LDA and SVC algorithms. For the binary between Intact and Scrambled, each patient's model is tested using a stratified 5-fold cross-validation technique. The accuracy scores correspond to the success of classifying between Intact and Scrambled conditions during each fold of the test-train split (5-folds per model). The Accuracy score for a given fold corresponds to the number of correct predictions over the number of labels in the test set. Each model ends up with five accuracy scores, and the bars in the graph represent the mean accuracy of each participant's model. Error bars represent the 95% confidence intervals computed from 1000 bootstrap samples.

Finally, I used the individual patient accuracy scores that I obtained using various combinations of the top models (LDA and SVC) and measures (LL and CWPEn) to know if a patient's 'ability' to discriminate between intact and scrambled audio clips is a good potential

predictor of their outcome. I compared the mean accuracy scores from the two groups, Favourable and Unfavourable, using a Mann-Whitney U test for each model-measure combination. Surprisingly, the resulting group-level accuracy scores were the same for patients with favourable and unfavourable outcomes (SVC & LL: $p = 0.657$; SVC & CWPEn: $p = 0.983$; LDA & LL: $p = 0.202$; LDA & CWPEn: $p = 0.351$; Figure 7). This means that patients' accuracy scores obtained using a combination of models SVM or LDA and measures LL or CWPEn are not related to a patient's outcome.

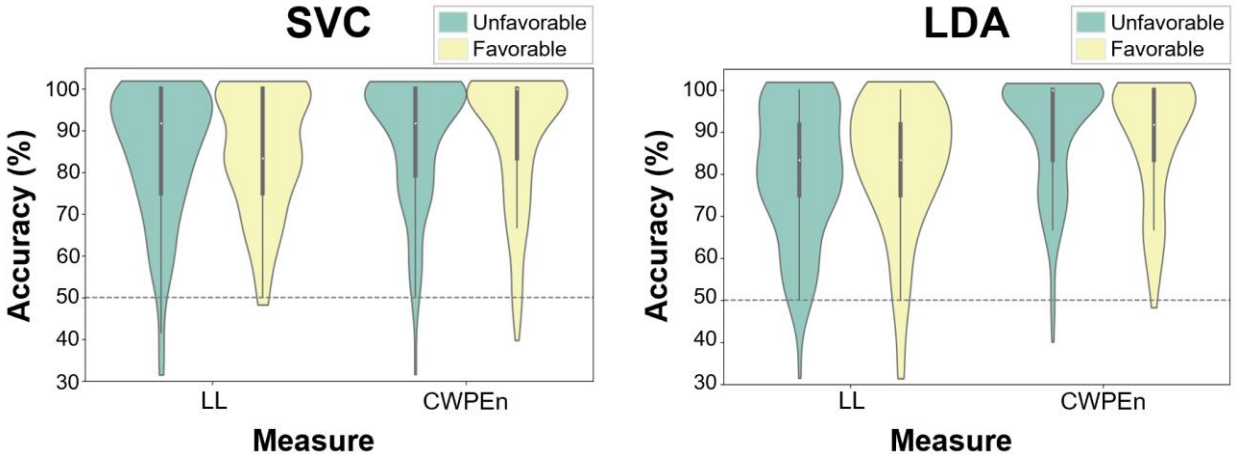


Figure 7. Group-level patient accuracy scores stratified by outcome. Violin plots compare patients with favourable and unfavourable outcomes at the group level by looking at differences in individual accuracy scores from intra-subject binary classifications of Tasks. The accuracy scores are based on model performance using LDA or SVC algorithms with CWPEn or LL measures.

It is possible that there were no group differences because the metric I chose to represent a patient's ability to discriminate between Intact and Scrambled was not sufficient for making this inference. Perhaps the wrong selection of measures and models was made due to some bias that was introduced in our method of narrowing down the options based on their performance with

healthy controls. Therefore, as a final step for Part I, to be sure that the negative findings were not because of the wrong combination of complexity measures and machine learning algorithms, I repeated the analysis for patients, this time using all possible combinations of models and measures.

The resulting performance metrics for classifying Intact vs. Scrambled for all combinations showed no group-level differences between patients with favourable vs. unfavourable outcomes. Individual patient accuracy scores and a bar graph representing means for all combinations are in Appendix B. Means after splitting by Outcome, and the table of statistics are in Appendix C. These results were consistent with my initial findings that using complexity measures to classify between Intact and Scrambled is not related to a patient's outcome.

Notably, the accuracy scores for many patients performed very well at discriminating between Intact and Scrambled. For example, LL could correctly classify between Intact and Scrambled with 83% accuracy. However, this does not support our hypothesis so long as performance is not related to outcome. Part I was set up to answer my original research question (more on this and accuracy scores in the Discussion). We do not know if the reason for negative findings is due to the lack of task-based differences or if existing task-based differences are not prognostic. All we can say for certain is that complexity discrimination of Intact and Scrambled does not index prognosis for acute brain-injured patients. Therefore, I should investigate if complexity, irrespective of the task, can be prognostic.

Part II: Assessing Differences for Outcome and Task using Lempel-Ziv Complexity

The next analysis used spatial-temporal LZc to delineate whether task differences are important for prognosis or if complexity measures themselves are sufficient at the group level. The mixed-factorial analysis produced a significant main effect of Outcome ($F(2, 55) = 5.02$, $p = 0.010$) suggesting that complexity is not the same across favourable, unfavourable, and healthy control subgroups (Figure 8a). However, there was no main effect of Task ($F(1,55) = 1.17$, $p = 0.284$), which may explain the unexpected results from Part I that relied on the accuracy of task discrimination (Table 3). Specifically, patients with unfavourable outcomes had significantly higher LZc (Intact: $M = 0.86$, $SD = 0.07$; Scrambled: $M = 0.85$, $SD = 0.07$) compared to healthy controls (Intact: $M = 0.77$, $SD = 0.08$; Scrambled: $M = 0.78$, $SD = 0.08$) for both Intact ($t(33) = -3.88$, $p = 0.003$) and Scrambled ($t(35) = -2.86$, $p = 0.021$) conditions. However, there was no significant difference (Intact: $t(27) = -0.54$, $p = 0.595$; Scrambled: $t(31) = -0.65$, $p = 0.595$) between healthy controls and patients with favourable outcomes (Intact: $M = 0.79$, $SD = 0.12$; Scrambled: $M = 0.80$, $SD = 0.09$) (Figure 8a, Tables 4-5). Also this analysis revealed a trend in which lower LZc appears to be related to better outcomes.

While there was no main effect of Task, the interaction effect between Task and Outcome was marginally non-significant, $F(2,55) = 2.69$, $p = 0.077$ (see Table 3). This suggests that the effect of the task type on scores showed a trend toward differing across the outcome groups. The post-hoc analysis in Table 5 suggests that Task potentially moderates the difference in complexity between Outcome groups, increasing their difference during intact. This is illustrated by the larger effect sizes seen during the Intact condition as compared to Scrambled. For instance, in the Intact condition there is a significant difference in complexity scores between HC and Unfavourable

outcome groups, with a large effect size $t(33.22) = -3.88$, $p < 0.001$, Hedge's g = -1.22. By contrast, these effects were smaller in the Scrambled condition; HC and Unfavorable are still significantly different with a moderate effect size $t(34.7) = -2.86$, $p = 0.021$, but smaller than the effect seen during intact (Scrambled Hedge's $g = -0.89$; Intact Hedge's $g = -1.22$; Table 5).

Similarly, during the Intact condition complexity scores between patients with Favourable and Unfavourable outcomes are marginally non-significant, with a moderate effect size $t(22.82) = 2.15$, $p = 0.085$, Hedge's $g = 0.72$. On the other hand during Scrambled there are no differences between favourable and unfavorable patients ($t(27.6) = 1.7$, $p = n.s$), and the effect size is once again smaller (Hege's $g = 0.56$; Table 5).

Table 3. ANOVA table for Figure 8a.

Source	SS	DF1	DF2	MS	F	p-value	ng2	eps
Outcome	0.14	2	55	0.07	5.02	**0.010**	0.15	N/A
Task	0	1	55	0	1.17	0.284	0	1
Interaction	0	2	55	0	2.69	0.077	0.01	N/A

Table 4. Descriptive statistics for Figure 8.

Outcome	Task	*n*	*M*	*SD*	*SEM*	Skewness	Kurtosis
Unfavourable	Intact	23	0.8593	0.0670	0.0140	-0.396	-0.5760
	Scrambled	23	0.8496	0.0680	0.0142	-0.470	-0.3590
Favourable	Intact	17	0.7877	0.1245	0.0302	-2.166	4.3265
	Scrambled	17	0.8037	0.0947	0.0230	-1.462	1.4697
Healthy Control	Intact	18	0.7686	0.0795	0.0187	-0.187	-0.7442
	Scrambled	18	0.7847	0.0753	0.0178	-0.387	-0.5965

Table 5. Post Hoc Analysis table for Figure 8a.

Contrast	Task	A	B	T	DOF	p-unc	p-corr	BF10	hedges
Task	-	Intact	Scrambled	-1.05	57	0.297	N/A	0.243	-0.06
Outcome	-	Healthy Control	Unfavourable	-3.52	35.16	**0.001**	**0.004**	28.42	-1.1
Outcome	-	Healthy Control	Favourable	-0.61	27.83	0.547	0.547	0.367	-0.2
Outcome	-	Unfavourable	Favourable	1.98	24.97	0.059	0.088	1.418	0.66
Task*Outcome	Intact	Healthy Control	Unfavourable	-3.88	33.22	**<.001**	**0.003**	67.38	-1.22
Task*Outcome	Intact	Healthy Control	Favourable	-0.54	26.94	0.595	0.595	0.364	-0.18
Task*Outcome	Intact	Unfavourable	Favourable	2.15	22.82	**0.042**	0.085	1.841	0.72
Task*Outcome	Scrambled	Healthy Control	Unfavourable	-2.86	34.7	**0.007**	**0.021**	6.678	-0.89
Task*Outcome	Scrambled	Healthy Control	Favourable	-0.65	30.56	0.518	0.595	0.384	-0.22
Task*Outcome	Scrambled	Unfavourable	Favourable	1.7	27.6	0.100	0.150	0.961	0.56

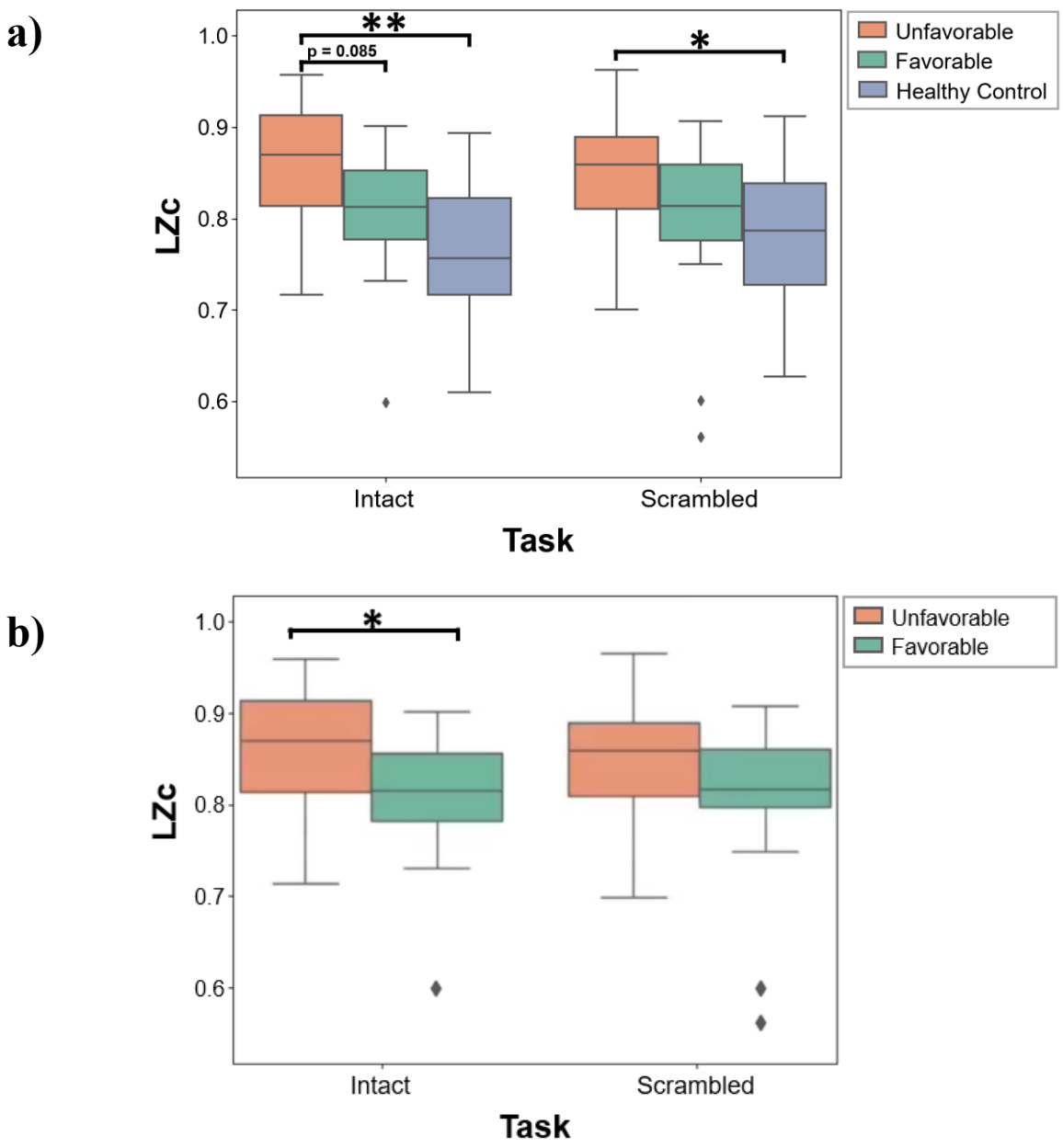


Figure 8. Mixed-factorial plot of LZC for outcome and task (Intact VS Scrambled). Boxplot of spatial-temporal LZc scores grouped by Task and Outcome in a mixed-factorial design study. **(a)** The first graph includes Healthy control as part of the analysis as an additional between subjects group; **(b)** The second graph only includes the two patient groups, Favorable and Unfavorable.

This interaction effect is further accentuated by the second mixed factorial analysis, which excluded the HC subgroup this time. (Figure 8b). Similar to previous results, I found a significant main effect of Outcome ($F(1, 38) = 4.5$, $p = 0.041$) and no main effect of Task ($F(1,38) = 0.05$, *n.s*). However, this time, the interaction effect between Task and Outcome was statistically significant ($F(1,38) = 5.51$, $p = 0.024$; Table 6).

Once again, the follow-up post-hoc analysis in Table 7 reveals the nature of this interaction. Unfavorable vs Favorable in the Intact condition is marginally non-significant ($t(22.82) = 2.15$, *p-unc*=0.042, *p-corr* = 0.085), with a moderate-high effect size (Hedge's $g = 0.72$). Compared to the lack of difference seen during the Scrambled condition ($t(27.6) = 1.7$, *n.s*, Hedge's $g = 0.56$), it would seem that the Task moderates complexity scores in different outcome groups, with the Unfavourable group showing larger LZc scores compared to the Favourable, especially in the Intact task.

The follow-up post-hoc analysis presented in Table 7 reveals the nature of this interaction. The comparison between the Unfavorable and Favorable groups in the Intact condition is marginally non-significant ($t(22.82) = 2.15$, *p-unc*=0.042, *p-corr* = 0.085), with a moderate-high effect size (Hedge's $g = 0.72$). In contrast, there was no significant difference observed in the Scrambled condition $t(27.6) = 1.7$, *n.s*, Hedge's $g = 0.56$). Which is the same pattern reported previously.

Table 6. ANOVA table for Figure 8b.

Source	SS	DF1	DF2	MS	F	p-value	ng2	eps
Outcome	0.07	1	38	0.07	4.5	**0.041**	0.1	N/A
Task	0	1	38	0	0.05	0.817	0	1
Interaction	0	1	38	0	5.51	**0.024**	0.01	N/A

Table 7. Post Hoc analysis table for Figure 8b.

Contrast	Task	A	B	T	DOF	p-unc	p-corr	BF10	hedges
Task	-	Intact	Scrambled	-0.22	39	0.827	N/A	0.243	-0.06
Outcome	-	Unfavourable	Favourable	1.98	24.97	0.059	N/A	1.418	0.66
Task*Outcome	Intact	Unfavourable	Favourable	2.15	22.82	**0.042**	0.085	1.841	0.72
Task*Outcome	Scrambled	Unfavourable	Favourable	1.7	27.6	0.100	0.100	0.961	0.56

These interaction effects indicate that the Task type can moderate complexity scores to exacerbate differences seen between outcome groups, such as higher complexity in Unfavorable patients compared to the Favourable and Healthy Control groups, especially in the Intact task. While the nature of the Task does affect complexity, this influence is still not significant enough to detect an individual's internal changes in complexity in response to Task-stimulation. This is evident from the inconclusive results in Part I and the lack of a significant group-level difference in task performance between the Intact and Scrambled conditions.

Importantly, to investigate the lack of task-based difference, I wanted to assess if the contrast between Scrambled and Intact was not large enough to see a difference, as many studies use Resting State as a baseline. Therefore, I used the Rest condition instead to see if such differences can be extrapolated with the larger contrast between tasks. The results indicate that there was no main effect of Task ($F(1,55) = 0$, $p = 0.997$), even when I replaced the Scrambled condition with the Resting State (Appendix D). Moreover, the effect of Outcome persisted as well ($F(2,55) = 7.21$, $p = 0.002$). Notably, this trend was even more drastic, with patients with favourable outcomes having significantly lower LZc scores ($M = 0.79$, $SD = 0.07$) than those with unfavourable outcomes ($M = 0.84$, $SD = 0.07$) in the Resting State condition ($t(35) = 2.26$, $p = 0.045$ (Appendix D). The confirmatory analysis once again showed that there was no difference between groups based on Task, but a significant main effect of Outcome.

It is important to note that the direction of the effect observed was in the opposite direction than I expected. Although many papers report higher complexity for more favourable outcomes, I observed greater values in the Unfavourable group compared to Favourable and Healthy Controls. However, it is suggested that the time scale parameter used to calculate complexity can affect the directional trend of the results (Shi et al., 2017). This has been demonstrated by a sleep study that found that complexity was higher during wakefulness than deep sleep at larger time scales but lower when smaller time scales were used (<0.04 s; Shi et al., 2017). Therefore, I calculated LZc using shorter time scales (<0.04 s) to see if the relative differences between LZc scores in each group were affected by this parameter. My results are consistent with these findings, as Appendix E shows that when I use smaller time scales, the results are flipped. Indeed, LZc scores for healthy controls (Intact: $M = 0.94$, $SD = 0.05$, Scrambled: $M = 0.94$, $SD = 0.05$) were now significantly higher than patients with unfavourable outcomes (Intact: $M = 0.89$, $SD = 0.07$, Scrambled: $M = 0.88$, $SD = 0.07$) for both Intact ($t(38) = 3.0$, $p = 0.018$) and Scrambled conditions ($t(39) = 2.91$, $p = 0.018$). When using Resting State conditions, LZc was also significantly higher for patients with favourable outcomes ($M = 0.94$, $SD = 0.06$) when compared to unfavourable outcomes ($M = 0.89$, $SD = 0.07$) ($t(35) = -2.52$, $p = 0.033$). This is an important point of consideration for methods of analysis and for interpreting the meaning of LZc scores.

Additionally, to ensure that my results were due to an actual difference in prognosis and not related to increased movement inherent in patients with high GCS scores, I performed a correlational analysis between LZc and GCS scores (Appendix F). I found that LZc and GCS were not correlated ($r = -0.20$, $p = 0.192$).

To further characterize LZc differences between outcomes, I also split my analysis by frequency band to see if the differences were being driven by a specific frequency band. After

calculating LZc for each frequency band, I found that the Delta and Alpha bands were likely driving the differences between outcomes (Figure 9). For the Delta band, LZc scores were significantly higher for patients with unfavourable outcomes (Intact: M = 0.71, SD = 0.06; Scrambled: M = 0.72, SD = 0.07) than healthy controls (Intact: M = 0.65, SD = 0.10; Scrambled: M = 0.66, SD = 0.09) for both Intact ($t(35)$ = -3.79, p = 0.003) and Scrambled ($t(37)$ = -3.04, p = 0.013) conditions. The same finding was also found for the Alpha band with LZc scores being significantly higher in patients with unfavourable outcomes (Intact: M = 0.80, SD = 0.05; Scrambled: M = 0.80, SD = 0.06) than in healthy controls (Intact: M = 0.72, SD = 0.08; Scrambled: M = 0.71, SD = 0.09) for both Intact ($t(38)$ = -4.29, p = 0.001) and Scrambled ($t(38)$ = -3.63, p = 0.002) conditions. However, this difference was not identified in the other frequency bands that I assessed.

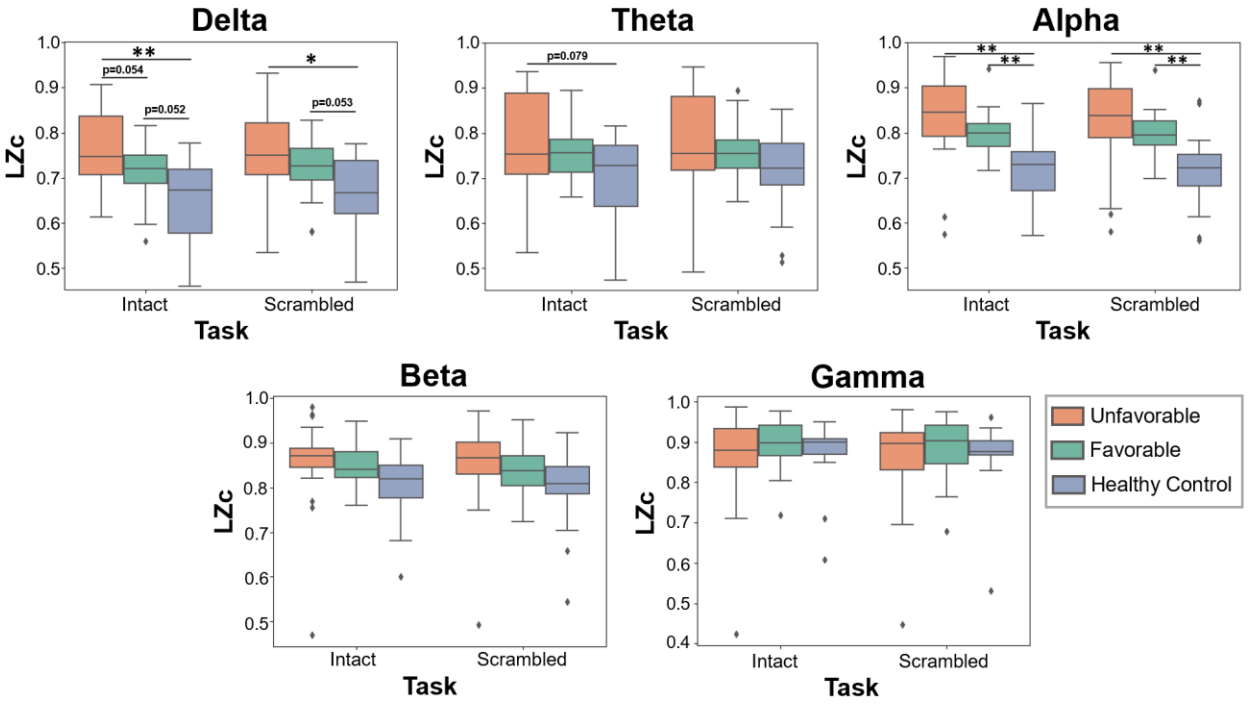


Figure 9. Mixed-factorial plots of LZc for different frequency bands.

During our study, we had a case in which a patient (05_1) with Guillain-Barré syndrome (GBS) recovered and was able to be rescanned 12 months later (Bauer et al., 1979). GBS is an autoimmune condition that affects the peripheral nervous system and can lead to total paralysis, rendering patients behaviorally unresponsive. This created a very unique opportunity to map her data onto our existing results to assess the potential of using LZc for assessing covert consciousness. It also allowed for a within-subject comparison of LZc.

Following this analysis using the original parameters, I found that when this patient was still in the ICU, they had a higher LZc compared to when they recovered (Figure 10). While the patient still fell within the range of favourable outcomes, I would expect this patient to have similar scores to healthy controls in both their scans since GBS patients retain the capacity for conscious awareness. However, it is difficult to know exactly what conscious state the patient was in during the time of testing, which may have possibly affected their LZc score. If so, this is an important limitation to consider, as the current method may underestimate a patient's true residual cognition. Nonetheless, scores from their 1st and 2nd scan seem to slightly align within the range of the Favourable and Healthy Control groups, respectively. Suggesting that, at the very least, scores seem to sometimes capture aspects of healthy brain function.

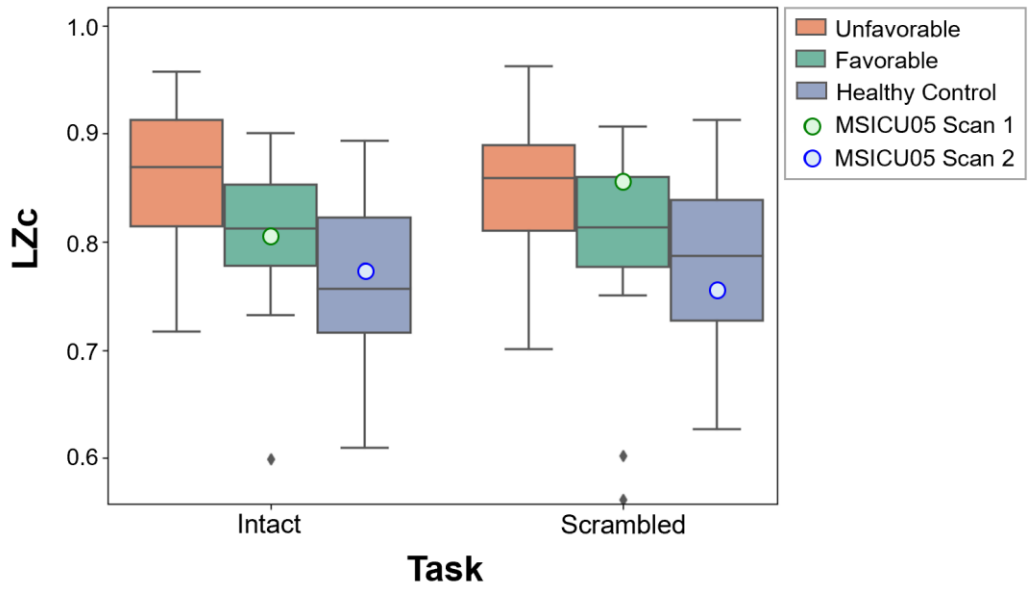


Figure 10. Guillain-Barré syndrome case study: LZC scores of a patient rescanned following recovery.

Finally, I also investigated other measures of complexity to make sure that results from Part II using spatial-temporal LZc are consistent across other measures and that the lack of task-based differences is not unique to differences in calculation. Therefore, an analysis similar to what was done with spatial-temporal LZc, was conducted for all other temporal complexity measures (Appendix G). Interestingly, while the main effects of Outcome were observed for some measures, no main effect of Task was seen in any of the measures (Appendix G). This further supports our conclusion that the true value of complexity measures is in their prognostic abilities rather than discriminating between tasks.

PART III: Predicting Outcome at an Individual Level

In Part II, I established that complexity measures are not sensitive enough to capture task-based differences, which could explain the null results from Part I, which relied heavily on this assumption. Conversely, Part II revealed that complexity scores are able to show group-level differences in Outcomes. Therefore, for Part III, I decided to explore various complexity measures with machine learning once again without classifying tasks. Instead, I classified patients based on Outcome. In other words, I evaluated individual-level classification of acute brain-injured patients using various complexity measures to assess its potential for prognosis.

The features for the models in Part III were created using only Resting State data. This decision was made because my previous results indicated that complexity scores varied between patients with favourable and unfavourable outcomes, regardless of the task used. Therefore, for Part III, I did not use task-based features. Additionally, this would allow for a larger sample size as more patients had resting-state scans. A total of 45 samples/patients were included to create and test the models, with 25 having an unfavourable outcome and 20 having a favourable outcome.

In this analysis, I used the K-Nearest Neighbours (KNN) algorithm to do a binary classification of Favourable and Unfavourable outcomes. For each model, a different complexity measure was used to calculate the respective model's feature set. Confusion matrices were then generated to allow for the calculation of accuracy, sensitivity (True Positive Rate; TPR), and specificity (True Negative Rate; TNR) (Figure 11). These calculated values can be found in Table 8.

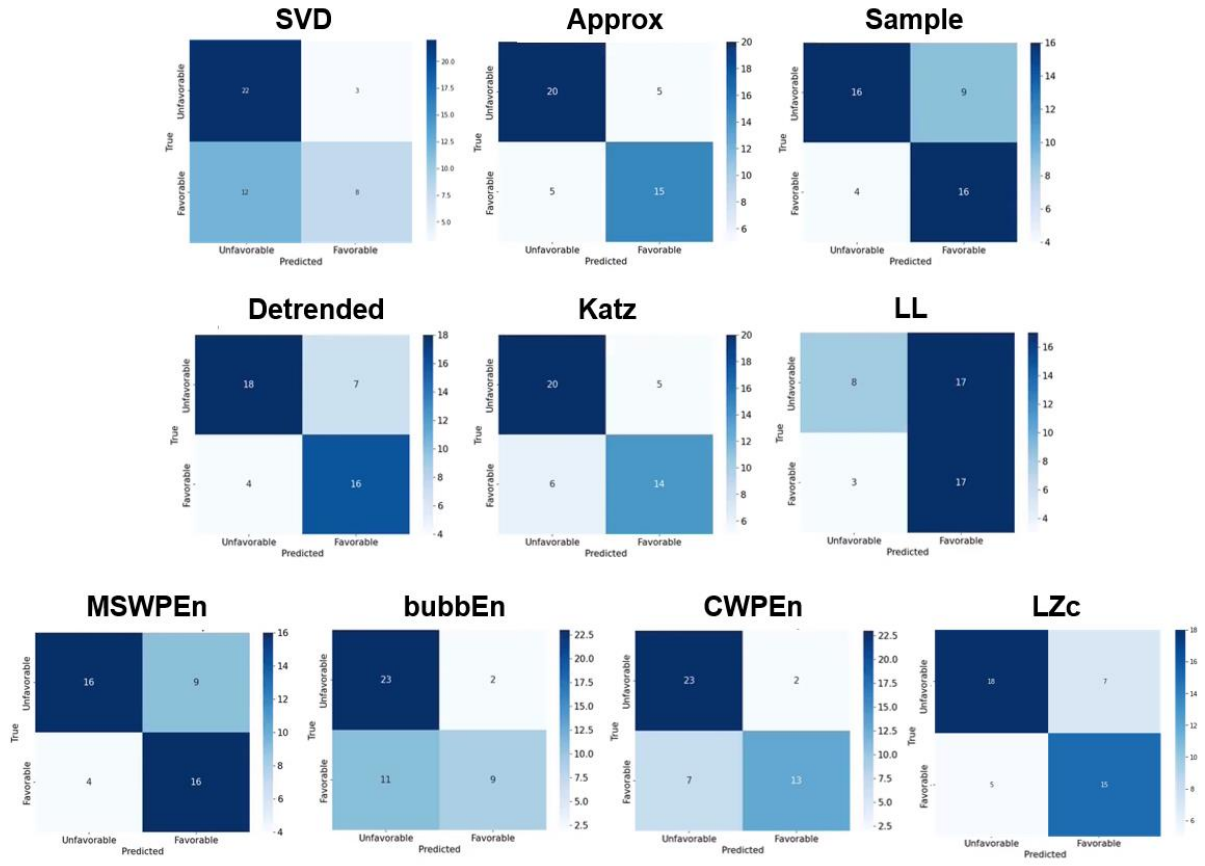


Figure 11. Confusion matrices for each KNN-based model using different measures of complexity to classify patient outcomes. Each confusion matrix is a 2x2 graph which denotes the performance of a model in predicting the outcome of a patient. For the matrix, the Y-axis refers to the true outcome of the patient (the bottom row is Favourable, and the top row is Unfavourable), and the X-axis refers to the models' predictions (the first column is Unfavourable prediction, the second column is Favourable prediction). The integers in the matrix show the number of samples that fall within each square. The darker shade of blue the higher the integer value is. The models are created from the KNN algorithm, and a different complexity measure was used to calculate each respective model's feature set. Resting-state data was used to calculate the features of each sample (patient) in the model. Thus, we can compare the relative performance of the various complexity measures at classifying patient outcome.

Overall, the models were reasonably accurate in classifying the outcomes of patients, with the most accurate model being CWPEn, with an accuracy of 80%, and the least accurate model being LL, with an accuracy of 55% (Table 8). When looking at sensitivity, the most sensitive model was LL (*TPR* = 0.85), and the least sensitive model was SVD (*TPR* = 0.40). Moreover, the most specific models were CPWEn (*TNR* = 0.92) and bubbEn (*TNR* = 0.92), while the least specific model was LL (*TNR* = 0.32).

Table 8. Area under the curve, accuracy, sensitivity, and specificity of all measures using the KNN algorithm.

Measure	*Accuracy*	*Sensitivity (TPR)*	*Specificity (TNR)*	*Area Under the Curve (AUC)*
LZc	0.73	0.75	0.72	0.79
SVD	0.66	0.4	0.88	0.78
Approx	0.78	0.75	0.80	0.77
Sample	0.71	0.80	0.64	0.78
Detrended	0.76	0.80	0.72	0.80
Katz	0.76	0.70	0.80	0.79
LL	0.55	0.85	0.32	0.62
MSWPEn	0.71	0.80	0.64	0.78
bubbEn	0.71	0.45	0.92	0.61
CWPEn	0.80	0.65	0.92	0.83

It is important to note that these performance metrics are not always an accurate depiction of a model's overall performance. For example, LL had the highest sensitivity but the lowest accuracy and specificity. Therefore, I performed an additional analysis to calculate the AUC of ROC curves, which provides a better well-rounded measure of a model's overall performance (Table 8; Figure 12). Overall, the AUC assesses the separability and ability to distinguish between Favourable and Unfavourable outcomes for each measure. Based on the calculated AUCs, the best overall performing models were CPWEn (AUC = 0.83) and Detrended (AUC = 0.80). Both of these top models also had relatively high accuracy (CPWEn = 80%, Detrended = 76%), sensitivity

(CPWEn = 0.65, Detrended = 0.8), and specificity (CPWEn = 0.92, Detrended = 0.72). Overall, this demonstrates the potential application of utilizing EEG complexity measures as features for prognostic use in the ICU.

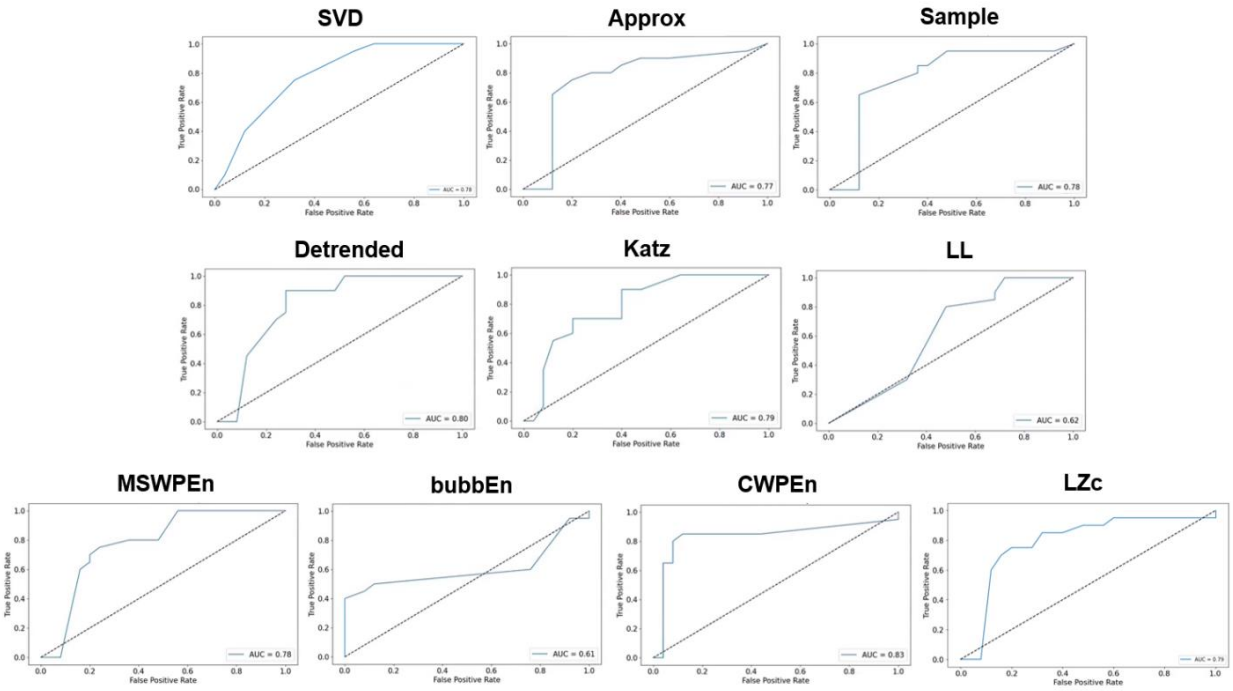


Figure 12. ROC Curves of each KNN-based model using different measures of complexity to classify patient outcomes. Each graph uses an ROC curve to denote the performance of a model in predicting the outcome of a patient as Favourable or Unfavourable. The X-axis represents the false positive rate, and the Y-axis represents the true positive rate. The ROC curve shows the degree of separability or overlaps with the diagonal line. The AUC is a single metric that summarizes the models' overall performance, with values above 0.5 denoting above-chance performance, and values closer to 1 would suggest the model's discriminative ability is even better. The models are created from the KNN algorithm, and a different complexity measure was used to calculate each respective model's feature set. Resting-state data was used to calculate the features of each sample/patient in the model. Thus, we can compare the relative performance of the various complexity measures in classifying patient outcomes.

Chapter 4: Discussion

Summary of Results

My study investigated the potential of using EEG complexity measures as early brain-based markers of recovery from serious brain injury. I examined differences in EEG complexity between auditory naturalistic stimulation and its scrambled version. The study involved two main parts:

In Part I, I used intra-subject level binary classification to avoid bias and overfitting. I found that Support Vector Machine (SVM) and Linear Discriminant Analysis (LDA) were the top-performing algorithms, while Fractal Line Length index (LL) and Conditional Weighted Permutation Entropy (CWPEn) were the best complexity measures for discriminating between intact and scrambled clips. However, these models were not able to predict patient outcomes, as accuracy scores were similar across groups.

In Part II, I focused on Lempel-Ziv complexity (LZc) and used a two-way mixed-factorial design. I found a main effect of Outcome but no main effect of Task and no interaction effect. This suggests that while LZc measures differed significantly between patients with Favorable and Unfavorable outcomes, the task-based differences were not sensitive enough in our patient population or paradigm.

The findings from Part I & II highlight the challenges in using EEG complexity measures to predict patient recovery and suggest that the quality of complex brain activity, rather than task-based differences, may be more informative for prognosis in acute brain-injured patients. I explored this claim further with an additional analysis step that takes findings from Parts I and II into consideration. In part III of this study, I aimed to investigate the potential of various complexity measures for individual-level classification of acute brain-injured patients based on

their outcomes. This analysis yielded promising results, with several measures showing good discriminative ability. For instance, the Lempel-Ziv complexity (LZc) measure achieved an Area Under the Curve (AUC) of 0.79 and an accuracy of 0.73 in classifying patient outcomes. Other measures, such as Detrended Fluctuation Analysis and Conditional Weighted Permutation Entropy, also performed well, with AUCs of 0.80 and 0.83, respectively. These findings suggest that EEG complexity measures have potential as prognostic tools for individual patients.

Now, let us discuss in more detail the findings from each part of my study and their implications.

Part I Discussion

It is important to acknowledge that in this analysis, I made the assumption that accuracy scores are the metric needed to infer a patient's 'ability' to discriminate between the two tasks. Therefore, to see a difference between the two outcomes, there would need to be task differences to begin with. Thus, I do not know if the negative results in Part I were due to a lack of task-based differences or if the task-based differences themselves are not prognostic. This necessitated my Part II analysis.

Notably, the accuracy scores for discriminating between Intact and Scrambled were quite high for many participants (Fig. 6). However, this does not support our hypothesis so long as performance is not related to outcome. Although it is understandable that upon seeing this, one might presume that complexity is amazing for task discrimination, just not sensitive for the purpose of prognosis, I caution against this claim as there could be alternative explanations. For instance, it is more likely that accuracy scores are just over-estimated due to overfitting, which was confirmed by my Part II results. Inherently, intra-subject level binary classification means all

samples involved in creating and testing the models all come from the same participant. In other words, no matter how I split the data for training and subsequent testing, the model will have seen similar data during training. Therefore, an overestimation of the resulting accuracy scores is to be expected.

Importantly, an overestimation of accuracy scores does not take away from the results in Part I. If task differences did exist and these differences were related to outcome, my Part I method would still be able to capture this difference at the group level. Indeed, if the multiple algorithms (representing each participant separately) are all being employed with the same set of features (complexity measures) from a population of people who are sensitive to task-based differences, there would be a systematic difference in accuracy scores between that group and a group of people who are not sensitive to task-based differences. This means that the absolute values of the accuracy scores should be interpreted with caution.

Intra-subject level classification is used in Part I of the study for several important reasons. This approach was chosen as an initial step to thoroughly investigate the research question while preserving as much individual-level data as possible. By using intra-subject classification, I could analyze each participant's data separately, potentially revealing patterns in their complexity that might be obscured otherwise. The benefit is that this analysis is not restricted by the high dimensional data that is so common in EEG and complexity research (D. Engemann et al., 2015). It was very important to design the experiment in this way, especially as a first step to prevent the loss of information through averaging (Garrett et al., 2010). Additionally, Part I allows for a comprehensive assessment of a broad number of measures with multiple models to remain unbiased and avoid cherry-picking when evaluating complexity measures for the purpose of testing

my hypothesis. This is needed, especially because there are so many measures to represent complexity mathematically (Shalizi, 2006).

Despite the lack of group-level differences, examining individual scores can still provide valuable insights into how complexity measures relate to patient outcomes. A notable example is the GBS patient (patient 03), who achieved the highest accuracy scores across all complexity measures (Figure 6). This patient, known to be conscious during the scan despite being unable to communicate, consistently demonstrated high accuracy scores. These results suggest that complexity measures may indeed be sensitive to a patient's level of consciousness and cognitive function. The fact that this patient's scores were markedly higher than those of patients with less favourable outcomes offers some evidence that accuracy scores from complexity measures could be linked to patient outcomes.

However, it is crucial to exercise caution when interpreting these findings. This is a single case, and we cannot draw definitive conclusions from one example alone. As mentioned earlier, there were no significant differences at the group level (Figure 7), indicating that the relationship between accuracy scores and outcomes requires further investigation. Part II of our study will help determine whether the results of Part I stem from a lack of task-based differences or if complexity measures themselves might be sufficient for prognosis at the group level.

Part II Discussion

In Part II, I examined whether the complexity of EEG with the “Taken” paradigm analyzed during acute ICU admission would differ from that of healthy adults. I found a main effect of the between-subjects variable, Outcome. I also identified a trend in which complexity scores for patients with greater recovery and emergence from a coma were different from patients who did

not recover. Similarly, the complexity scores of healthy controls were found to be statistically different from patients who did not recover or both Intact ($t(33) = -3.88$, $p = 0.003$) and Scrambled ($t(35) = -2.86$, $p = 0.021$). What's more, complexity was highly similar between healthy controls and those patients who eventually recovered. Even where subtle differences are seen between healthy controls and those patients who eventually recovered, the trend remains in a consistent direction. It correlates with the general 'health' of the respective groups: highest in the Unfavourable condition, intermediate in the Favourable condition, and least for the Healthy Control group. This directional trend was further confirmed by our case study with a GBS patient (Figure 10). Together, this suggested that complexity values themselves may offer a valuable predictor of current functional integrity as well as future functional outcomes in brain-damaged patients.

Directionality of Findings

Importantly, the direction of this effect was opposite from what I expected. Most papers report higher complexity in healthy states and lower complexity in DoC (Casali et al., 2013; Casarotto et al., 2016; D. A. Engemann et al., 2018; Lei et al., 2022; Luppi et al., 2019; Sitt et al., 2014). More broadly, there is a conventional belief that higher complexity always indicates a healthy physiological system (Lipsitz & Goldberger, 1992). This is because, from both a conceptual and quantification point of view, complexity is often interpreted as irregularity, unpredictability, desynchrony, or randomness (Stam, 2005). However, complex systems are neither completely ordered nor absolutely random (Costa et al., 2005; Goldberger, Peng, et al., 2002). As mentioned in the introduction, complexity is in an intermediate state between randomness and order (Stam, 2005; Tononi & Edelman, 1998; Yang & Tsai, 2013). Thus, an

abnormal brain complexity would give rise to either highly ordered or highly random brain patterns. Both regular and random patterns can be indicative of pathology and represent a deviation from complexity (Goldberger, Peng, et al., 2002). The challenge is that some measures of complexity actually index randomness (i.e., entropy) and are not direct estimates of physiological complexity (Costa et al., 2005; Goldberger, Peng, et al., 2002). Measures like LZc, ApEn, and SampEn give high values with completely random data (Costa et al., 2002, 2005; Lempel & Ziv, 1976), whereas detrended fluctuation analysis is a measure of predictability (Lau et al., 2022). Consistent with this view, we see that while my results for LZc appear to be negatively correlated with prognosis (figure 8), the opposite is true using the Detrended measure, showing a significant main effect of the Outcome but in the opposite direction $p = 0.033$ (Appendix G).

Another issue to consider on the topic of directionality is the way we calculate complexity and the hyperparameters that need to be considered. In the literature, higher values of complexity are generally found in schizophrenia (Hoptman et al., 2010). However, some studies have also noted lower complexity in schizophrenics (Fernández et al., 2013). Similarly, lower complexity is generally reported for Alzheimer's disease (Dauwels et al., 2010; Escudero et al., 2006; Stam, 2005; Stam et al., 2009; Takahashi, 2013). However, recent Alzheimer's studies have reported mixed results, demonstrating a reduction in complexity in severe Alzheimer's at larger time scales, but the opposite effect tended to occur at coarser time scales (Escudero et al., 2006; Mizuno et al., 2010; Park et al., 2007). Changing time scales as a reason for opposite results is further corroborated by a sleep study, where at small scales (<0.04 s), the entropy is higher during wakefulness and increasing time scales in contrast, entropy is higher during deep sleep and lower with increasing time scales. (W. Shi et al., 2017). My results are consistent with these findings; Appendix E shows that when I use small scales (<0.04 s), my results are flipped. Thus, complexity

constantly appears reliably to discriminate Unfavourable, Favourable, and Healthy Controls, albeit in opposite ways.

Lastly, this issue of directionality would have been alleviated if I could use task-based differences to index prognosis. This is what I originally intended since all I would need to look for is if the task induces a change. In the absence of tasks, I would be limited to making inferences on directionality. Task-based change in complexity is an important metric for this reason.

While complexity is related to state-based differences in brain activity, it can be difficult to make concrete inferences about an individual's experience with absolute values alone. Conceptual and computational limitations make such inferences difficult, as different theories and measures of complexity may have opposing yet discriminatory applications and results. Moreover, complexity values can be highly variable because of many inter-dependent factors like the measures used, the neuroimaging method, analysis techniques, and the recording environment. For instance, some measures of complexity sensitive to noise can show abnormally high values of complexity due to external stimuli from the environment. Thus, interpreting the difference between noisy data and abnormal brain activity can be difficult without a baseline measure. That is why controlling the environment and finding baseline metrics through the use of tasks can be informative for understanding state-related differences in brain activity.

Task-Based Differences

In Part II, my mixed-factorial design study revealed a main effect of Outcome but no main effect of Task and no Interaction Effect, meaning that the tasks (Intact and Scrambled) did not significantly influence EEG complexity scores. This was true not only for LZc but for all the other measures that I tested as well (see summary table in Appendix G). Additionally, when I used the

Rest condition instead of the Scrambled condition to check for differences, my results still showed no main effect of Task, but a significant main effect of Outcome (Appendix D).

On the other hand, we observed an interaction effect indicating that tasks can moderate complexity scores across different outcome groups. Specifically, we found that the differences in complexity between Unfavorable and Favorable outcomes were larger when the Intact task was used compared to when the Scrambled task was employed. While this is promising, the effect was not strong enough to predict patient outcomes, possibly because listening to the intact version of the audio clip did not provoke sufficiently different neural responses to be detected by the complexity measures.

There are a couple of reasons as to how this lack of task-based differences may arise. Firstly, the detection of task-based differences may be constrained by data quality. Task-related variations in brain activity may exhibit less pronounced moment-to-moment changes compared to overall differences between different brain states. Therefore, more sensitive measures are required to capture these subtle changes in complexity induced by tasks. Data noise can reduce the signal-to-noise ratio, posing challenges in capturing the nuanced differences between tasks.

Secondly, the ability to detect task-based differences may be dependent on the population being tested. Previous research indicates that older and poorer-performing adults exhibit smaller transitions between internal and external brain states and less overall signal variability, which is related to complexity (Garrett et al., 2010). In contrast, studies in younger adults have shown that brain signal variability increases during task performance compared to rest, which correlates with better task performance (Wutte et al., 2011). Initially, this can be used as evidence which supports the idea of using task-based change for differentiating between populations. However, when we consider that the task itself may be too difficult for certain populations, we also need to consider

fine-tuning the task to be easy enough to maximally differentiate the specific population of interest. Mixed results reported in the literature suggest that the relationship between task difficulty and brain variability may depend on the cognitive demands of the task and the participants' abilities (Garrett et al., 2013; He, 2011; Takahashi et al., 2009; Wutte et al., 2011).

Thirdly, the lack of observable task-based differences in complexity might be due to the need for a more substantial change in conscious content. Previous research has shown that complexity measures are sensitive to significant alterations in cognitive states, such as the transition between sleep and wakefulness (Tosun et al., 2017) or the effects of psychedelic substances (Lebedev et al., 2016; Petri et al., 2014; M. M. Schartner et al., 2017; Tagliazucchi et al., 2014; Viol et al., 2017). These changes in brain signal patterns occur alongside the intense alterations in perception, cognition, and emotional states that are commonly reported during psychedelic experiences. While these parallel observations are intriguing, it is important to note that the exact relationship between changes in brain signal complexity and subjective experiences remains an area of ongoing research and debate in the field of neuroscience.

In any case, it is important to note that our Taken task was originally designed to measure inter-subject synchrony rather than to induce a dramatic change in conscious content that could significantly alter the underlying complexity of brain signals. Initially, I thought this approach of using task-based differences in complexity would be a good idea for predicting patient outcomes. The rationale for choosing naturalistic auditory stimuli was based on the premise that such stimuli would require the preservation of higher-order cognitive functions, such as language processing and plot following (Naci et al., 2017). The idea is that these stimuli would perturb brain regions associated with more complex cognitive functions, which could be captured through complexity measures. Capturing such a change was inspired by previous studies that demonstrated the

sensitivity of complexity measures to subtle changes in cognitive states and task engagement. For example, studies have shown that complexity varies depending on the task we are engaged in, such as language processing, motor imagery, and working memory (Omidvarnia et al., 2022), as well as during different states of vigilance (L.C. Shi et al., 2013). Similarly, studies investigating meditation (Aftanas & Golocheikine, 2002) or creative thinking (Mölle et al., 1999) have shown notable changes in complexity.

These considerations highlight the complexity of using task-based differences in EEG complexity measures for prognostic purposes in brain-injured patients. The variability in task performance across different populations and the potential impact of task difficulty on brain signal variability suggest that careful task design and selection are crucial for even considering the prognostic value of these measures. Future research should focus on developing tasks that are sensitive enough to elicit detectable changes in brain complexity across a range of patient populations, while also considering the cognitive capabilities of severely brain-injured individuals.

Part III Discussion

In Part III of my study, I employed machine learning techniques for binary classification of patient outcomes. This approach allowed us to evaluate the prognostic utility of EEG complexity measures on an individual level, building upon the group-level findings from Parts I and II.

Unlike Part I, which used intra-subject machine learning, Part III incorporated all participants into a single machine learning model. During the train-test split, the model was tested on previously unseen participants. This approach avoided data bias and enhanced the algorithm's generalizability, allowing for more robust inferences from the performance metrics.

Our results demonstrated relatively high accuracy, sensitivity, and specificity in predicting individual prognosis using complexity measures. The strong performance of our machine learning models suggests that the complexity measures I employed capture meaningful differences in brain activity between patients with different outcomes. This supports my previous findings from Parts I and II, while also highlighting the limitations of my initial approach in Part I. Although my complexity measures weren't sensitive enough to detect task differences as identified in Part II, my research showed that they can still be prognostic, as evidenced by the group-level differences observed in Part II (see Appendix G). These insights shaped the design of Part III and contributed to its success in demonstrating individual differences.

This finding is particularly significant as it supports the potential clinical application of these measures in acute brain injury cases. The success of this approach in distinguishing between Favorable and Unfavorable outcomes on an individual basis provides strong evidence for the prognostic value of EEG complexity analysis in the ICU setting. The ability to predict individual outcomes using these measures could have significant implications for clinical practice, potentially aiding in early decision-making regarding treatment strategies, resource allocation, and family counselling in acute brain injury cases. However, it is crucial to emphasize that while these results are promising, they should be interpreted cautiously and in conjunction with other clinical indicators.

The sensitivity and specificity scores may not yet be high enough for clinical use. This limitation might stem from relying solely on absolute values of complexity to discriminate between groups. However, Part III's purpose was to explore the informativeness of these measures at the individual level without task perturbation. I found that complexity can inform prognosis, even without utilizing task-based changes or perturbations seen in other studies. This suggests that

individual prediction may be even more promising with future research that builds on my findings, given that absolute values performed relatively well even without task-based differences.

Conversely, my results indicate that complexity is at least valuable as a feature that can be combined with other EEG features, such as sensory-evoked potentials and event-related potentials, for examining prognosis (Formisano et al., 2019). Given EEG's high accessibility in the ICU, neuroimaging exams that use various EEG features can be crucial for prognosis. Future research should also explore how these predictive models can improve alongside other features or with multiple imaging modalities (Kazazian et al., 2021). While the models I present may not be perfect as standalone tools, they show promise for use as features within a larger model that combines various clinical EEG features.

Furthermore, future research should focus on validating these findings in larger, more diverse patient populations and exploring how these predictive models perform across different types of brain injuries. Additionally, investigating the relationship between specific complexity measures and particular aspects of brain function or recovery could provide deeper insights into the neurophysiological basis of these prognostic indicators.

An important aspect of my analysis was the use of AUC calculations, which proved particularly valuable in our context. AUC provides a balanced measure of model performance across all possible classification thresholds, which is especially crucial when dealing with imbalanced datasets common in medical prognostics. This issue of imbalanced data was even highlighted within my own analysis in the LL model (Fig. 11), which had a high sensitivity of 85% but a low specificity of 32%. Unlike individual performance metrics which depend on a specific threshold, AUC evaluates the model's performance across all possible thresholds, making it more suitable for

clinical settings where the optimal threshold may vary depending on the specific use case or patient population.

Theoretical Implications

The relationship between the brain's inherent complexity and its information-processing capacity forms the theoretical foundation for our use of complexity measures in this study (Garrett et al., 2013). My findings, which demonstrate significant differences in EEG complexity between patients with favourable and unfavourable outcomes, align with this theoretical framework. The brain's function as a nonlinear dynamic system, characterized by ever-changing interactions between neurons and neural networks, creates significant variability in brain signals (Garrett et al., 2013). Complexity measures are thought to capture this variability (Costa et al., 2002).

This approach has shown promise in assessing changes in consciousness and brain aging (Keshmiri, 2020), which aligns with my application of these measures in acute brain injury cases. My results, particularly the success of my machine learning models in predicting individual outcomes, further validate the use of complexity measures as a prognostic tool. These findings suggest that our EEG complexity measures are indeed capturing meaningful differences in brain activity between patients with different outcomes, reflecting variations in their brain's information processing capacity and overall functional state (Keshmiri, 2020). This provides a robust framework for evaluating their potential in clinical decision-making for acute brain-injured patients.

Limitations

While my study provides valuable insights into the use of EEG complexity measures for prognostic purposes in acute brain injury cases, it is important to acknowledge several limitations. Firstly, while my machine learning models showed promising results, they would benefit from further validation on larger, independent datasets to ensure their robustness and generalizability across different clinical settings and patient populations. Moreover, the heterogeneity of brain injuries in our patient cohort could have introduced variability in my results, potentially masking subtle differences in complexity measures across different types of injuries.

Another limitation is the potential impact of medication and other medical interventions on EEG complexity measures, which I couldn't fully account for in my analysis. I excluded patients who received sedative medication. However, for some patients, sedation was discontinued just before the scanning session began. Consequently, we're uncertain about the amount of residual medication in their systems and whether it varied across patients. Conversely, I aim to identify measures that aren't overly sensitive to such variations, as this may better reflect real-world clinical scenarios.

Additionally, my study was limited by the lack of long-term follow-up data. Follow-up with many of the patients was based on short-term outcomes as I did not want to exclude too many patients from the analysis. Future research should aim to track patient outcomes over extended periods to better understand the prognostic value of EEG complexity measures in predicting long-term recovery and functional outcomes.

Lastly, determining the true specificity and sensitivity of prognostic tests for patients in the ICU is challenging because life-sustaining measures are withdrawn following a poor prognosis. This creates a "ground truth" problem, as it is impossible to know if a patient might have recovered

if life support had not been withdrawn, potentially leading to a self-fulfilling prophecy where predictions of poor outcomes influence the decision to withdraw life support (Becker et al., 2001). This introduces bias in assessing prognostic tests' effectiveness. Perhaps examining long-term outcomes in contexts where life support withdrawal is less common to understand the impact of such biases.

Concluding Remarks

Complexity has been used in a wide array of interdisciplinary research. Recently, it has also made waves in neuroimaging research and may even one day take over as the gold-standard measure in EEG analysis. However, in its current state, the use of complexity in EEG analysis also comes with a few challenges. As it stands, it may be too susceptible to noise and is quite difficult to calculate and conceptualize. Additionally, due to its recency in the field, the current measures of complexity may not be sensitive enough for our current functional neuroimaging tools and techniques. Importantly, however, many complexity measures are still coming out today. This is promising as newer measures may be able to overcome these challenges and further improve the prognostic abilities of complexity I have already started to demonstrate in this study.

References

Abdalmalak, A., Milej, D., Norton, L., Debicki, D. B., Owen, A. M., & Lawrence, K. S. (2021). The Potential Role of fNIRS in Evaluating Levels of Consciousness. *Frontiers in Human Neuroscience*, *15*. https://doi.org/10.3389/fnhum.2021.703405

Aftanas, L. I., & Golocheikine, S. A. (2002). Non-linear dynamic complexity of the human EEG during meditation. *Neuroscience Letters*, *330*(2), 143–146. https://doi.org/10.1016/s0304-3940(02)00745-0

Alej, Torre-Luque, R. D. L., & Bornas, X. (2017). Complexity and Irregularity in the Brain Oscillations of Depressive Patients: A Systematic Review. *Neuropsychiatry*, *7*(5), 466–477. https://doi.org/10.4172/Neuropsychiatry.1000238

Bai, Y., Lin, Y., & Ziemann, U. (2021). Managing disorders of consciousness: The role of electroencephalography. *Journal of Neurology*, *268*(11), 4033–4065. https://doi.org/10.1007/s00415-020-10095-z

Bartels, A., & Zeki, S. (2004). The neural correlates of maternal and romantic love. *NeuroImage*, *21*(3), 1155–1166. https://doi.org/10.1016/j.neuroimage.2003.11.003

Bartels, A., & Zeki, S. (2005). Brain dynamics during natural viewing conditions—A new guide for mapping connectivity in vivo. *NeuroImage*, *24*(2), 339–349. https://doi.org/10.1016/j.neuroimage.2004.08.044

Bauer, G., Gerstenbrand, F., & Rumpl, E. (1979). Varieties of the locked-in syndrome. *Journal of Neurology*, *221*(2), 77–91. https://doi.org/10.1007/BF00313105

Becker, K. J., Baxter, A. B., Cohen, W. A., Bybee, H. M., Tirschwell, D. L., Newell, D. W., Winn, H. R., & Longstreth, W. T. (2001). Withdrawal of support in intracerebral hemorrhage may

lead to self-fulfilling prophecies. *Neurology*, *56*(6), 766–772. https://doi.org/10.1212/wnl.56.6.766

Bodien, Y. G., Chatelle, C., & Edlow, B. L. (2017). Functional Networks in Disorders of Consciousness. *Seminars in Neurology*, *37*(5), 485–502. https://doi.org/10.1055/s-0037-1607310

Boly, M., Coleman, M. R., Davis, M. H., Hampshire, A., Bor, D., Moonen, G., Maquet, P. A., Pickard, J. D., Laureys, S., & Owen, A. M. (2007). When thoughts become action: An fMRI paradigm to study volitional brain activity in non-communicative brain injured patients. *NeuroImage*, *36*(3), 979–992. https://doi.org/10.1016/j.neuroimage.2007.02.047

Boly, M., Sasai, S., Gosseries, O., Oizumi, M., Casali, A., Massimini, M., & Tononi, G. (2015). Stimulus Set Meaningfulness and Neurophysiological Differentiation: A Functional Magnetic Resonance Imaging Study. *PLoS ONE*, *10*(5). https://doi.org/10.1371/journal.pone.0125337

Boly, M., & Seth, A. K. (2012). Modes and models in disorders of consciousness science. *Archives Italiennes De Biologie*, *150*(2–3), 172–184. https://doi.org/10.4449/aib.v150i2.1372

Bornas, X., Llabrés, J., Morillas-Romero, A., Aguayo-Siquier, B., Balle, M., & Tortella-Feliu, M. (2013). Complexity of everyday life heart rate dynamics and attentional control in healthy students. *Nonlinear Dynamics, Psychology, and Life Sciences*, *17*(3), 345–360.

Callaway, C. W., Schmicker, R. H., Brown, S. P., Albrich, J. M., Andrusiek, D. L., Aufderheide, T. P., Christenson, J., Daya, M. R., Falconer, D., Husa, R. D., Idris, A. H., Ornato, J. P., Rac, V. E., Rea, T. D., Rittenberger, J. C., Sears, G., & Stiell, I. G. (2014). Early Coronary Angiography and Induced Hypothermia are Associated with Survival and Functional

Recovery after Out-of-Hospital Cardiac Arrest. *Resuscitation*, *85*(5), 657–663. https://doi.org/10.1016/j.resuscitation.2013.12.028

Casali, A. G., Gosseries, O., Rosanova, M., Boly, M., Sarasso, S., Casali, K. R., Casarotto, S., Bruno, M.-A., Laureys, S., Tononi, G., & Massimini, M. (2013). A theoretically based index of consciousness independent of sensory processing and behavior. *Science Translational Medicine*, *5*(198), 198ra105. https://doi.org/10.1126/scitranslmed.3006294

Casarotto, S., Comanducci, A., Rosanova, M., Sarasso, S., Fecchio, M., Napolitani, M., Pigorini, A., G Casali, A., Trimarchi, P. D., Boly, M., Gosseries, O., Bodart, O., Curto, F., Landi, C., Mariotti, M., Devalle, G., Laureys, S., Tononi, G., & Massimini, M. (2016). Stratification of unresponsive patients by an independently validated index of brain complexity. *Annals of Neurology*, *80*(5), 718–729. https://doi.org/10.1002/ana.24779

Claassen, J., Doyle, K., Matory, A., Couch, C., Burger, K. M., Velazquez, A., Okonkwo, J. U., King, J.-R., Park, S., Agarwal, S., Roh, D., Megjhani, M., Eliseyev, A., Connolly, E. S., & Rohaut, B. (2019). Detection of Brain Activation in Unresponsive Patients with Acute Brain Injury. *The New England Journal of Medicine*, *380*(26), 2497–2505. https://doi.org/10.1056/NEJMoa1812757

Coleman, M. R., Davis, M. H., Rodd, J. M., Robson, T., Ali, A., Owen, A. M., & Pickard, J. D. (2009). Towards the routine use of brain imaging to aid the clinical diagnosis of disorders of consciousness. *Brain*, *132*(9), 2541–2552. https://doi.org/10.1093/brain/awp183

Coleman, M. R., Rodd, J. M., Davis, M. H., Johnsrude, I. S., Menon, D. K., Pickard, J. D., & Owen, A. M. (2007). Do vegetative patients retain aspects of language comprehension? Evidence from fMRI. *Brain: A Journal of Neurology*, *130*(Pt 10), 2494–2507. https://doi.org/10.1093/brain/awm170

Costa, M., Goldberger, A. L., & Peng, C.-K. (2002). Multiscale entropy analysis of complex physiologic time series. *Physical Review Letters*, *89*(6), 068102. https://doi.org/10.1103/PhysRevLett.89.068102

Costa, M., Goldberger, A. L., & Peng, C.-K. (2005). Multiscale entropy analysis of biological signals. *Physical Review E*, *71*(2), Article 2. https://doi.org/10.1103/PhysRevE.71.021906

Cruse, D., Chennu, S., Chatelle, C., Bekinschtein, T. A., Fernández-Espejo, D., Pickard, J. D., Laureys, S., & Owen, A. M. (2011). Bedside detection of awareness in the vegetative state: A cohort study. *Lancet (London, England)*, *378*(9809), 2088–2094. https://doi.org/10.1016/S0140-6736(11)61224-5

Cruse, D., Chennu, S., Fernández-Espejo, D., Payne, W. L., Young, G. B., & Owen, A. M. (2012). Detecting Awareness in the Vegetative State: Electroencephalographic Evidence for Attempted Movements to Command. *PLOS ONE*, *7*(11), e49933. https://doi.org/10.1371/journal.pone.0049933

Cruse, D., Norton, L., Gofton, T., Young, G. B., & Owen, A. M. (2014). Positive prognostication from median-nerve somatosensory evoked cortical potentials. *Neurocritical Care*, *21*(2), 238–244. https://doi.org/10.1007/s12028-014-9982-y

Damoiseaux, J. S., Rombouts, S. a. R. B., Barkhof, F., Scheltens, P., Stam, C. J., Smith, S. M., & Beckmann, C. F. (2006). Consistent resting-state networks across healthy subjects. *Proceedings of the National Academy of Sciences of the United States of America*, *103*(37), 13848–13853. https://doi.org/10.1073/pnas.0601417103

Dauwels, J., Vialatte, F., Musha, T., & Cichocki, A. (2010). A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG. *NeuroImage*, *49*(1), 668–693. https://doi.org/10.1016/j.neuroimage.2009.06.056

Dávalos, A., Jabloun, M., Ravier, P., & Buttelli, O. (2019). On the Statistical Properties of Multiscale Permutation Entropy: Characterization of the Estimator's Variance. *Entropy (Basel, Switzerland)*, *21*(5), 450. https://doi.org/10.3390/e21050450

Davis, M. H., Coleman, M. R., Absalom, A. R., Rodd, J. M., Johnsrude, I. S., Matta, B. F., Owen, A. M., & Menon, D. K. (2007). Dissociating speech perception and comprehension at reduced levels of awareness. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(41), 16032–16037. https://doi.org/10.1073/pnas.0701309104

Donoghue, T., Dominguez, J., & Voytek, B. (2020). Electrophysiological Frequency Band Ratio Measures Conflate Periodic and Aperiodic Neural Activity. *eNeuro*, *7*(6), ENEURO.0192-20.2020. https://doi.org/10.1523/ENEURO.0192-20.2020

Dueck, M. H., Klimek, M., Appenrodt, S., Weigand, C., & Boerner, U. (2005). Trends but not individual values of central venous oxygen saturation agree with mixed venous oxygen saturation during varying hemodynamic conditions. *Anesthesiology*, *103*(2), 249–257. https://doi.org/10.1097/00000542-200508000-00007

Eagleman, S. L., & MacIver, M. B. (2018). Can you hear me now? Information processing in primary auditory cortex at loss of consciousness. *British Journal of Anaesthesia*, *121*(3), 526–529. https://doi.org/10.1016/j.bja.2018.06.008

Edlow, B. L., Chatelle, C., Spencer, C. A., Chu, C. J., Bodien, Y. G., O'Connor, K. L., Hirschberg, R. E., Hochberg, L. R., Giacino, J. T., Rosenthal, E. S., & Wu, O. (2017). Early detection of consciousness in patients with acute severe traumatic brain injury. *Brain*, *140*(9), 2399–2414. https://doi.org/10.1093/brain/awx176

Edlow, B. L., Claassen, J., Schiff, N. D., & Greer, D. M. (2021). Recovery from disorders of consciousness: Mechanisms, prognosis and emerging therapies. *Nature Reviews Neurology*, *17*(3), 135–156. https://doi.org/10.1038/s41582-020-00428-x

Engemann, D. A., Raimondo, F., King, J.-R., Rohaut, B., Louppe, G., Faugeras, F., Annen, J., Cassol, H., Gosseries, O., Fernandez-Slezak, D., Laureys, S., Naccache, L., Dehaene, S., & Sitt, J. D. (2018). Robust EEG-based cross-site and cross-protocol classification of states of consciousness. *Brain: A Journal of Neurology*, *141*(11), 3179–3192. https://doi.org/10.1093/brain/awy251

Engemann, D., Raimondo, F., King, J.-R., Jas, M., Gramfort, A., Dehaene, S., Naccache, L., & Sitt, J. D. (2015). *Automated Measurement and Prediction of Consciousness in Vegetative and Minimally Conscious Patients*.

Escudero, J., Abásolo, D., Hornero, R., Espino, P., & López, M. (2006). Analysis of electroencephalograms in Alzheimer's disease patients with multiscale entropy. *Physiological Measurement*, *27*(11), 1091–1106. https://doi.org/10.1088/0967-3334/27/11/004

Esteller, R., Echauz, J., Tcheng, T., Litt, B., & Pless, B. (2001). Line length: An efficient feature for seizure onset detection. *2001 Conference Proceedings of the 23rd Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, *2*, 1707–1710 vol.2. https://doi.org/10.1109/IEMBS.2001.1020545

Ferenets, R., Lipping, T., Anier, A., Jäntti, V., Melto, S., & Hovilehto, S. (2006). Comparison of entropy and complexity measures for the assessment of depth of sedation. *IEEE Transactions on Bio-Medical Engineering*, *53*(6), 1067–1077. https://doi.org/10.1109/TBME.2006.873543

Fernández, A., Gómez, C., Hornero, R., & López-Ibor, J. J. (2013). Complexity and schizophrenia. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *45*, 267–276. https://doi.org/10.1016/j.pnpbp.2012.03.015

Fernández-Espejo, D., & Owen, A. M. (2013). Detecting awareness after severe brain injury. *Nature Reviews. Neuroscience*, *14*(11), 801–809. https://doi.org/10.1038/nrn3608

Ferrarelli, F., Massimini, M., Sarasso, S., Casali, A., Riedner, B. A., Angelini, G., Tononi, G., & Pearce, R. A. (2010). Breakdown in cortical effective connectivity during midazolam-induced loss of consciousness. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(6), 2681–2686. https://doi.org/10.1073/pnas.0913008107

Finn, E. S. (2021). Is it time to put rest to rest? *Trends in Cognitive Sciences*, *25*(12), 1021–1032. https://doi.org/10.1016/j.tics.2021.09.005

Formisano, R., Aloisi, M., & Contrada, M. (2022). Chapter 16 - Post-coma syndrome in the context of severe acquired brain injury: Traumatic brain injury and beyond. In R. Rajendram, V. R. Preedy, & C. R. Martin (Eds.), *Diagnosis and Treatment of Traumatic Brain Injury* (pp. 205–219). Academic Press. https://doi.org/10.1016/B978-0-12-823347-4.00029-4

Formisano, R., Giustini, M., Aloisi, M., Contrada, M., Schnakers, C., Zasler, N., & Estraneo, A. (2019). An International survey on diagnostic and prognostic protocols in patients with disorder of consciousness. *Brain Injury*, *33*(8), Article 8. https://doi.org/10.1080/02699052.2019.1622785

Friston, K. J., Holmes, A. P., Worsley, K. J., Poline, J.-P., Frith, C. D., & Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*(4), Article 4. https://doi.org/10.1002/hbm.460020402

Garrett, D. D., Kovacevic, N., McIntosh, A. R., & Grady, C. L. (2010). Blood Oxygen Level-Dependent Signal Variability Is More than Just Noise. *The Journal of Neuroscience*, *30*(14), Article 14. https://doi.org/10.1523/JNEUROSCI.5166-09.2010

Garrett, D. D., Samanez-Larkin, G. R., MacDonald, S. W. S., Lindenberger, U., McIntosh, A. R., & Grady, C. L. (2013). Moment-to-moment brain signal variability: A next frontier in human brain mapping? *Neuroscience and Biobehavioral Reviews*, *37*(4), 610–624. https://doi.org/10.1016/j.neubiorev.2013.02.015

Gell-Mann, M. (1994). *The Quark and the Jaguar: Adventures in the Simple and the Complex*. W.H. Freeman.

Giacino, J. T., Ashwal, S., Childs, N., Cranford, R., Jennett, B., Katz, D. I., Kelly, J. P., Rosenberg, J. H., Whyte, J., Zafonte, R. D., & Zasler, N. D. (2002). The minimally conscious state: Definition and diagnostic criteria. *Neurology*, *58*(3), 349–353. https://doi.org/10.1212/wnl.58.3.349

Giacino, J. T., Fins, J. J., Laureys, S., & Schiff, N. D. (2014). Disorders of consciousness after acquired brain injury: The state of the science. *Nature Reviews. Neurology*, *10*(2), 99–114. https://doi.org/10.1038/nrneurol.2013.279

Giacino, J. T., & Kalmar, K. (1997). The Vegetative and Minimally Conscious States: A Comparison of Clinical Features and Functional Outcome. *The Journal of Head Trauma Rehabilitation*, *12*(4), 36.

Giacino, J. T., Katz, D. I., Schiff, N. D., Whyte, J., Ashman, E. J., Ashwal, S., Barbano, R., Hammond, F. M., Laureys, S., Ling, G. S. F., Nakase-Richardson, R., Seel, R. T., Yablon, S., Getchius, T. S. D., Gronseth, G. S., & Armstrong, M. J. (2018). Practice guideline update

recommendations summary: Disorders of consciousness. *Neurology*, *91*(10), Article 10. https://doi.org/10.1212/WNL.0000000000005926

Giacino, J. T., & Zasler, N. D. (1995). Outcome after severe traumatic brain injury: Coma, the vegetative state, and the minimally responsive state. *The Journal of Head Trauma Rehabilitation*, *10*(1), 41.

Gill-Thwaites, H. (2006). Lotteries, loopholes and luck: Misdiagnosis in the vegetative state patient. *Brain Injury*, *20*(13–14), 1321–1328. https://doi.org/10.1080/02699050601081802

Gofton, T. E., Chouinard, P. A., Young, G. B., Bihari, F., Nicolle, M. W., Lee, D. H., Sharpe, M. D., Yen, Y.-F., Takahashi, A. M., & Mirsattari, S. M. (2009). Functional MRI study of the primary somatosensory cortex in comatose survivors of cardiac arrest. *Experimental Neurology*, *217*(2), 320–327. https://doi.org/10.1016/j.expneurol.2009.03.011

Goldberger, A. L., Amaral, L. A. N., Hausdorff, J. M., Ivanov, P. Ch., Peng, C.-K., & Stanley, H. E. (2002). Fractal dynamics in physiology: Alterations with disease and aging. *Proceedings of the National Academy of Sciences*, *99*(suppl_1), 2466–2472. https://doi.org/10.1073/pnas.012579499

Goldberger, A. L., Peng, C.-K., & Lipsitz, L. A. (2002). What is physiologic complexity and how does it change with aging and disease? *Neurobiology of Aging*, *23*(1), 23–26. https://doi.org/10.1016/s0197-4580(01)00266-4

Grassberger, P. (1989). Information content and predictability of lumped and distributed dynamical systems. *Physica Scripta*, *40*(3), 346. https://doi.org/10.1088/0031-8949/40/3/016

Grassberger, P. (1991). Information and Complexity Measures in Dynamical Systems. In H. Atmanspacher & H. Scheingraber (Eds.), *Information Dynamics* (pp. 15–33). Springer US. https://doi.org/10.1007/978-1-4899-2305-9_2

Green, S. M. (2011). Cheerio, laddie! Bidding farewell to the Glasgow Coma Scale. *Annals of Emergency Medicine*, *58*(5), 427–430. https://doi.org/10.1016/j.annemergmed.2011.06.009

Green, T., Rosen, S., Faulkner, A., & Paterson, R. (2013). Adaptation to spectrally-rotated speech. *The Journal of the Acoustical Society of America*, *134*(2), 1369–1377. https://doi.org/10.1121/1.4812759

Gu, C., Liu, Z.-X., & Woltering, S. (2022). Electroencephalography complexity in resting and task states in adults with attention-deficit/hyperactivity disorder. *Brain Communications*, *4*(2), fcac054. https://doi.org/10.1093/braincomms/fcac054

Hardstone, R., Poil, S.-S., Schiavone, G., Jansen, R., Nikulin, V. V., Mansvelder, H. D., & Linkenkaer-Hansen, K. (2012). Detrended Fluctuation Analysis: A Scale-Free View on Neuronal Oscillations. *Frontiers in Physiology*, *3*, 450. https://doi.org/10.3389/fphys.2012.00450

Hasson, U., Malach, R., & Heeger, D. J. (2010). Reliability of cortical activity during natural stimulation. *Trends in Cognitive Sciences*, *14*(1), 40. https://doi.org/10.1016/j.tics.2009.10.011

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., & Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science (New York, N.Y.)*, *303*(5664), 1634–1640. https://doi.org/10.1126/science.1089506

Hasson, U., Yang, E., Vallines, I., Heeger, D. J., & Rubin, N. (2008). A hierarchy of temporal receptive windows in human cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *28*(10), 2539–2550. https://doi.org/10.1523/JNEUROSCI.5487-07.2008

He, B. J. (2011). Scale-Free Properties of the Functional Magnetic Resonance Imaging Signal during Rest and Task. *Journal of Neuroscience*, *31*(39), Article 39. https://doi.org/10.1523/JNEUROSCI.2111-11.2011

Heinke, W., Fiebach, C. J., Schwarzbauer, C., Meyer, M., Olthoff, D., & Alter, K. (2004). Sequential effects of propofol on functional brain activation induced by auditory language processing: An event-related functional magnetic resonance imaging study. *BJA: British Journal of Anaesthesia*, *92*(5), 641–650. https://doi.org/10.1093/bja/aeh133

Hoptman, M. J., Zuo, X.-N., Butler, P. D., Javitt, D. C., D'Angelo, D., Mauro, C. J., & Milham, M. P. (2010). Amplitude of low-frequency oscillations in schizophrenia: A resting state fMRI study. *Schizophrenia Research*, *117*(1), 13–20. https://doi.org/10.1016/j.schres.2009.09.030

Huberman, B. A., & Hogg, T. (1986). Complexity and Adaptation. *Physica D: Nonlinear Phenomena*, *22*(1), 376–384. https://doi.org/10.1016/0167-2789(86)90308-1

Ibáñez-Molina, A. J., & Iglesias-Parro, S. (2014). Fractal characterization of internally and externally generated conscious experiences. *Brain and Cognition*, *87*, 69–75. https://doi.org/10.1016/j.bandc.2014.03.002

Jennett, B., & Bond, M. (1975). Assessment of outcome after severe brain damage. *Lancet (London, England)*, *1*(7905), 480–484. https://doi.org/10.1016/s0140-6736(75)92830-5

Jennett, B., & Plum, F. (1972). Persistent vegetative state after brain damage. A syndrome in search of a name. *Lancet (London, England)*, *1*(7753), 734–737. https://doi.org/10.1016/s0140-6736(72)90242-5

Jeong, J. (2002). Nonlinear dynamics of EEG in Alzheimer's disease. *Drug Development Research*, *56*(2), 57–66. https://doi.org/10.1002/ddr.10061

Johnson, C. W. (2006). What are emergent properties and how do they affect the engineering of complex systems? *Reliability Engineering & System Safety*, *91*(12), 1475–1481. https://doi.org/10.1016/j.ress.2006.01.008

Katz, M. J. (1988). Fractals and the analysis of waveforms. *Computers in Biology and Medicine*, *18*(3), 145–156. https://doi.org/10.1016/0010-4825(88)90041-8

Kazazian, K., Norton, L., Gofton, T. E., Debicki, D., & Owen, A. M. (2020). Cortical Function in Acute Severe Traumatic Brain Injury and at Recovery: A Longitudinal fMRI Case Study. *Brain Sciences*, *10*(9), 604. https://doi.org/10.3390/brainsci10090604

Kazazian, K., Norton, L., Laforge, G., Abdalmalak, A., Gofton, T., Debicki, D., Slessarev, M., Hollywood, S., St Lawrence, K., & Owen, A. (2021). Improving Diagnosis and Prognosis in Acute Severe Brain Injury: A Multimodal Imaging Protocol. *Frontiers in Neurology*, *12*. https://doi.org/10.3389/fneur.2021.757219

Kerssens, C., Hamann, S., Peltier, S., Hu, X. P., Byas-Smith, M. G., & Sebel, P. S. (2005). Attenuated Brain Response to Auditory Word Stimulation with Sevoflurane: A Functional Magnetic Resonance Imaging Study in Humans. *Anesthesiology*, *103*(1), 11–19. https://doi.org/10.1097/00000542-200507000-00006

Keshmiri, S. (2020). Entropy and the Brain: An Overview. *Entropy*, *22*(9), Article 9. https://doi.org/10.3390/e22090917

Koch, C., Massimini, M., Boly, M., & Tononi, G. (2016). Neural correlates of consciousness: Progress and problems. *Nature Reviews Neuroscience*, *17*(5), 307–321. https://doi.org/10.1038/nrn.2016.22

Koenig, M. A., Holt, J. L., Ernst, T., Buchthal, S. D., Nakagawa, K., Stenger, V. A., & Chang, L. (2014). MRI Default Mode Network Connectivity is Associated with Functional Outcome

after Cardiopulmonary Arrest. *Neurocritical Care*, *20*(3), 348–357. https://doi.org/10.1007/s12028-014-9953-3

Kondziella, D., Bender, A., Diserens, K., van Erp, W., Estraneo, A., Formisano, R., Laureys, S., Naccache, L., Ozturk, S., Rohaut, B., Sitt, J. D., Stender, J., Tiainen, M., Rossetti, A. O., Gosseries, O., Chatelle, C., & EAN Panel on Coma, Disorders of Consciousness. (2020). European Academy of Neurology guideline on the diagnosis of coma and other disorders of consciousness. *European Journal of Neurology*, *27*(5), 741–756. https://doi.org/10.1111/ene.14151

Kondziella, D., Friberg, C. K., Frokjaer, V. G., Fabricius, M., & Møller, K. (2016). Preserved consciousness in vegetative and minimal conscious states: Systematic review and meta-analysis. *Journal of Neurology, Neurosurgery, and Psychiatry*, *87*(5), 485–492. https://doi.org/10.1136/jnnp-2015-310958

Laforge, G. (2017). Identifying Electrophysiological Components of Covert Awareness in Patients with Disorders of Consciousness. *Electronic Thesis and Dissertation Repository*. https://ir.lib.uwo.ca/etd/4883

Laforge, G. (2022). Movies on the Mind: Using Naturalistic Stimuli to Assess Perception, Cognition, and Awareness in Patients with Disorders of Consciousness. *Electronic Thesis and Dissertation Repository*. https://ir.lib.uwo.ca/etd/8427

Laforge, G., Gonzalez-Lara, L. E., Owen, A. M., & Stojanoski, B. (2020). Individualized assessment of residual cognition in patients with disorders of consciousness. *NeuroImage: Clinical*, *28*, 102472. https://doi.org/10.1016/j.nicl.2020.102472

Lau, Z. J., Pham, T., Chen, S. H. A., & Makowski, D. (2022). Brain entropy, fractal dimensions and predictability: A review of complexity measures for EEG in healthy and

neuropsychiatric populations. *European Journal of Neuroscience*, *56*(7), Article 7. https://doi.org/10.1111/ejn.15800

Laureys, S. (2005). The neural correlate of (un)awareness: Lessons from the vegetative state. *Trends in Cognitive Sciences*, *9*(12), 556–559. https://doi.org/10.1016/j.tics.2005.10.010

Laureys, S., Pellas, F., Van Eeckhout, P., Ghorbel, S., Schnakers, C., Perrin, F., Berré, J., Faymonville, M.-E., Pantke, K.-H., Damas, F., Lamy, M., Moonen, G., & Goldman, S. (2005). The locked-in syndrome: What is it like to be conscious but paralyzed and voiceless? *Progress in Brain Research*, *150*, 495–511. https://doi.org/10.1016/S0079-6123(05)50034-7

Lebedev, A. v., Kaelen, M., Lövdén, M., Nilsson, J., Feilding, A., Nutt, D. j., & Carhart-Harris, R. l. (2016). LSD-induced entropic brain activity predicts subsequent personality change. *Human Brain Mapping*, *37*(9), 3203–3213. https://doi.org/10.1002/hbm.23234

Lei, L., Liu, K., Yang, Y., Doubliez, A., Hu, X., Xu, Y., & Zhou, Y. (2022). Spatio-temporal analysis of EEG features during consciousness recovery in patients with disorders of consciousness. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *133*, 135–144. https://doi.org/10.1016/j.clinph.2021.08.027

Lempel, A., & Ziv, J. (1976). On the Complexity of Finite Sequences. *IEEE Transactions on Information Theory*, *22*(1), Article 1. IEEE Transactions on Information Theory. https://doi.org/10.1109/TIT.1976.1055501

Li, J., Shen, J., Liu, S., Chauvel, M., Yang, W., Mei, J., Lei, L., Wu, L., Gao, J., & Yang, Y. (2018). Responses of Patients with Disorders of Consciousness to Habit Stimulation: A Quantitative EEG Study. *Neuroscience Bulletin*, *34*(4), 691–699. https://doi.org/10.1007/s12264-018-0258-y

Lipsitz, L. A., & Goldberger, A. L. (1992). Loss of "complexity" and aging. Potential applications of fractals and chaos theory to senescence. *JAMA*, *267*(13), 1806–1809.

Lloyd, S. (2001). Measures of complexity: A nonexhaustive list. *IEEE Control Systems Magazine*, *21*(4), 7–8. IEEE Control Systems Magazine. https://doi.org/10.1109/MCS.2001.939938

Luppi, A. I., Craig, M. M., Pappas, I., Finoia, P., Williams, G. B., Allanson, J., Pickard, J. D., Owen, A. M., Naci, L., Menon, D. K., & Stamatakis, E. A. (2019). Consciousness-specific dynamic interactions of brain integration and functional diversity. *Nature Communications*, *10*(1), 4616. https://doi.org/10.1038/s41467-019-12658-9

Ma, Y., Shi, W., Peng, C.-K., & Yang, A. C. (2018). Nonlinear dynamical analysis of sleep electroencephalography using fractal and entropy approaches. *Sleep Medicine Reviews*, *37*, 85–93. https://doi.org/10.1016/j.smrv.2017.01.003

Makowski, D., Te, A. S., Pham, T., Lau, Z. J., & Chen, S. H. A. (2022). The Structure of Chaos: An Empirical Comparison of Fractal Physiology Complexity Indices Using NeuroKit2. *Entropy*, *24*(8), Article 8. https://doi.org/10.3390/e24081036

Manis, G., Aktaruzzaman, M., & Sassi, R. (2017). Bubble Entropy: An Entropy Almost Free of Parameters. *IEEE Transactions on Bio-Medical Engineering*, *64*(11), 2711–2718. https://doi.org/10.1109/TBME.2017.2664105

Massimini, M., Boly, M., Casali, A., Rosanova, M., & Tononi, G. (2009). *A perturbational approach for evaluating the brain's capacity for consciousness*. *177*, 201–214. https://doi.org/10.1016/S0079-6123(09)17714-2

Massimini, M., Ferrarelli, F., Huber, R., Esser, S. K., Singh, H., & Tononi, G. (2005). Breakdown of cortical effective connectivity during sleep. *Science (New York, N.Y.)*, *309*(5744), 2228–2232. https://doi.org/10.1126/science.1117256

Massimini, M., Ferrarelli, F., Murphy, M., Huber, R., Riedner, B., Casarotto, S., & Tononi, G. (2010). Cortical reactivity and effective connectivity during REM sleep in humans. *Cognitive Neuroscience*, *1*(3), 176–183. https://doi.org/10.1080/17588921003731578

Mayer, S. A., & Kossoff, S. B. (1999). Withdrawal of life support in the neurological intensive care unit. *Neurology*, *52*(8), 1602–1609. https://doi.org/10.1212/wnl.52.8.1602

Mitchell, S. D. (2009). *Unsimple Truths: Science, Complexity, and Policy*. University of Chicago Press.

Mizuno, T., Takahashi, T., Cho, R. Y., Kikuchi, M., Murata, T., Takahashi, K., & Wada, Y. (2010). Assessment of EEG dynamical complexity in Alzheimer's disease using multiscale entropy. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *121*(9), 1438–1446. https://doi.org/10.1016/j.clinph.2010.03.025

Mölle, M., Marshall, L., Wolf, B., Fehm, H. L., & Born, J. (1999). EEG complexity and performance measures of creative thinking. *Psychophysiology*, *36*(1), 95–104. https://doi.org/10.1017/s0048577299961619

Monti, M. M., Vanhaudenhuyse, A., Coleman, M. R., Boly, M., Pickard, J. D., Tshibanda, L., Owen, A. M., & Laureys, S. (2010). Willful modulation of brain activity in disorders of consciousness. *The New England Journal of Medicine*, *362*(7), 579–589. https://doi.org/10.1056/NEJMoa0905370

Morgan, B. L. (2023). Introduction to Neuro Assessment in Critical Care: Part 1 in a Series. *The Canadian Journal of Critical Care Nursing*, *34*(3), 25–30.

Moritz, C. H., Rowley, H. A., Haughton, V. M., Swartz, K. R., Jones, J., & Badie, B. (2001). Functional MR imaging assessment of a non-responsive brain injured patient. *Magnetic Resonance Imaging*, *19*(8), 1129–1132. https://doi.org/10.1016/s0730-725x(01)00432-5

Morlet, D., & Fischer, C. (2014). MMN and novelty P3 in coma and other altered states of consciousness: A review. *Brain Topography*, *27*(4), 467–479. https://doi.org/10.1007/s10548-013-0335-5

Naci, L., Cusack, R., Anello, M., & Owen, A. M. (2014). A common neural code for similar conscious experiences in different individuals. *Proceedings of the National Academy of Sciences*, *111*(39), Article 39. https://doi.org/10.1073/pnas.1407007111

Naci, L., & Owen, A. M. (2013). Making every word count for nonresponsive patients. *JAMA Neurology*, *70*(10), 1235–1241. https://doi.org/10.1001/jamaneurol.2013.3686

Naci, L., Sinai, L., & Owen, A. M. (2017). Detecting and interpreting conscious experiences in behaviorally non-responsive patients. *NeuroImage*, *145*, 304–313. https://doi.org/10.1016/j.neuroimage.2015.11.059

Norton, L. (2017). Functional Magnetic Resonance Imaging as an Assessment Tool in Critically Ill Patients. *Electronic Thesis and Dissertation Repository*. https://ir.lib.uwo.ca/etd/4812

Norton, L., Hutchison, R. M., Young, G. B., Lee, D. H., Sharpe, M. D., & Mirsattari, S. M. (2012). Disruptions of functional connectivity in the default mode network of comatose patients. *Neurology*, *78*(3), 175–181. https://doi.org/10.1212/WNL.0b013e31823fcd61

Omidvarnia, A., Liégeois, R., Amico, E., Preti, M. G., Zalesky, A., & Van De Ville, D. (2022). On the Spatial Distribution of Temporal Complexity in Resting State and Task Functional MRI. *Entropy (Basel, Switzerland)*, *24*(8), 1148. https://doi.org/10.3390/e24081148

Owen, A. M., Coleman, M. R., Boly, M., Davis, M. H., Laureys, S., & Pickard, J. D. (2006). Detecting awareness in the vegetative state. *Science (New York, N.Y.)*, *313*(5792), 1402. https://doi.org/10.1126/science.1130197

Park, J.-H., Kim, S., Kim, C.-H., Cichocki, A., & Kim, K. (2007). Multiscale entropy analysis of eeg from patients under different pathological conditions. *Fractals*, *15*(04), 399–404. https://doi.org/10.1142/S0218348X07003691

Peng, C.-K., Buldyrev, S. V., Havlin, S., Simons, M., Stanley, H. E., & Goldberger, A. L. (1994). Mosaic organization of DNA nucleotides. *Physical Review E*, *49*(2), 1685–1689. https://doi.org/10.1103/PhysRevE.49.1685

Petri, G., Expert, P., Turkheimer, F., Carhart-Harris, R., Nutt, D., Hellyer, P. J., & Vaccarino, F. (2014). Homological scaffolds of brain functional networks. *Journal of The Royal Society Interface*, *11*(101), 20140873. https://doi.org/10.1098/rsif.2014.0873

Pincus, S. M. (1991). Approximate entropy as a measure of system complexity. *Proceedings of the National Academy of Sciences*, *88*(6), 2297–2301. https://doi.org/10.1073/pnas.88.6.2297

Plourde, G., Belin, P., Chartrand, D., Fiset, P., Backman, S. B., Xie, G., & Zatorre, R. J. (2006). Cortical Processing of Complex Auditory Stimuli during Alterations of Consciousness with the General Anesthetic Propofol. *Anesthesiology*, *104*(3), 448–457. https://doi.org/10.1097/00000542-200603000-00011

Posner, J. B., Saper, C. B., Schiff, N. D., & Claassen, J. (2019). *Plum and Posner's Diagnosis and Treatment of Stupor and Coma*. https://academic.oup.com/book/29503

Pugin, D., Hofmeister, J., Gasche, Y., Vulliemoz, S., Lövblad, K.-O., Ville, D. V. D., & Haller, S. (2020). Resting-State Brain Activity for Early Prediction Outcome in Postanoxic Patients in a Coma with Indeterminate Clinical Prognosis. *American Journal of Neuroradiology*, *41*(6), 1022–1030. https://doi.org/10.3174/ajnr.A6572

Richman, J. S., & Moorman, J. R. (2000). Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology. Heart and Circulatory Physiology*, *278*(6), H2039-2049. https://doi.org/10.1152/ajpheart.2000.278.6.H2039

Roberts, S. J., Penny, W., & Rezek, I. (1999). Temporal and spatial complexity measures for electroencephalogram based brain-computer interfacing. *Medical & Biological Engineering & Computing*, *37*(1), 93–98. https://doi.org/10.1007/BF02513272

Rosanova, M., Fecchio, M., Casarotto, S., Sarasso, S., Casali, A. G., Pigorini, A., Comanducci, A., Seregni, F., Devalle, G., Citerio, G., Bodart, O., Boly, M., Gosseries, O., Laureys, S., & Massimini, M. (2018). Sleep-like cortical OFF-periods disrupt causality and complexity in the brain of unresponsive wakefulness syndrome patients. *Nature Communications*, *9*(1), 4427. https://doi.org/10.1038/s41467-018-06871-1

Rosanova, M., Gosseries, O., Casarotto, S., Boly, M., Casali, A. G., Bruno, M.-A., Mariotti, M., Boveroux, P., Tononi, G., Laureys, S., & Massimini, M. (2012). Recovery of cortical effective connectivity and recovery of consciousness in vegetative patients. *Brain: A Journal of Neurology*, *135*(Pt 4), 1308–1320. https://doi.org/10.1093/brain/awr340

Sair, H. I., Hannawi, Y., Li, S., Kornbluth, J., Demertzi, A., Di Perri, C., Chabanne, R., Jean, B., Benali, H., Perlbarg, V., Pekar, J., Luyt, C.-E., Galanaud, D., Velly, L., Puybasset, L., Laureys, S., Caffo, B., Stevens, R. D., & Neuroimaging for Coma Emergence and Recovery (NICER) Consortium. (2018). Early Functional Connectome Integrity and 1-Year Recovery in Comatose Survivors of Cardiac Arrest. *Radiology*, *287*(1), 247–255. https://doi.org/10.1148/radiol.2017162161

Sarasso, S., Boly, M., Napolitani, M., Gosseries, O., Charland-Verville, V., Casarotto, S., Rosanova, M., Casali, A. G., Brichant, J.-F., Boveroux, P., Rex, S., Tononi, G., Laureys, S.,

& Massimini, M. (2015). Consciousness and Complexity during Unresponsiveness Induced by Propofol, Xenon, and Ketamine. *Current Biology: CB*, *25*(23), 3099–3105. https://doi.org/10.1016/j.cub.2015.10.014

Sarasso, S., Casali, A. G., Casarotto, S., Rosanova, M., Sinigaglia, C., & Massimini, M. (2021). Consciousness and complexity: A consilience of evidence. *Neuroscience of Consciousness*, niab023. https://doi.org/10.1093/nc/niab023

Sauseng, P., Griesmayr, B., Freunberger, R., & Klimesch, W. (2010). Control mechanisms in working memory: A possible function of EEG theta oscillations. *Neuroscience and Biobehavioral Reviews*, *34*(7), 1015–1022. https://doi.org/10.1016/j.neubiorev.2009.12.006

Schartner, M. M., Carhart-Harris, R. L., Barrett, A. B., Seth, A. K., & Muthukumaraswamy, S. D. (2017). Increased spontaneous MEG signal diversity for psychoactive doses of ketamine, LSD and psilocybin. *Scientific Reports*, *7*(1), 46421. https://doi.org/10.1038/srep46421

Schartner, M., Seth, A., Noirhomme, Q., Boly, M., Bruno, M.-A., Laureys, S., & Barrett, A. (2015). Complexity of Multi-Dimensional Spontaneous EEG Decreases during Propofol Induced General Anaesthesia. *PloS One*, *10*(8), e0133532. https://doi.org/10.1371/journal.pone.0133532

Schiff, N. D. (2015). Cognitive Motor Dissociation Following Severe Brain Injuries. *JAMA Neurology*, *72*(12), 1413–1415. https://doi.org/10.1001/jamaneurol.2015.2899

Shalizi, C. R. (2006). Methods and Techniques of Complex Systems Science: An Overview. In T. S. Deisboeck & J. Y. Kresh (Eds.), *Complex Systems Science in Biomedicine* (pp. 33–114). Springer US. https://doi.org/10.1007/978-0-387-33532-2_2

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*(3), 379–423. The Bell System Technical Journal. https://doi.org/10.1002/j.1538-7305.1948.tb01338.x

Shi, L.-C., Jiao, Y.-Y., & Lu, B.-L. (2013). Differential entropy feature for EEG-based vigilance estimation. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, *2013*, 6627–6630. https://doi.org/10.1109/EMBC.2013.6611075

Shi, W., Shang, P., Ma, Y., Sun, S., & Yeh, C.-H. (2017). A comparison study on stages of sleep: Quantifying multiscale complexity using higher moments on coarse-graining. *Communications in Nonlinear Science and Numerical Simulation*, *44*, 292–303. https://doi.org/10.1016/j.cnsns.2016.08.019

Silva, S., de Pasquale, F., Vuillaume, C., Riu, B., Loubinoux, I., Geeraerts, T., Seguin, T., Bounes, V., Fourcade, O., Demonet, J.-F., & Péran, P. (2015). Disruption of posteromedial large-scale neural communication predicts recovery from coma. *Neurology*, *85*(23), 2036–2044. https://doi.org/10.1212/WNL.0000000000002196

Sitt, J. D., King, J.-R., El Karoui, I., Rohaut, B., Faugeras, F., Gramfort, A., Cohen, L., Sigman, M., Dehaene, S., & Naccache, L. (2014). Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state. *Brain: A Journal of Neurology*, *137*(Pt 8), 2258–2270. https://doi.org/10.1093/brain/awu141

Soddu, A., Vanhaudenhuyse, A., Demertzi, A., Bruno, M.-A., Tshibanda, L., Di, H., Boly, M., Papa, M., Laureys, S., & Noirhomme, Q. (2011). Resting state activity in patients with disorders of consciousness. *Functional Neurology*, *26*(1), 37–43.

Sokoliuk, R., Degano, G., Banellis, L., Melloni, L., Hayton, T., Sturman, S., Veenith, T., Yakoub, K. M., Belli, A., Noppeney, U., & Cruse, D. (2021). Covert Speech Comprehension Predicts Recovery From Acute Unresponsive States. *Annals of Neurology*, *89*(4), 646–656. https://doi.org/10.1002/ana.25995

Solhjoo, S., Motie Nasrabadi, A., & Hashemi Golpayegani, M. R. (2005). EEG-Based Mental Task Classification in Hypnotized and Normal Subjects. *Conference Proceedings: ... Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual Conference*, *2005*, 2041–2043. https://doi.org/10.1109/IEMBS.2005.1616858

Song, M., Zhang, Y., Cui, Y., Yang, Y., & Jiang, T. (2018). Brain Network Studies in Chronic Disorders of Consciousness: Advances and Perspectives. *Neuroscience Bulletin*, *34*(4), 592–604. https://doi.org/10.1007/s12264-018-0243-5

Spiers, H. J., & Maguire, E. A. (2007). Decoding human brain activity during real-world experiences. *Trends in Cognitive Sciences*, *11*(8), 356–365. https://doi.org/10.1016/j.tics.2007.06.002

Stam, C. J. (2005). Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field. *Clinical Neurophysiology*, *116*(10), Article 10. https://doi.org/10.1016/j.clinph.2005.06.011

Stam, C. J., de Haan, W., Daffertshofer, A., Jones, B. F., Manshanden, I., van Cappellen van Walsum, A. M., Montez, T., Verbunt, J. P. A., de Munck, J. C., van Dijk, B. W., Berendse, H. W., & Scheltens, P. (2009). Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease. *Brain: A Journal of Neurology*, *132*(Pt 1), 213–224. https://doi.org/10.1093/brain/awn262

Stam, C. J., van Cappellen van Walsum, A.-M., & Micheloyannis, S. (2002). Variability of EEG synchronization during a working memory task in healthy subjects. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *46*(1), 53–66. https://doi.org/10.1016/s0167-8760(02)00041-7

Stender, J., Gosseries, O., Bruno, M.-A., Charland-Verville, V., Vanhaudenhuyse, A., Demertzi, A., Chatelle, C., Thonnard, M., Thibaut, A., Heine, L., Soddu, A., Boly, M., Schnakers, C., Gjedde, A., & Laureys, S. (2014). Diagnostic precision of PET imaging and functional MRI in disorders of consciousness: A clinical validation study. *Lancet (London, England)*, *384*(9942), 514–522. https://doi.org/10.1016/S0140-6736(14)60042-8

Stevens, R. D., & Sutter, R. (2013). Prognosis in severe brain injury. *Critical Care Medicine*, *41*(4), 1104–1123. https://doi.org/10.1097/CCM.0b013e318287ee79

Tagliazucchi, E., Carhart-Harris, R., Leech, R., Nutt, D., & Chialvo, D. R. (2014). Enhanced repertoire of brain dynamical states during the psychedelic experience. *Human Brain Mapping*, *35*(11), 5442–5456. https://doi.org/10.1002/hbm.22562

Takahashi, T. (2013). Complexity of spontaneous brain activity in mental disorders. *Progress in Neuro-Psychopharmacology & Biological Psychiatry*, *45*, 258–266. https://doi.org/10.1016/j.pnpbp.2012.05.001

Takahashi, T., Cho, R. Y., Murata, T., Mizuno, T., Kikuchi, M., Mizukami, K., Kosaka, H., Takahashi, K., & Wada, Y. (2009). Age-related variation in EEG complexity to photic stimulation: A multiscale entropy analysis. *Clinical Neurophysiology*, *120*(3), Article 3. https://doi.org/10.1016/j.clinph.2008.12.043

Teasdale, G., & Jennett, B. (1974). ASSESSMENT OF COMA AND IMPAIRED CONSCIOUSNESS: A Practical Scale. *The Lancet*, *304*(7872), 81–84. https://doi.org/10.1016/S0140-6736(74)91639-0

Teasdale, G., Maas, A., Lecky, F., Manley, G., Stocchetti, N., & Murray, G. (2014). The Glasgow Coma Scale at 40 years: Standing the test of time. *The Lancet. Neurology*, *13*(8), 844–854. https://doi.org/10.1016/S1474-4422(14)70120-6

Tononi, G., & Edelman, G. M. (1998). Consciousness and Complexity. *Science*, *282*(5395), Article 5395. https://doi.org/10.1126/science.282.5395.1846

Tosun, P., Abásolo, D., Stenson, G., & Winsky-Sommerer, R. (2017). Characterisation of the Effects of Sleep Deprivation on the Electroencephalogram Using Permutation Lempel–Ziv Complexity, a Non-Linear Analysis Tool. *Entropy*, *19*(12), Article 12. https://doi.org/10.3390/e19120673

Turgeon, A. F., Lauzier, F., Burns, K. E. A., Meade, M. O., Scales, D. C., Zarychanski, R., Moore, L., Zygun, D. A., McIntyre, L. A., Kanji, S., Hébert, P. C., Murat, V., Pagliarello, G., Fergusson, D. A., & Canadian Critical Care Trials Group. (2013). Determination of neurologic prognosis and clinical decision making in adult patients with severe traumatic brain injury: A survey of Canadian intensivists, neurosurgeons, and neurologists. *Critical Care Medicine*, *41*(4), 1086–1093. https://doi.org/10.1097/CCM.0b013e318275d046

Turgeon, A. F., Lauzier, F., Simard, J.-F., Scales, D. C., Burns, K. E. A., Moore, L., Zygun, D. A., Bernard, F., Meade, M. O., Dung, T. C., Ratnapalan, M., Todd, S., Harlock, J., Fergusson, D. A., & Canadian Critical Care Trials Group. (2011). Mortality associated with withdrawal of life-sustaining therapy for patients with severe traumatic brain injury: A Canadian multicentre cohort study. *CMAJ: Canadian Medical Association Journal = Journal de*

l'Association Medicale Canadienne, *183*(14), 1581–1588. https://doi.org/10.1503/cmaj.101786

Unakafov, A. M., & Keller, K. (2014). *Conditional entropy of ordinal patterns* (arXiv:1407.5390). arXiv. https://doi.org/10.48550/arXiv.1407.5390

Vanhaudenhuyse, A., Laureys, S., & Perrin, F. (2008). Cognitive Event-Related Potentials in Comatose and Post-Comatose States. *Neurocritical Care*, *8*(2), 262–270. https://doi.org/10.1007/s12028-007-9016-0

Viol, A., Palhano-Fontes, F., Onias, H., de Araujo, D. B., & Viswanathan, G. M. (2017). Shannon entropy of brain functional complex networks under the influence of the psychedelic Ayahuasca. *Scientific Reports*, *7*(1), 7388. https://doi.org/10.1038/s41598-017-06854-0

Wang, X., & Ding, M. (2011). Relation between P300 and event-related theta-band synchronization: A single-trial analysis. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *122*(5), 916–924. https://doi.org/10.1016/j.clinph.2010.09.011

Weijer, C., Bruni, T., Gofton, T., Young, G. B., Norton, L., Peterson, A., & Owen, A. M. (2016). Ethical considerations in functional magnetic resonance imaging research in acutely comatose patients. *Brain: A Journal of Neurology*, *139*(Pt 1), 292–299. https://doi.org/10.1093/brain/awv272

Whyte, J., Cifu, D., Dikmen, S., & Temkin, N. (2001). Prediction of functional outcomes after traumatic brain injury: A comparison of 2 measures of duration of unconsciousness. *Archives of Physical Medicine and Rehabilitation*, *82*(10), 1355–1359. https://doi.org/10.1053/apmr.2001.26091

Wilson, J. T., Pettigrew, L. E., & Teasdale, G. M. (1998). Structured interviews for the Glasgow Outcome Scale and the extended Glasgow Outcome Scale: Guidelines for their use. *Journal of Neurotrauma*, *15*(8), 573–585. https://doi.org/10.1089/neu.1998.15.573

Wutte, M. G., Smith, M. T., Flanagin, V. L., & Wolbers, T. (2011). Physiological Signal Variability in hMT+ Reflects Performance on a Direction Discrimination Task. *Frontiers in Psychology*, *2*, 185. https://doi.org/10.3389/fpsyg.2011.00185

Yang, A. C., & Tsai, S.-J. (2013). Is mental illness complex? From behavior to brain. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *45*, 253–257. https://doi.org/10.1016/j.pnpbp.2012.09.015

Yi, L. I., Xiao-ping, L. I. U., Xian-hong, L., Jing-qi, L. I., Wen-wei, Y., Dan-ke, Z., Li-hua, L. I., & Yong, Y. (2014). Mapping Brain Injury with Symmetrical-channels' EEG Signal Analysis – A Pilot Study. *Scientific Reports*, *4*(1), 5023. https://doi.org/10.1038/srep05023

Young, G. B. (2009). Coma. *Annals of the New York Academy of Sciences*, *1157*, 32–47. https://doi.org/10.1111/j.1749-6632.2009.04471.x

Zhang, X. S., Roy, R. J., & Jensen, E. W. (2001). EEG complexity as a measure of depth of anesthesia for patients. *IEEE Transactions on Bio-Medical Engineering*, *48*(12), 1424–1433. https://doi.org/10.1109/10.966601

Appendix A: Ethics

Western Research

Date: 5 February 2024

To: Dr. Adrian M. Owen

Project ID: 114967

Review Reference: 2024-114967-88867

Study Title: Improving Diagnosis and Prognosis in Acute Brain Injury: A Multimodal Imaging Approach (MIMIC Study - Multimodal IMaging in Intensive Care)

Application Type: Continuing Ethics Review (CER) Form

Review Type: Delegated

Date Approval Issued: 05/Feb/2024 14:44

REB Approval Expiry Date: 07/Feb/2025

Dear Dr. Adrian M. Owen,

The Western University Research Ethics Board has reviewed the application. This study, including all currently approved documents, has been re-approved until the expiry date noted above.

REB members involved in the research project do not participate in the review, discussion or decision.

Western University REB operates in compliance with, and is constituted in accordance with, the requirements of the Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans (TCPS 2); the International Conference on Harmonisation Good Clinical Practice Consolidated Guideline (ICH GCP); Part C, Division 5 of the Food and Drug Regulations; Part 4 of the Natural Health Products Regulations; Part 3 of the Medical Devices Regulations and the provisions of the Ontario Personal Health Information Protection Act (PHIPA 2004) and its applicable regulations. The REB is registered with the U.S. Department of Health & Human Services under the IRB registration number IRB 00000940.

Please do not hesitate to contact us if you have any questions.

Electronically signed by:

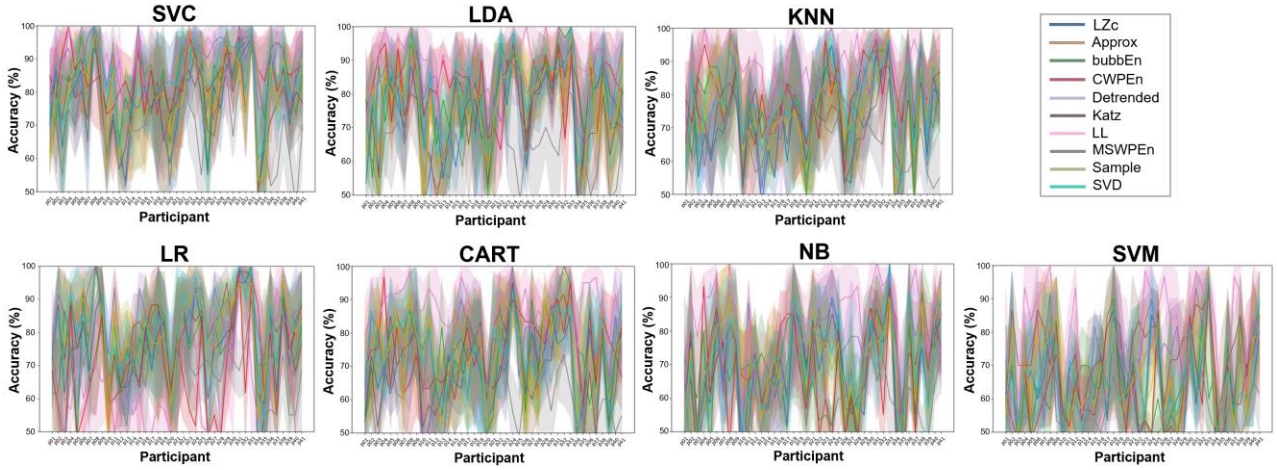
Mr. Joshua Hatherley, Ethics Coordinator on behalf of Dr. N. Poonai, HSREB Chair 05/Feb/2024 14:44

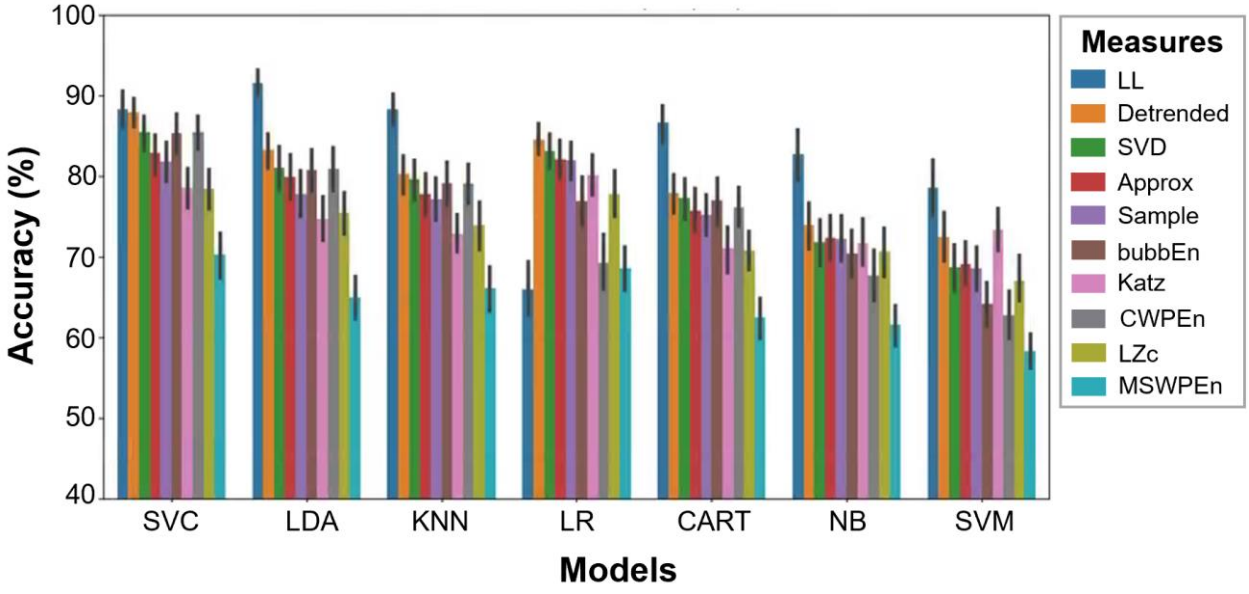
Reason: I am approving this document

Note: This correspondence includes an electronic signature (validation and approval via an online system that is compliant with all regulations).

Appendix B: Performance of Various Models and Measures for Patient Intra-Subject Level Binary Classification of Intact VS Scrambled.

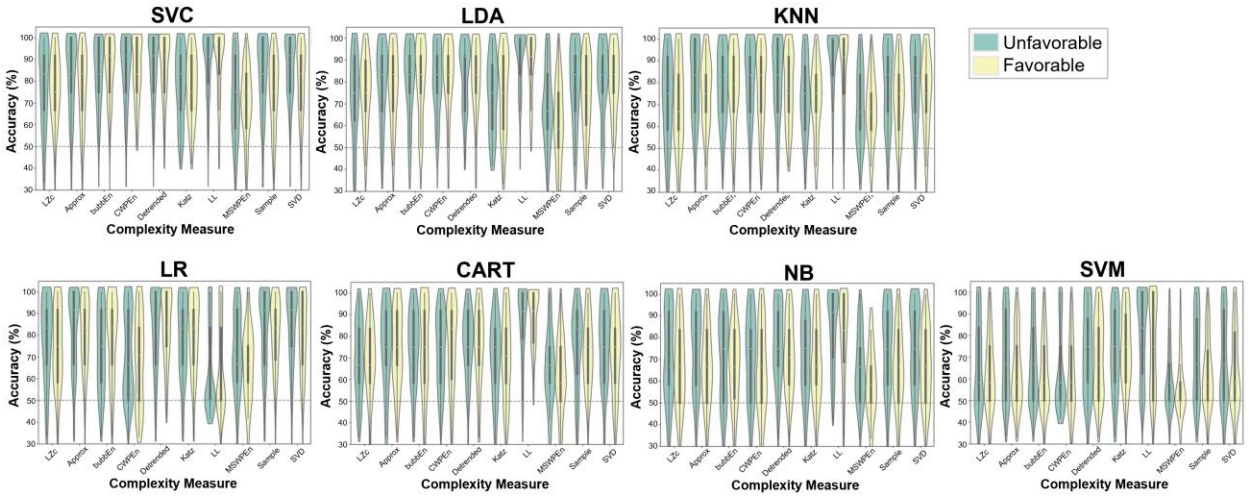
The data in the following graphs are derived from patient data only. I assessed the performance discriminating between Intact and Scrambled of various intra-subject level binary classification models using various complexity measures. The models analyzed include SVC, LDA, KNN, LR, CART, NB, and SVM. The measures used for comparison include LL, Detrended, SVD, Approx, Sample, bubbEn, Katz, CWPEn, LZc, and MSWPEn. The first plot represents the accuracy scores of each patient in the various combinations. The second plot is a bar graph of the overall means of each combination, represented in descending order.




Accuracy (%)
100
90
80
70
60
50
40
SVC
LDA
KNN
LR
CART
NB
SVM
Models
Measures
LL
Detrended
SVD
Approx
Sample
bubbEn
Katz
CWPEn
LZc
MSWPEn

Appendix C: Additional Part I Analysis with Other Measures

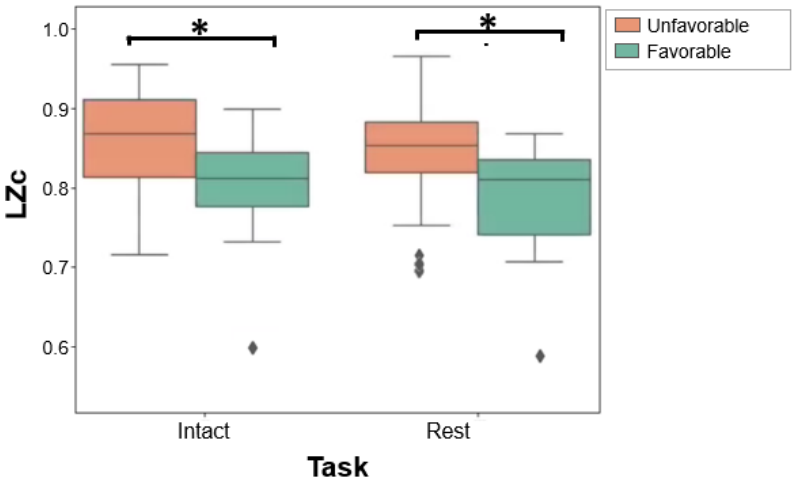
The data in the following graph are derived from patient data only. Following my analysis in Appendix B, I looked at differences between accuracy scores in the Favourable and Unfavourable conditions depicted in the violin plots below. The table shows descriptive and inferential statistics.

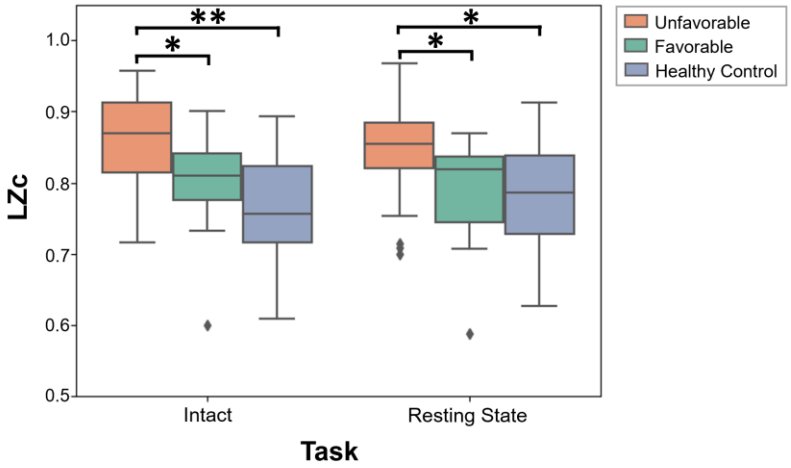


Model	Measure	Mean_Favorable	Mean_Unfavorable	tstat	pvalue	sig
SVC	ll	0.89	0.89	5349.5	0.329	*ns.*
SVC	detrended	0.87	0.86	5263.5	0.414	*ns.*
SVC	svd	0.81	0.86	4180	0.992	*ns.*
SVC	approx	0.77	0.82	4162	0.993	*ns.*
SVC	sample	0.76	0.83	3885.5	0.999	*ns.*
SVC	bubben	0.85	0.83	5563.5	0.173	*ns.*
SVC	katz	0.77	0.78	4971	0.688	*ns.*
SVC	cwpen	0.85	0.85	5184.5	0.491	*ns.*
SVC	LZc	0.76	0.78	4829.5	0.797	*ns.*
SVC	mswpen	0.7	0.72	4973.5	0.686	*ns.*
LDA	ll	0.88	0.9	4672	0.9	*ns.*
LDA	detrended	0.82	0.83	4995	0.669	*ns.*
LDA	svd	0.8	0.8	5060.5	0.609	*ns.*
LDA	approx	0.78	0.8	4884	0.758	*ns.*
LDA	sample	0.74	0.79	4384	0.971	*ns.*
LDA	bubben	0.81	0.8	5276	0.404	*ns.*
LDA	katz	0.75	0.74	5446	0.258	*ns.*
LDA	cwpen	0.82	0.81	5562.5	0.175	*ns.*
LDA	LZc	0.76	0.73	5440	0.263	*ns.*
LDA	mswpen	0.62	0.68	4172	0.992	*ns.*
KNN	ll	0.87	0.9	4812	0.824	*ns.*
KNN	detrended	0.77	0.81	4498	0.948	*ns.*

KNN	svd	0.75	0.8	4334.5	0.978	*ns.*
KNN	approx	0.75	0.79	4393	0.97	*ns.*
KNN	sample	0.72	0.78	4192	0.991	*ns.*
KNN	bubben	0.78	0.78	5232.5	0.446	*ns.*
KNN	katz	0.73	0.73	5126.5	0.547	*ns.*
KNN	cwpen	0.78	0.8	4808.5	0.811	*ns.*
KNN	LZc	0.7	0.72	4808.5	0.81	*ns.*
KNN	mswpen	0.67	0.69	5087.5	0.584	*ns.*
LR	ll	0.66	0.65	5272	0.401	*ns.*
LR	detrended	0.84	0.83	5216.5	0.46	*ns.*
LR	svd	0.8	0.84	4415	0.968	*ns.*
LR	approx	0.8	0.82	4661	0.893	*ns.*
LR	sample	0.8	0.82	4655.5	0.896	*ns.*
LR	bubben	0.79	0.74	6051.5	0.018	*ns.*
LR	katz	0.8	0.79	5426.5	0.273	*ns.*
LR	cwpen	0.69	0.71	4731	0.856	*ns.*
LR	LZc	0.74	0.78	4575.5	0.925	*ns.*
LR	mswpen	0.67	0.73	4226.5	0.988	*ns.*
CART	ll	0.89	0.86	5394.5	0.292	*ns.*
CART	detrended	0.75	0.77	5036.5	0.631	*ns.*
CART	svd	0.73	0.76	4751	0.845	*ns.*
CART	approx	0.75	0.75	4988.5	0.673	*ns.*
CART	sample	0.72	0.75	4480	0.952	*ns.*
CART	bubben	0.75	0.75	5287	0.395	*ns.*
CART	katz	0.7	0.7	5168	0.507	*ns.*
CART	cwpen	0.76	0.74	5629	0.138	*ns.*
CART	LZc	0.7	0.69	5353	0.335	*ns.*
CART	mswpen	0.62	0.64	4847	0.785	*ns.*
NB	ll	0.8	0.84	4819.5	0.807	*ns.*
NB	detrended	0.71	0.77	4273.5	0.985	*ns.*
NB	svd	0.67	0.74	3985	0.998	*ns.*
NB	approx	0.66	0.76	3786	1	*ns.*
NB	sample	0.65	0.75	3761	1	*ns.*
NB	bubben	0.68	0.72	4541	0.935	*ns.*
NB	katz	0.69	0.72	4709.5	0.868	*ns.*
NB	cwpen	0.65	0.71	4352	0.976	*ns.*
NB	LZc	0.66	0.71	4432	0.962	*ns.*
NB	mswpen	0.58	0.66	4013	0.997	*ns.*
SVM	ll	0.75	0.8	4539	0.94	*ns.*
SVM	detrended	0.7	0.72	4918.5	0.731	*ns.*
SVM	svd	0.64	0.68	4644.5	0.9	*ns.*
SVM	approx	0.64	0.71	4189.5	0.991	*ns.*
SVM	sample	0.63	0.7	4118	0.994	*ns.*
SVM	bubben	0.62	0.66	4607	0.916	*ns.*
SVM	katz	0.74	0.73	5211.5	0.466	*ns.*
SVM	cwpen	0.61	0.64	4893	0.754	*ns.*
SVM	LZc	0.62	0.66	4750.5	0.847	*ns.*
SVM	mswpen	0.55	0.6	4312.5	0.984	*ns.*

Appendix D: LZC for Outcome and Task (Resting State VS Intact)

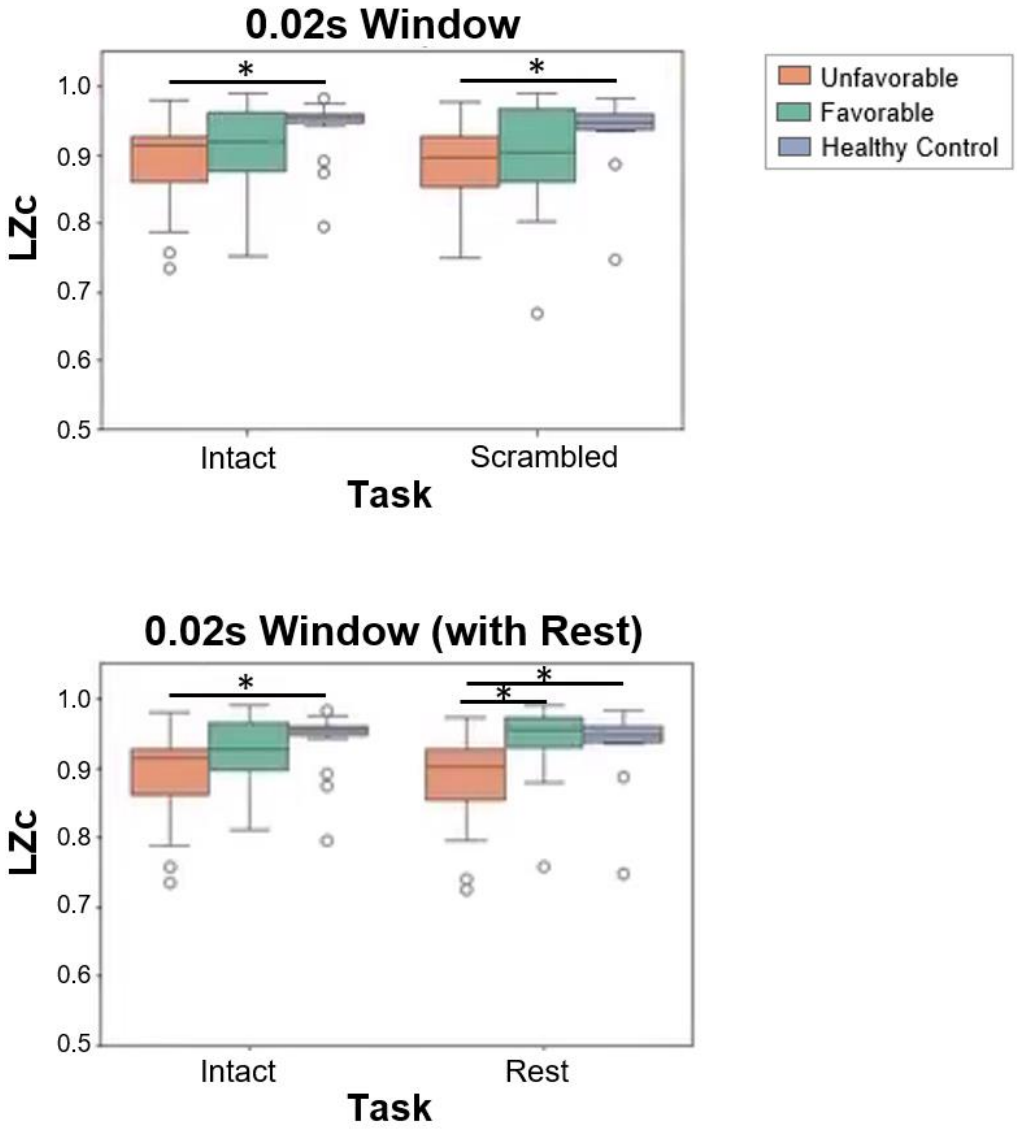




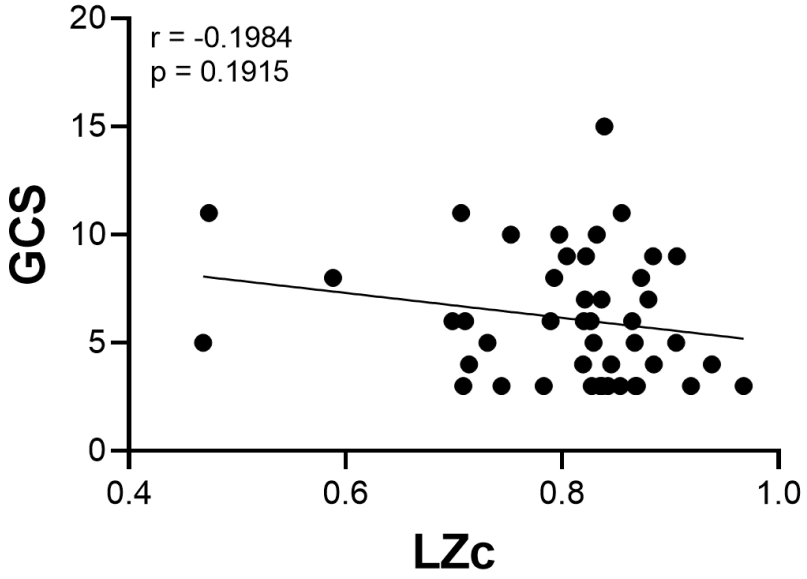
Descriptive Statistics for Appendix D

Outcome	Task	*n*	*M*	*SD*	*SEM*	Skewness	Kurtosis
Unfavourable	Intact	23	0.8593	0.0670	0.0140	-0.396	-0.5760
	Rest	23	0.8496	0.0680	0.0142	-0.470	-0.3590
Favourable	Intact	17	0.7877	0.1245	0.0302	-2.166	4.3265
	Rest	17	0.8037	0.0947	0.0230	-1.462	1.4697
Healthy Control	Intact	18	0.7686	0.0795	0.0187	-0.187	-0.7442
	Scrambled	18	0.7847	0.0753	0.0178	-0.387	-0.5965

Appendix E: LZC for Outcome and Task Calculated with Shortened Time Scales

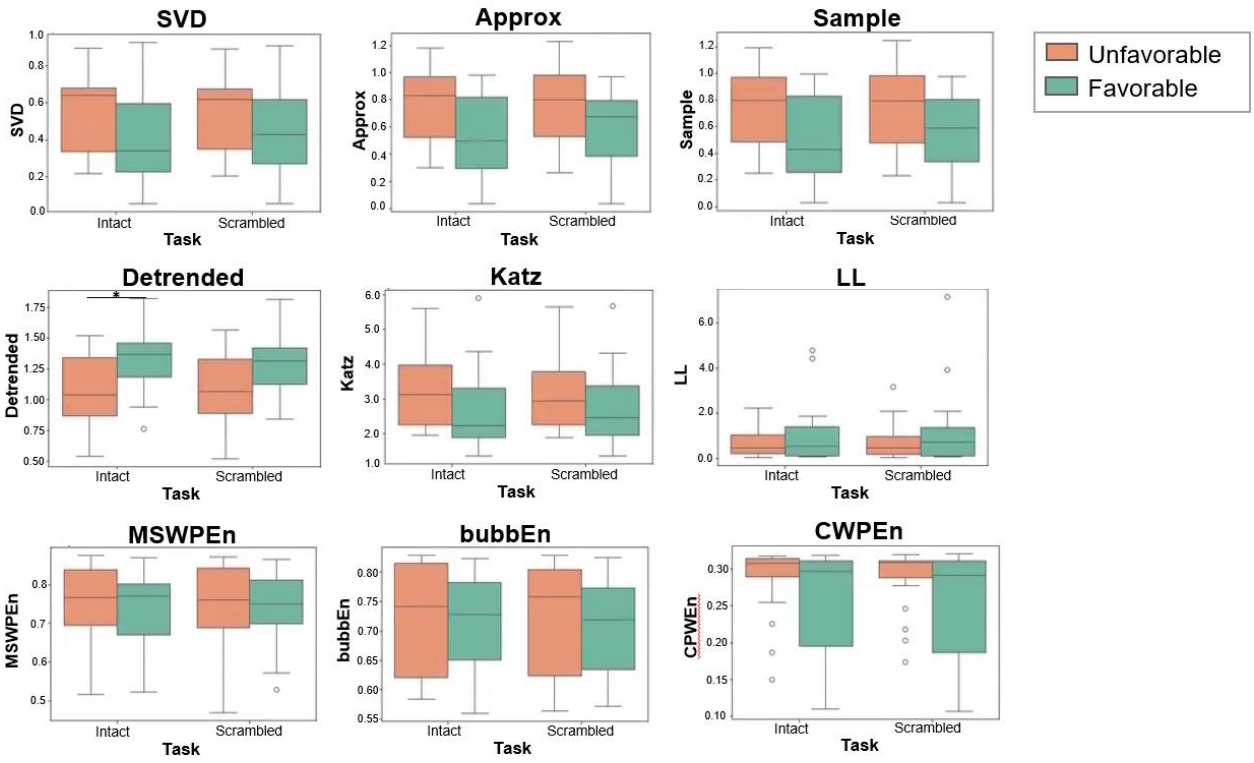


Appendix F: GCS and LZC Correlation

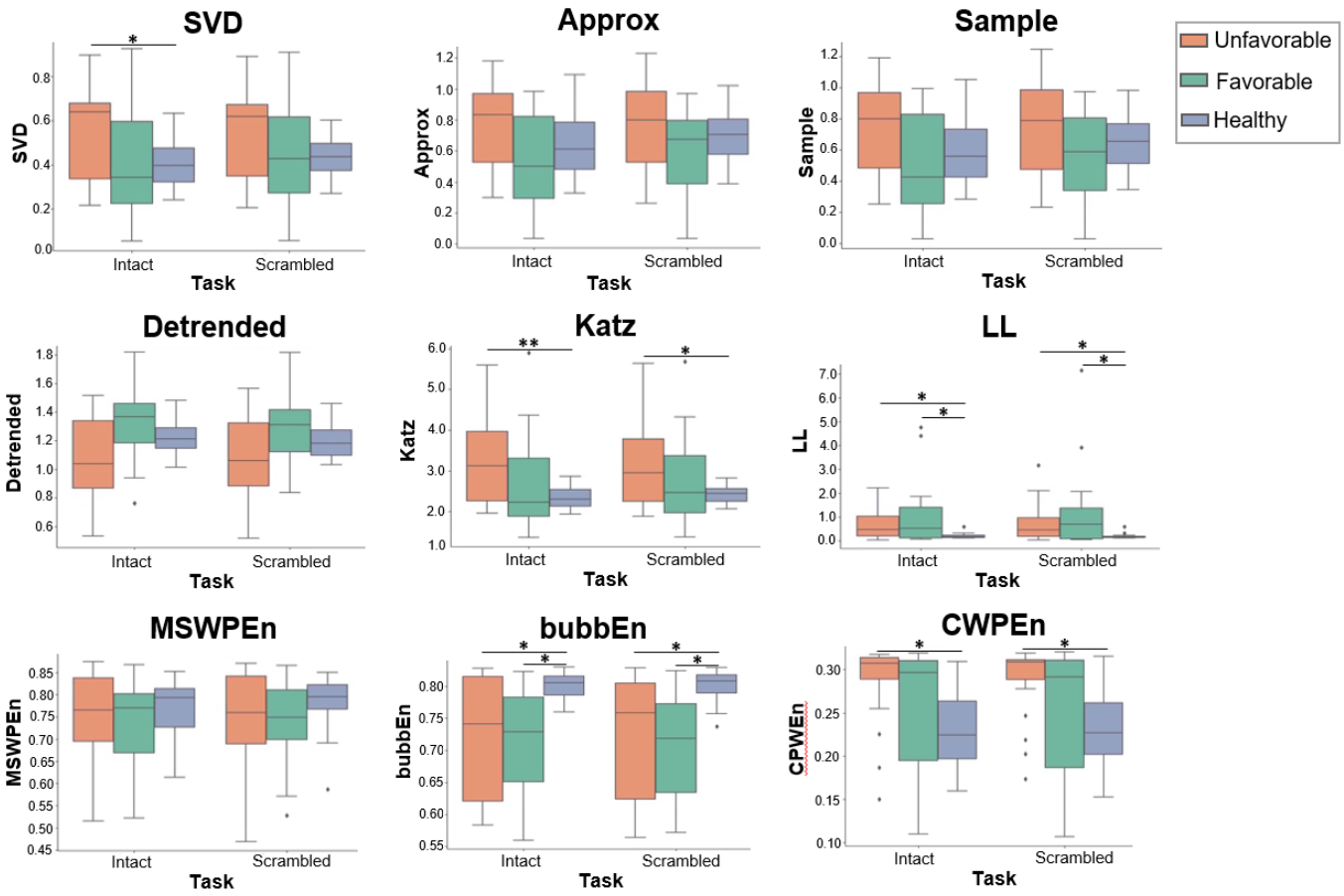


Appendix G: Additional Part II Analysis with Other Measures

The following Analysis used the same methods the one used in Part II LZc, but for other measures. Below are graphs for all the other temporal complexity measures averaged across time and electrodes to get a single value for each participant so that I can use normal statistical methods. The table below summarizes the main effects or interaction effects that can be observed for each of the measures that were used in the parametric comparison for within-subject difference of Intact and Scrambled and between-subjects differences of patients with favourable outcomes, unfavorable outcomes.



Furthermore, here are graphs that include healthy controls as an addition between-subject group.



Model	Outcome	Task	Interaction
LZc (Spatial-Temporal)	X		
SVD			X
Approx	X		
Sample	x		
Detrended	X		x
Katz			x
LL			
MSWPEn			
bubbEN			X
CWPEn	X		

**Note: X = significant main effect*

x = trending main effect (p<0.1)

Table results based on above graphs including HC group

Curriculum Vitae

EDUCATION

Master of Science (M.Sc.), University of Western Ontario **Sept. 2022 – Present**
London, ON

- Department of Psychology
- Cognitive, Developmental and Brain Sciences Program
- Supervisors: Dr. Adrian Owen & Dr. Derek Debicki

Bachelor of Arts (B.A.), King's University College **Sept. 2018 – Apr. 2022**
London, ON

- Honours Specialization in Psychology
- Thesis title: Cortical Function of Super Refractory Status Epilepticus: An fMRI Case Study
- Supervisor: Dr. Loretta Norton

RESEARCH WORK EXPERIENCE

Teaching Assistant – Psychology as a Social Science (PSYCH 1003) **Jan. 2024 – Present**
University of Western Ontario, London, ON

Teaching Assistant – Psychology as a Natural Science (PSYCH 1002) **Sept. 2023 – Dec. 2023**
University of Western Ontario, London, ON

Teaching Assistant – Research Methods in Psychology (PSYCH 2801) **Sept. 2022 – Dec. 2022**
University of Western Ontario, London, ON

Research Assistant – The Owen Lab **Oct. 2021 – Aug. 2022**
University of Western Ontario, London, ON

- Collected electroencephalogram (EEG) data in the Intensive Care Unit (ICU)
- Analyzed Functional Magnetic Resonance Imaging (fMRI) data

AWARDS, HONORS, & DISTINCTIONS

- 2024 Ontario Graduate Scholarship (OGS) funded by the Province of Ontario and the University of Western Ontario (valued at $15,500).
- 2023 Recipient of the Reva Gerstein Fellowship for Master's Study in Psychology (valued at $4,000).
- Gold medalist in the Greco-Roman category at the 2023 Canadian National Wrestling Championships.
- Gold medalist at the 2023 Ontario University Athletics (OUA) Wrestling Championship.
- Silver medalist at the 2023 U-Sports Wrestling Championship.
- 2022 Ontario Graduate Scholarship (OGS) funded by the Province of Ontario and the University of Western Ontario (valued at $15,500).
- 2021-2022 Ontario Wrestling Carding (valued at $3,000).
- U-Sports Academic All-Canadian Award (2019-2023).
- King's University College Dean's List (2019-2022).
- Peter Lockyer Wrestling Award (2018-2022) awarded by the University of Western Ontario's Varsity Wrestling Team (total value of $10,500).

- Gold medalist at the 2021 Ontario Senior Wrestling Provincial Championship.
- 2021 Recipient of the Bronze W Award presented by the University of Western Ontario's Varsity Wrestling Team.
- King's University College Continuing Admission Academic Scholarship (2018-2021) (total value of $4,500).
- King's University College Academic Athletic Scholarship (2018-2021) (total value of $12,000).
- 2019 & 2020 Most Dedicated Athlete awarded by the University of Western Ontario's Varsity Wrestling Team.
- 2019 Male Rookie of the Year awarded by the University of Western Ontario's Varsity Wrestling Team.

PUBLICATIONS

Conference Abstracts:

- **Al-Hayawi H**, Laforge G, Novi SL, Wang X, Debicki DB, Norton L, Owen AM: Leveraging electroencephalography and machine learning for predicting neurologic recovery after acute brain injury. London Health Research Day. London, Canada. May 2024.
- **Al-Hayawi H**, Laforge G, Norton L, Owen AM: Complexity Modulation with Naturalistic Narrative Stimuli for Prognosis of Acute Brain-Injured Patients. Cognitive Neuroscience Society 2024 31st Annual Meeting. Toronto, Canada. April 2024.
- **Al-Hayawi H**, Laforge G, Novi SL, Wang X, Debicki DB, Norton L, Owen AM: Using EEG and machine learning for predicting neurologic recovery after acute brain injury. 51st Annual Lake Ontario Visionary Establishment (L.O.V.E) Conference. Niagara Falls, Canada. February 2024.
- **Al-Hayawi H**, Norton L, Kazazian K, Gofton TE, Owen AM: Cortical Function of Super Refractory Status Epilepticus: An fMRI Case Study. London Health Research Day. London, Canada. May 2022.

Publications:

- Maschke C, Norton L, Duclos C, Dolhan K, Han M, Laforge G, Frantz A, Wang X, **Al-Hayawi H**, Zhang T, Lavoie R, Owen AM, Blain-Moraes S: EEG brain response during sedation interruption complements behavioral assessment following brain injury (manuscript in preparation).

VOLUNTEERSHIP AND EXTRACURRICULARS

Team Captain of the UWO Varsity Wrestling Team **Sept. 2021 – Apr. 2023**
University of Western Ontario, London, ON

Cognitive Neuroscience Journal Club Member **Sept. 2021 – Apr. 2022**
King's University College, London, ON

- Led by Dr. Loretta Norton

Psychology Research Club Member **Sept. 2020 – Sept. 2022**
King's University College, London, ON

- Led by Dr. Christopher Roney

Alumni Sports Fundraising Volunteer **Sept. 2019 – Apr. 2023**
London, ON

- Ran the cash register during BINGO charity events to raise funds for the London-Western Wrestling Club.
- Helped run two fundraising alumni golf tournaments to raise funds for varsity sports.