Characterizing the Role of Neural Dynamics in the Treatment of Depression

by

Sravya Atluri

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy

Institute of Biomaterials and Biomedical Engineering
University of Toronto

Characterizing the Role of Neural Dynamics in the Treatment of Depression

Sravya Atluri
Doctor of Philosophy
Institute of Biomaterials and Biomedical Engineering
University of Toronto
2019

Abstract

Selecting an appropriate treatment for patients with depression is challenging for several reasons. There is no clear understanding on (i) the pathophysiology of depression, (ii) heterogeneity in depression, and (iii) targets for successful treatment outcome. As such, although treatments for depression are effective, their average efficacy seems to be poor. It is widely accepted that seizures induced in the brain are highly effective for severe, treatment-resistant cases of depression. Seizures are also known to impact the dynamics of neural activity. Based on this knowledge, we investigated whether treatments for depression impact neural dynamics for therapeutic efficacy. We also evaluated whether measures of neural dynamics can predict response. Seizure therapy (electroconvulsive therapy and magnetic seizure therapy) and pharmacotherapy (escitalopram) were studied. It is hypothesized that modulations of neural dynamics in several frequencies, timescales, regions and networks, previously shown to be affected in depression, are associated with therapeutic outcome. These modulations are also hypothesized to be distinct from modulations associated with non-response. In this work, measures of neural dynamics were derived from power spectral density analysis, multiscale entropy analysis and microstate analysis of resting-state, eyes-closed EEG data. Results suggest

that successful seizure therapy potentially impacts several characteristics of neural dynamics for therapeutic efficacy. In responders of seizure therapy, modulation of neural dynamics was observed in regions (posterior cingulate cortex, precuneus, occipital pole) and networks (salience, fronto-parietal) previously known to be impaired in depression. In responders of escitalopram, modulation of neural dynamics was observed after 2 weeks into the 8-week course of escitalopram treatment. These changes were observed in regions known to be impaired in depression (posterior cingulate cortex, precuneus, posterior cingulate cortex) but not within networks. In non-responders of escitalopram, an early modulation of neural dynamics (i.e., baseline to 2 weeks) was observed. Finally, using measures of neural dynamics, prediction of response to escitalopram achieved an accuracy of 83.2%. Knowledge from this work will guide the development of antidepressant response prediction tools and potentially improve treatment efficacy in depression.

Acknowledgements

First and foremost, I would like to thank my supervisor **Dr. Faranak Farzan**. Words fall short when describing my gratitude for you. Time after time, your strength inspired me, your kindness motivated me, and your support encouraged me to overcome my obstacles and embrace my toughest challenges. Thank you for always pushing me to be the best intellectual version of myself. I will always have immense respect and admiration for you in my heart.

I would also like to express my heartfelt gratitude to my co-supervisor, **Dr. Willy Wong**. One of the most brilliant minds in academia I know, but also modest and humble. Because of you, I will never lose my curiosity and will always think about the 'why not' instead of the 'why'. Your endless dedication and care towards your students, and your admirable work ethic is truly inspiring. I hope to carry these qualities into the next stage of my life.

My work would be incomplete without the tremendous support and guidance of my committee members, **Dr. Zafiris Jeff Daskalakis** and **Dr. Daniel Blumberger**. Your consistent feedback and critical questions have contributed to the improved quality of my research. I am extremely grateful to have had the opportunity to learn from both of you. I would also like to sincerely thank my external reviewers, **Dr. George Foussias** and **Dr. Gregory Light**, for their valuable time and input on this thesis.

The last four years of my life spent at Temerty would have been incomplete without my fellow lab mates, colleagues and dear friends. **Apoorva:** thank you for your constant encouragement and support. Your time management skills are my goals. **Yinming, Colin and Reza:** thank you for your generous input in my work. I am grateful for the technical and critical-thinking skills I gained because of you. **Matt:** thank you dealing with my overwhelming need for perfection and for always pulling me out of it. **Mawahib:** thank you for the entertaining conversations and sharing my love for food. They were saviours during my most tiring days. **Stacey:** thank you for teaching me TMS-EEG and most importantly, for building my biceps in the process. Last, but in no way the least, a big thank you to all of my lab roomies **(Michelle, Jen, Paul, Jeanette and**

our newest addition Elnaz) for your constant support, encouragement and advice. I will miss you all dearly.

My dearest friends from Ottawa who have been so patient with me and stayed in touch even when I have fallen back on that promise, thank you from the bottom of my heart. To my first friend in Toronto, **Tsukiko**, thank you for teaching me the importance of balance in academia and for your valuable friendship. From Temerty friend and roommate to best friend and sister: **Felicity**, thank you for your immense support over the past four years. Your quirks (and humour) lighten up my worst days and your curiosity is contagious. My partner **Naren**, thank you for staying by my side in the toughest time. Your infinite patience, care and support are the pillars of my strength. You have changed my life for the better.

I would not be the person I am today without my family. Many of my principles and values are reflections of my parents: **Uma Devi** and **Durga Prasad**. Thank you for your love, your sacrifices, and despite our objection, for always prioritizing your children's happiness over yours. My success, now and always, will be because of you. To my siblings **Swetha, Sruthi and Sai**, thank you for always keeping me grounded with your uncensored honesty. I would not have it any other way. To the love of my life, my niece **Ishika**, thank you for all the giggles and hugs exactly when I needed them. I dedicate this thesis to all of you.

Table of Contents

List of Tables

List of Figures

List of Supplementary Figures

List of Abbreviations

ACC	Anterior Cingulate Cortex
aMCC	Anterior Mid-Cingulate Cortex
AMI	Autobiographical Memory Interview
ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
ATR	Antidepressant Treatment Response
AUC	Area Under the Curve
BDI	Beck's Depression Inventory
BDNF	Brain Derived Neurotrophic Factor
BT-ECT	Bitemporal Electroconvulsive Therapy
CAMH	Centre for Addiction and Mental Health
CAN-BIND	Canadian Biomarker Integration Network in Depression
DMN	Default-mode Network
dPCC	dorsal Posterior Cingulate Cortex
DSM	Diagnostic and Statistical Manual
ECT	Electroconvulsive Therapy
EEG	Electroencephalography
fMRI	functional Magnetic Resonance Imaging
GABA	Gamma-Aminobutyric Acid
GUI	Graphical User Interface
HAMD	Hamilton Rating Scale for Depression
HRSD	Hamilton Rating Scale for Depression
HSD	Honest Significant Difference
ICA	Independent Component Analysis
LORETA	Low Resolution Brain Electromagnetic Tomography
MADRS	Montgomery-Åsberg Depression Rating Scale
MAOI	Monoamine Oxidase Inhibitor
MeFG	Medial Frontal Gyrus
MiFG	Middle Frontal Gyrus

MINI	Mini International Neuropsychiatric Inventory
MNI	Montreal Neurological Institute
MoCA	Montreal Cognitive Assessment
MSE	Multiscale Entropy
MST	Magnetic Seizure Therapy
PCC	Posterior Cingulate Cortex
pMCC	Posterior Mid-Cingulate Cortex
QNS	Queens University
rACC	rostral Anterior Cingulate Cortex
RBF	Radial Basis Function
ROC	Receiver Operating Characteristic
ROI	Region of Interest
rTMS	repetitive Transcranial Brain Stimulation
RUL-UB ECT	Right Unilateral Ultra Brief Electroconvulsive Therapy
SCID	Structured Clinical Interview for DSM
SFG	Superior Frontal Gyrus
SNRI	Serotonin Norepinephrine Re-uptake Inhibitor
SSI	Scale for Suicidal Ideation
SSRI	Selective Serotonin Re-uptake Inhibitor
SVM	Support Vector Machine
TGH	Toronto General Hospital
TMS	Transcranial Magnetic Stimulation
TRD	Treatment-Resistant Depression
UBC	University of British Columbia
vPCC	ventral Posterior Cingulate Cortex

Section I: Introduction, Objectives, and Methods

Chapter 1 - Insight into Depression and Current Treatments for Depression from Previous Literature

1.1 Motivating Problem

Depression is the leading cause of years lost due to disability (World Health Organization 2009) and is ranked as the fourth medical condition with the greatest disease burden worldwide (Vos, Allen et al. 2016). This is expected to rise to second by 2020 (Murray, Lopez et al. 1996). One major cause for the social and economic burden of depression is the number of years it can take to find an ideal treatment. Currently, a wide range of treatments are available for depression including psychotherapy, medications and brain stimulation (Duval, Lebowitz et al. 2006). The Food and Drug Administration alone has approved 30+ drugs for the treatment of depression based on double-blind placebo-controlled studies (Food and Drug Administration 2013). Despite the availability of effective treatment, poor understanding on the pathophysiology of depression, the cause for heterogeneity in depression and targets for successful antidepressant response, have made it challenging for clinicians to identify an effective treatment for patients. As a result, patients undergo a trial-and-error process, receiving multiple courses of treatment, before noticing benefits. Some may never even reach remission (the complete disappearance of symptoms) (Paykel, Ramana et al. 1995, Rush, Trivedi et al. 2006, Souery, Papakostas et al. 2006, Trivedi, Rush et al. 2006). To reduce the time spent in failed trials and avoid the debilitating impact of untreated depression, reliable predictors of treatment response must be identified. Although clinical scales are useful for diagnosis, they may not be sufficient for predicting treatment outcome (Serretti, Olgiati et al. 2007, Howland, Wilson et al. 2008, Chekroud, Zotti et al. 2016). In this thesis, we aim to identify potential neurophysiological targets of successful treatments for depression and predictors of treatment response.

Patients who show similar symptoms at the behavioral level can have distinct sources at the biological level (Vuilleumier 2005). Therefore, we hypothesize that biological markers may be sensitive enough to reliably predict differences in benefits and/or adverse effects of a treatment for an individual patient. In this thesis, we focus on non-invasive measures of neural activity (mechanistic and predictive) for their accessibility and potential for clinical translation. We study these measures over the course of different treatments for depression (i.e., first line of treatment vs. treatments for severe depression) that vary in their clinical efficacy and side effects. The knowledge gained from this work is hoped to benefit the development of clinical decision-

making tools for those who do not respond to standard treatments and improve the success rate of treatments for depression.

1.2 Characterizing Depression

Depression is a debilitating disorder with a significant negative socioeconomic and quality-of-life impact (Üstün, Ayuso-Mateos et al. 2004, Lerner and Henke 2008, Olchanski, Myers et al. 2013). For a majority of patients, depression is a reoccurring condition characterized by the persistence of symptoms over multiple episodes that may be months, or years apart (Keller and Boland 1998, Richards 2011). Two cases of depression are studied in this thesis: major depressive disorder and treatment-resistant depression. Clinical definitions are provided in the following sections.

1.2.1 Major Depressive Disorder

Major depressive disorder is a medical condition characterized by a wide range of symptoms. During a single major depressive episode, symptoms range from abnormalities in mood, cognition, psychomotor activity as well as sleep and appetite disturbances (Fava and Kendler 2000, Belmaker and Agam 2008). Diagnosis criteria of major depressive disorder is specified in the 5th edition of the Diagnostic and Statistical Manual of Mental Disorders (DSM-5) (American Psychiatric Association 2013). It includes the presence of 5 or more symptoms listed in the manual during a 2-week period, where at least one symptom is depressed mood or loss of interest. The list of symptoms (taken from (American Psychiatric Association 2013)) may include:

*"**1.** Depressed mood most of the day, nearly every day, as indicated by either subjective report (e.g., feels sad, empty, hopeless) or observation made by others (e.g., appears tearful). (Note: In children and adolescents, can be irritable mood.)*
***2.** Markedly diminished interest or pleasure in all, or almost all, activities most of the day, nearly every day (as indicated by either subjective account or observation.)*

3. *Significant weight loss when not dieting or weight gain (e.g., a change of more than 5% of body weight in a month), or decrease or increase in appetite nearly every day. (Note: In children, consider failure to make expected weight gain.)*

4. *Insomnia or hypersomnia nearly every day.*

5. *Psychomotor agitation or retardation nearly every day (observable by others, not merely subjective feelings of restlessness or being slowed down).*

6. *Fatigue or loss of energy nearly every day.*

7. *Feelings of worthlessness or excessive or inappropriate guilt (which may be delusional) nearly every day (not merely self-reproach or guilt about being sick).*

8. *Diminished ability to think or concentrate, or indecisiveness, nearly every day (either by subjective account or as observed by others).*

9. *Recurrent thoughts of death (not just fear of dying), recurrent suicidal ideation without a specific plan, or a suicide attempt or a specific plan for committing suicide."*

Several clinical instruments are available for the quantitative measurement of illness severity, ranging from clinician-rated scales to self-report scales (Cusin, Yang et al. 2009). Some common scales used by clinicians are the Hamilton Rating Scale for Depression (HRSD) and the Montgomery-Åsberg Depression Rating Scale (MADRS). These scales measure multiple items across the domain of the illness to provide a better insight into the severity of major depression. One commonly used self-report scale is the Beck's Depression Inventory (BDI). Self-report scales avoid potential clinician or researcher biases and provide a unique evaluation (Uher, Perlis et al. 2012). However, they may be confounded by cognitive impairment or thought disorder (Austin, Deary et al. 1998).

Based on consensus from the literature and clinical practice, response is usually defined as a 50% or greater reduction in the clinical assessment score from baseline. Remission after treatment is defined as a full recovery (i.e., no symptoms) or a score below or equal to a certain number for each scale. For the HRSD-17 scale for example, remission is usually characterized by a score less than or equal to 6.

1.2.2 Treatment-Resistant Depression

Approximately one third of the major depression patient population does not respond to standard, adequate treatments for depression (Fava 2003). An adequate dose is close to the maximum recommended dose. An adequate treatment length is at least 4 consecutive weeks, where an adequate dose is provided for at least 3 of those 4 weeks (Thase and Rush 1995). If a patient does not respond to 2 or more successive courses of adequate treatment (psychotherapy or medications), they are termed as treatment-resistant depression patients. It is estimated that the prevalence of treatment-resistance in primary care is 21.7% in Canada (Rizvi, Grima et al. 2014). In the STAR*D trial, the prevalence rate was estimated to be even higher at 30% (Warden, Trivedi et al. 2007). Studies that have examined clinical differences between resistant and non-resistant depression (Fagiolini and Kupfer 2003) identified greater severity in risk factors associated with resistance such as high recurrence rates (up to 80%) (Fekadu, Wooderson et al. 2009), psychiatric comorbidity (Souery, Oswald et al. 2007), undetected hypomania (Rush, Trivedi et al. 2006), and even mortality (Fekadu, Wooderson et al. 2009).

1.3 Theories on the Pathophysiology of Depression

The pathology of depression is not fully understood. In addition to the observable clinical characteristics associated with depression, several physiological abnormalities have been reported at the genetic level, at the cellular level, in functional brain network connectivity and in global brain function (Belmaker and Agam 2008). These abnormalities likely vary between patients and lead to the heterogeneity of depression.

1.3.1 Genetic Level

It is estimated that 40-50% of depression cases may be caused by genetic factors (Lohoff 2010). Due to the heterogeneity seen in clinical symptoms, multiple genes are hypothesized to be involved in depression. Here, we highlight two that have been more widely explored.

(1) **Serotonin Transporter Gene** is involved in the regulation of serotonin re-uptake transporters. A high level of these transporters increases the reabsorption of serotonin and decreases the level of the serotonin neurotransmitter in the brain (Deakin 1991, Lucki

1998, Mann 1999, Parsey, Hastings et al. 2006). Low levels of serotonin have been widely associated with depression (explained in section 1.3.2).

(2) **Brain Derived Neurotrophic Factor (BDNF) Gene** has an important role in neurogenesis and mood disorders (Lee and Kim 2010). Low levels of BDNF in the hippocampus have been associated with chronic stress and depressive personality traits (Karege, Perret et al. 2002). Specifically, the *Val66Met* polymorphism in the BDNF gene is observed in individuals with a history of depression (Chen, Jing et al. 2006).

1.3.2 Cellular Neurotransmission

Neurotransmitters are the chemical messengers of neuronal communication in the brain. The junction between two neurons is called the synapse. A message is transmitted when a presynaptic neuron releases a neurotransmitter (i.e., sends a message) and the postsynaptic neuron absorbs the neurotransmitter through receptors (i.e., receives the message). Excess molecules are reabsorbed by the presynaptic neuron (Kandel, Schwartz et al. 2012). Dysfunction in neuronal communication may result during different stages of this process and lead to: (1) low levels of the neurotransmitter in the presynaptic neuron, (2) reduced expression or availability of receptors at the postsynaptic neuron, (3) faster reabsorption by the presynaptic neuron than absorption by postsynaptic neurons, and/or (4) reduced availability of molecules guiding this process. Several neurotransmitters are thought to have a key role in the emergence of depressive symptoms (Belmaker and Agam 2008, Krishnan and Nestler 2008).

(1) Serotonin is one neurotransmitter that is widely known to play a key role in depression (Owens and Nemeroff 1994). Several lines of evidence link serotonin with the regulation of mood, sleep, memory and other functions known to be impaired in depression (Owens and Nemeroff 1994). Specifically, low levels of serotonin are suggested to lead to depression.

(2) Dopamine has a complex role in neurological and mental functions. Dopamine neurons extend (through dopaminergic pathways) from the midbrain to a number of other brain regions (Lindvall and Björklund 1978) including the basal ganglia (movement), prefrontal cortex (problem-solving, intelligence, complex thoughts), amygdala (emotional processing) and the hippocampus (memory) (Robbins 2003). Dopamine also plays a

major role in motivation and reward. Due to its wide range of effects, low levels of dopamine have been linked to several psychiatric illnesses including depression (Brown and Gershon 1993).

(3) As a neurotransmitter, high levels of **norepinephrine** are associated with increased energy, alertness, concentration and cognitive ability (Ressler and Nemeroff 1999, Goddard, Ball et al. 2010). Reduced norepinephrine neurotransmission is associated with depressive symptoms (Moret and Briley 2011). Interestingly, modulation of serotonin levels may indirectly raise levels of norepinephrine and dopamine (Bymaster, Zhang et al. 2002).

(4) **Gamma-Aminobutyric Acid (GABA)** plays a major role in inhibitory neurotransmission and is crucial for controlling brain excitability (Chebib and Johnston 1999) or hyperconnectivity potentially seen in depression (Sheline, Price et al. 2010). Accumulating evidence suggests a specific GABAergic dysfunction in mood disorders (Kalueff and Nutt 2007). Low GABA levels seen in depression (Sanacora, Mason et al. 1999, Sanacora, Gueorguieva et al. 2004, Hasler, van der Veen et al. 2007, Bhagwagar, Wylezinska et al. 2008) may lead to decreased neurogenesis in the hippocampus (Earnheart, Schweizer et al. 2007) and reduced GABA metabolism (Shelp, Bown et al. 1999).

To summarize, patients with depression reveal low levels of neurotransmitter concentrations in regions associated with mood and emotion. The symptoms of depression observed at the behavioral level may be linked to dysfunctions in neuronal communication at the cellular level.

1.3.3 Brain Regions

The pathophysiology of depression may also be described by investigating abnormalities in three subdivisions of the brain:

(1) **Cortical regions** include the dorsal and medial prefrontal cortex, dorsal and ventral anterior cingulate cortex, subgenual cingulate cortex, orbital frontal cortex and the insula (Drevets, Price et al. 1997, Mayberg, Liotti et al. 1999, Bremner, Vythilingam et al. 2002, Kimbrell, Ketter et al. 2002, Anand, Li et al. 2005, Rigucci, Serafini et al. 2010, Sprengelmeyer, Steele et al. 2011). The prefrontal cortex is implicated in complex

cognitive behavior such as reasoning, planning and personality expression. Studies have shown that reduced metabolism in the prefrontal cortex, combined with the inclination to act on negative emotions, may result in suicidal behavior (Arango, Underwood et al. 1995, Desmyter, Van Heeringen et al. 2011). Depression is also linked to reduced metabolism in the anterior cingulate cortex and the subgenual cingulate cortex, regions implicated in the cognitive aspects of emotional processing (Drevets, Price et al. 1997, Mayberg, Liotti et al. 1999). A reduced volume of the orbitofrontal cortex is also seen in depression, and this region is known to be implicated in mood regulation (Bremner, Vythilingam et al. 2002). Finally, increased activation of the insula and insular regions is linked to depression (Anand, Li et al. 2005). The insula has a crucial role in how emotional experiences are processed, assessed and responded to (Modinos, Ormel et al. 2009, Lamm and Singer 2010).

(2) **Subcortical limbic regions** include the hippocampus, amygdala, and dorsomedial thalamus (Sheline, Gado et al. 1998, Bremner, Narayan et al. 2000, Schweitzer, Tuckwell et al. 2001, MacQueen, Yucel et al. 2008, Lorenzetti, Allen et al. 2009). Depression is associated with decreased volume and activation in these regions suggesting neurodegeneration.

(3) **Basal ganglia and the brainstem** (striatum) have an important role in reward and motivation (Balleine, Delgado et al. 2007). The brainstem contains the brain serotonergic neurons, norepinephrine neurons and the dopaminergic neurons. Some imaging studies have shown abnormalities in activation or metabolism in these regions. However, results are not consistent (Kumari, Mitterschiffthaler et al. 2003, Surguladze, Brammer et al. 2005, Knutson, Bhanji et al. 2008, Remijnse, Nielen et al. 2009).

1.3.4 Connectivity between Brain Regions

Although abnormalities in several, distinct brain regions are associated with depression, findings of these local changes in neural activity are not consistent. Some studies show an increase while others show a decrease in the activity of any single brain region (Drevets, Videen et al. 1992, Andreasen, Paradiso et al. 1998, McIntosh and Gonzalez-Lima 1998, Friston 2002, Lawrie, Buechel et al. 2002, Anand, Mathews et al. 2003, Mayberg 2003). Recent insight from

neuroimaging studies suggests that depression may be linked to abnormalities in the connectivity of several brain regions rather than discrete brain regions. For example, in patients with depression, reduced metabolism in the prefrontal cortex is linked with increased metabolism in limbic regions (striatum and thalamus) (Mayberg, Liotti et al. 1999). In addition, reduced metabolism is observed in the subgenual cingulate cortex (Greicius, Flores et al. 2007) and this region is connected to several other brain regions implicated in depression including the anterior cingulate cortex, amygdala and the dorsomedial thalamus (i.e., amygdala-striatal-pallidial-thalamic-cingulate cortex circuit) (Drevets, Price et al. 1997, Price and Drevets 2010).

1.3.5 Large-Scale Neural Networks

Evidence from functional magnetic resonance imaging (fMRI) studies suggests that abnormalities in three neural networks, and the relative activation between these networks, may be the facilitators of depressive symptoms: (1) the salience network, (2) the frontoparietal network, and (3) the default mode network.

(1) The **salience network** includes the anterior insula and the dorsal anterior cingulate cortex (Seeley, Menon et al. 2007). It also includes three subcortical structures: the amygdala, the ventral striatum and the substantia nigra (Seeley, Menon et al. 2007). The salience network is involved in several complex brain functions such as communication, social behavior and self-awareness (Menon and Uddin 2010). In addition, it plays an important role in the integration of sensory, emotional and cognitive information. Abnormalities in this network may result in subjective and aberrant detection or processing of emotional information, potentially leading to symptoms of depression (Harrison, Pujol et al. 2008, Manoliu, Meng et al. 2014, Kaiser, Andrews-Hanna et al. 2015).

(2) The **frontoparietal network** includes the inferior parietal lobe and the prefrontal cortex (Dosenbach, Fair et al. 2008, Zanto and Gazzaley 2013). Subcortical structures include the brainstem (Paus 2000). This network is involved in cognitive control and in the mediation of other regions/networks potentially impacted in depression (i.e., salience and default-mode network (Cole, Repovš et al. 2014, Kaiser, Andrews-Hanna et al. 2015).

(3) The **default-mode network** includes the precuneus, posterior cingulate gyrus, inferior parietal lobule, angular gyrus, the frontal pole and parts of the medial and lateral temporal

cortex (Greicius, Krasnow et al. 2003, Fransson and Marrelec 2008). Subcortical structures include the amygdala, hippocampus and the parahippocampus (Greicius, Supekar et al. 2009). This network is activated during self-referential thought and shows high activity during resting-state (Sheline, Barch et al. 2009, Kaiser, Andrews-Hanna et al. 2015).

A recent comprehensive meta-analysis of resting-state fMRI studies proposed a neurocognitive network model of depression (Kaiser, Andrews-Hanna et al. 2015). In this model, depression is associated with three aberrant network interactions: (i) default-mode network dominance over the frontoparietal network, (ii) abnormal switching between the default mode network and the frontoparietal network due to an impaired salience nnetwork, and (iii) ineffective frontoparietal network modulation of the default-mode network.

1.4 Treatments for Depression

Due to the complex pathophysiology of depression, no single treatment seems to show efficacy in all patients. For patients with mild to moderate severity of depression, initial treatment may include psychotherapy, medications or a combination of both. For patients with severe depression, antidepressants, antipsychotics, repetitive transcranial brain stimulation (rTMS) and/or electroconvulsive therapy (ECT) may be considered. For treatment-resistant depression, ECT is currently the most effective treatment.

1.4.1 Psychotherapy

Psychotherapy is often the first step in the treatment for depression (Cuijpers, van Straten et al. 2008). It can be effective when clinical symptoms include the presence of psychosocial stressors, interpersonal or intrapsychic conflicts, and comorbidities including personality disorders as specified in DSM-5. Two main types of psychotherapy are available: interpersonal psychotherapy (Klerman and Weissman 1994) and cognitive therapy (Beck 1979). The focus and goal of interpersonal psychotherapy is to improve communication within relationships and help develop a network that can support the individual during depressive episodes. Cognitive therapy,

on the other hand, focuses on identifying abnormalities in information processing during the perception and interpretation of emotions. By identifying these fundamental distortions, treatment can be personalized to modify this behavior. Although psychotherapy shows a high acceptance rate, it has marginal efficacy for more severe types of depression (Thase, Greenhouse et al. 1997).

1.4.2 Pharmacotherapy

In general, pharmacotherapy has shown efficacy in the treatment of major depressive disorder and more severe depression (Fava 2003). In fact, the Food and Drug Administration has approved 30+ drugs for the treatment of depression with the support of several double-blind, placebo-controlled studies (Food and Drug Administration 2013). The challenge, however, is identifying which drug is best suited for a patient at any given time. Remission rates are around 30% for the first trial and decline progressively with subsequent medication trials (Rush, Trivedi et al. 2006, Trivedi, Rush et al. 2006). Antidepressant medications can be categorized into different groups based on their effects at neuronal synapses.

(1) **Selective serotonin re-uptake inhibitors (SSRIs)** are the most commonly prescribed antidepressants and often preferred for their favorable short-term side-effect profile (Goodwin 1996, Goldstein and Goodnick 1998). SSRIs show efficacy by blocking the re-uptake (reabsorption) of serotonin and therefore boost the levels of serotonin in the brain (Stahl 1998). As mentioned before, low levels of serotonin in the brain have been associated with low mood and increasing the levels of serotonin is thought to elevate mood (Meltzer 1990). Although the short-term side effects are minimal, long-term side effects of SSRIs such as insomnia, weight gain, apathy, etc. may still be present (Masand and Gupta 2002). Examples of SSRIs include escitalopram, citalopram, fluoxetine, paroxetine, sertraline and vilazodone.

(2) **Serotonin norepinephrine re-uptake inhibitors (SNRIs)** block the uptake (absorption) of serotonin at lower doses and block the uptake of both serotonin and norepinephrine at higher doses (Stahl, Grady et al. 2005). Low levels of norepinephrine have been associated with low mood among other symptoms (Dell'Osso, Buoli et al. 2010). For efficacy, SNRIs may increase the levels of both serotonin and norepinephrine to elevate symptoms of

depression (Thase, Entsuah et al. 2001, Tran, Bymaster et al. 2003). Examples of SNRI medications include venlafaxine and duloxetine.

(3) **Atypical antidepressants** are another class of medications that vary greatly in their effects on neurotransmitters. For example, some atypical antidepressants such as bupropion target levels of norepinephrine and dopamine (Ascher, Cole et al. 1995) while others like mirtazapine target serotonin and norepinephrine (Stimmel, Dopheide et al. 1997). Additional examples include nefazodone, trazodone and vortioxetine.

(4) **Tricyclic and tetracyclic antidepressants**, identified by their distinct chemical structure, provide efficacy by blocking the re-uptake (absorption) of serotonin and norepinephrine. In the process, they increase the levels of these neurotransmitters in the brain (Gillman 2007). Although they were widely used before, due to side effects such as sedation, weight gain, hypotension, and effects on cardiac conduction, they are currently not the first choice for patients (Anderson 1998). Examples include amitriptyline, amoxapine, desipramine, doxepin, imipramine, nortriptyline etc.

(5) **Monoamine oxidase inhibitors (MAOIs)** inhibit the monoamine oxidase enzymes involved in reducing the levels of norepinephrine, serotonin and dopamine in the brain (Livingston and Livingston 1996, Amsterdam and Chopra 2001). As a result, the levels of these neurotransmitters are increased (Youdim, Edmondson et al. 2006, López-Muñoz and Alamo 2009). Several side effects of MAOIs have prevented their wide use including nausea, drowsiness, low blood pressure and weight gain (Evans, Davidson et al. 1982, Fallon, Foote et al. 1988, Fava 2000). MAOIs also require a specific diet since certain foods and medications can cause severely high blood pressure (Sullivan and Shulman 1984). Examples include isocarboxazid, phenelzine, selegiline and tranylcypromine.

1.4.3 Repetitive Transcranial Magnetic Stimulation (rTMS)

In rTMS, an electrical current is induced on the surface of the targeted brain region. The current is induced by a series of short magnetic pulses generated outside the brain but near the scalp (George, Wassermann et al. 1995). The dorsolateral prefrontal cortex is a common target based on evidence from neuroimaging studies suggesting impaired metabolism in this cortical region in depression (George, Wassermann et al. 1997). The stimulation of the left prefrontal cortex

through rTMS is currently approved for treatment-resistant depression by the Food and Drug Administration (Carpenter, Janicak et al. 2012, Connolly, Helmer et al. 2012) and by Health Canada in two provinces, Quebec and Saskatchewan. Although rTMS is a localized stimulation method, it is hypothesized to show therapeutic effect through top-down effects on the limbic system (striatum and amygdala) (George 2010, Fox, Buckner et al. 2012).

1.4.4 Seizure Therapy

The induction of seizures in the brain has proved to be highly effective in the treatment of severe depression with psychotic features and treatment-resistant depression. Yet its mechanism of action is not clearly understood (Fava 2003, Souery, Papakostas et al. 2006). Research suggests that the induction of seizures in brain regions and networks involved in depression may normalize their activity for therapeutic effect (Farzan, Boutros et al. 2014). Understanding the mechanism of action of successful seizure therapy could provide evidence towards the localization of depression. There are two types of seizure therapy.

(1) Electroconvulsive therapy (ECT) is the most effective treatment for treatment-resistant depression with efficacy rates around 60-70% (Thase and Rush 1995). During ECT, electrodes are applied directly to the scalp to generate electrical currents and electric fields in the brain that can trigger a brief, generalized seizure. When the current propagates beyond the site of stimulation, it is hypothesized to impact the dynamics of distributed but functionally-connected brain regions (i.e., functional brain networks) disrupted in depression (Farzan, Boutros et al. 2014). However, such non-specific stimulation of the brain can also affect the dynamics of brain networks involved in cognition, leading to the most common adverse effect of this treatment: memory impairment (Devanand, Sobin et al. 1995).

(2) Magnetic seizure therapy (MST) induces a seizure through magnetic fields. It is currently undergoing clinical trials to evaluate its efficacy and side effects relative to ECT. Although MST relies on the principles of seizure induction for therapeutic benefit, unlike ECT, the effect of MST stimulation is focal (Deng, Lisanby et al. 2011). In addition, based on the few clinical trials conducted to date, MST may improve depressive symptoms (Kayser, Bewernick et al. 2011) and suicidal ideation (Sun, Farzan et al. 2016) without the

cognitive side effects seen with ECT (Lisanby, Luber et al. 2003, Spellman, McClintock et al. 2008, Deng, McClintock et al. 2015). However, in its early stage of development, its efficacy relative to ECT requires further study (Kayser, Bewernick et al. 2011). Identifying the brain networks modified by ECT for therapeutic effect could allow the optimization of MST stimulation parameters.

1.5 Optimizing Treatments for Depression

As mentioned before, ECT has the highest efficacy for severe, treatment-resistant depression. Therefore, identifying the neurophysiological targets of ECT may help optimize other treatments for depression towards higher efficacy. Although the mechanism of action of ECT is unknown, seizures induced by ECT are known to have a significant impact on neural dynamics (Farzan, Boutros et al. 2014). In fact, the regularity (or irregularity) of neural activity is monitored in the ECT clinic to monitor seizure adequacy during treatment administration (Abrams 2002). Recent studies suggest that similar measures may also be used to characterize MST seizure characteristics (Lisanby, Luber et al. 2003, Kayser, Bewernick et al. 2011, Fitzgerald, Hoy et al. 2013, Kayser, Bewernick et al. 2013, Backhouse, Noda et al. 2018). If modulations in neural dynamics following ECT are associated with improvement in depressive symptoms, it may be possible to develop targets for treatments based on these markers of neural dynamics.

1.6 Understanding Depression and Treatments for Depression by Studying Neural Oscillations

The heterogeneous nature of depression and individual variation of symptoms seen in depression may be due to dysfunction at any of the biological levels mentioned above (i.e., cellular, molecular, network) or a combination of several levels (Belmaker and Agam 2008). Furthermore, abnormalities at one level may translate to another level. For example, dysfunctional processes at the cellular level can influence cognitive and emotional information processing at the network level (i.e., bottom-up effects). Likewise, abnormalities at the network level can translate to the level of a single cell (i.e., top-down effect) to influence neural communication (Leuchter, Hunter et al. 2015). To optimize the prediction of treatment response,

it is important to understand how treatments can target these dysfunctions. In this thesis, we investigate whether treatments for depression impact neural dynamics at the large-scale network level for therapeutic efficacy. With recent technological advancements, it is possible to directly monitor neural activity using non-invasive methods that are also cost-effective and accessible.

1.6.1 Neural Synchronization Mechanism of Action of Treatments for Depression

It still unclear how the acute neurophysiological mechanisms of treatments for depression might translate to an improvement in mood, but one well-known hypothesis is an increase in neuroplasticity (D'sa and Duman 2002, Brunoni, Lopes et al. 2008, Pittenger and Duman 2008). Neuroplasticity is the ability of the brain to form new connections or reorganize existing connections to improve neural transmission. An increase in the expression and signaling of BDNF (involved in increasing neuroplasticity in the hippocampus and cortex) is associated with antidepressant medication response (Russo-Neustadt, Beard et al. 2000, D'sa and Duman 2002). An increase in BDNF levels is also reported following ECT in the hippocampus and the cortex (Zetterström, Pei et al. 1998, Bocchio-Chiavetto, Zanardini et al. 2006, Brunoni, Baeken et al. 2014).

As outlined in **section 1.3**, previous evidence suggests that depression is associated with abnormalities at multiple biological levels from genetic to large-scale functional networks. One overarching hypothesis has linked changes seen at the cellular level to the network level. This hypothesis suggests that depression may be the result of dysfunctions in neuroplastic processes responsible for regulating synchronized neural oscillations (Leuchter, Hunter et al. 2015). Treatments in depression are suggested to show therapeutic efficacy by normalizing dysfunctions in these neuroplastic processes leading to an improved regulation of neural oscillatory activity (D'sa and Duman 2002, Leuchter, Hunter et al. 2014). Neural synchronization is necessary for effective neural communication and can occur at the level of single neurons to form microcircuits, or at a larger scale where microcircuits show synchronized activations to form large-scale neural networks (Buzsáki and Draguhn 2004, Fox, Snyder et al. 2005). Moreover, disturbances in oscillatory activity at the cellular level can translate to the network level through

bottom-up processes (Xu, An et al. 2013) and also from the network level to the cellular level through top-down processes (Engel, Fries et al. 2001, Steriade 2001, Yatham, Liddle et al. 2010, Lanzenberger, Baldinger et al. 2013). This is because the summation of individual neurons firing together generates neural oscillations and the electric field generated by network oscillations can induce electrical activity at individual neurons (Wang 2010, Anastassiou, Perin et al. 2011, Buzsáki, Anastassiou et al. 2012).

In this thesis, we focus on pharmacotherapy (escitalopram) and seizure therapy (ECT and MST) treatments for depression. Based on the above hypothesis for the mechanism of action of treatments for depression, pharmacotherapy may show therapeutic effect through the neurochemical regulation of bottom-up neuroplastic processes involved in synchronized oscillatory activity. Seizure therapy may show therapeutic effect through the neuro-electric regulation of top-down neuroplastic processes involved in synchronized oscillatory activity.

At the cellular level, pharmacotherapy can modulate neural dynamics and regulate neuroplasticity through direct or indirect effects on voltage-gated ion channels (Salomon and Cowan 2013). These channels are responsible for maintaining membrane potentials, which extend to electrical activity at the scalp recorded by electroencephalography readings (see section 1.6.2) (Buzsáki, Anastassiou et al. 2012). Therefore, factors affecting synaptic transmission at the cellular level may result in changes that modulate neuronal oscillations at the network level (Colwell 2011, Frederick, Bourget-Murray et al. 2014). Such factors may rely on the availability of presynaptic receptors for neurotransmitter re-uptake, the availability of postsynaptic receptors for signal transmission, and/or the levels of proteins and molecules that mediate synaptic transmission. At the network level, stimulation treatments such as seizure therapy are known to modulate neural dynamics (Arns, Drinkenburg et al. 2012, Olbrich and Arns 2013, Farzan, Boutros et al. 2014). The induction of a generalized seizure may significantly impact neural dynamics in the targeted region as well as other regions linked to the stimulated region (potentially associated with mood) for therapeutic effect (McNally and Blumenfeld 2004). In addition, induced neuronal oscillations are thought to modulate corticothalamic oscillations and facilitate the resetting of oscillatory network activity to regulate mood and other depressive symptoms (Paus, Sipila et al. 2001, Fuggetta and Noh 2013, Leuchter, Cook et al. 2013).

The role of neural dynamics in the mechanism of action of treatments for depression is speculative, but it is one of the few hypotheses that can bridge the effects of treatments at several levels of biological organization. Therefore, using this hypothesis as a framework, we studied several markers of neural dynamics to monitor the effects of pharmacotherapy for depression and seizure therapy for treatment-resistant depression. In addition, we evaluated whether these markers of neural dynamics can also predict response to pharmacotherapy.

1.6.2 Studying Neural Oscillations using Electroencephalography Data

The electrical potential of a single neuron or a group of neurons can fluctuate in a rhythmic pattern revealing synchronized activity (Llinás 1988, Hutcheon and Yarom 2000). When a large number of neurons oscillate together at a certain frequency, large-scale oscillations are generated (Varela, Lachaux et al. 2001, Buzsáki and Draguhn 2004) and can be detected by electroencephalography (EEG). EEG is a non-invasive recording of neural activity at the scalp and reflects cortical electrical activity (Kaiser 2007). EEG oscillations are measured at the cortex, but they are primarily generated by postsynaptic potentials (Creutzfeldt 1974). The signal intensity is extremely small and measured in microvolts. In addition, the high temporal resolution of EEG allows the proper investigation of fast-changing neural dynamics. Compared to fMRI, EEG is also affordable and accessible. For all the studies included within this thesis, EEG data was collected during the resting-state condition.

1.6.3 Advantage of the Resting-State Condition

Spontaneous neural activity is not a passive condition of the brain but rather a representation of the default functioning of the brain during the resting condition (Raichle and Snyder 2007). There are several advantages of studying neural activity during the resting-state condition. First, EEG data collected during the eyes-closed, resting-state condition allows for the examination of default neural activations without the bias and confounding effects associated with a task (Barry, Clarke et al. 2007, Van Diessen, Numan et al. 2015). In addition, since depression is associated

with internal biases in negative emotional information processing, studying neural activity during the resting-state condition can provide an assessment of self-referential neural activity (i.e., rumination and increased self-focus and self-critical nature) thought to underlie key symptoms of depression (Broyd, Demanuele et al. 2009). Such abnormalities in the default functioning of the brain may underlie the core symptoms of depression.

1.6.4 Resting-State EEG Markers of Neural Dynamics in Depression

In the following section, EEG measures of neural dynamics that were used in this thesis will be briefly introduced.

1.6.4.1 Frequency Analysis

According to Fourier's theorem, a periodic signal can be decomposed into a discrete set of sine and cosine functions, each with a specific amplitude and frequency (Bracewell and Bracewell 1986). Using this theorem, the EEG signal can be decomposed into frequency bands and the power at each of these frequency bands can be defined. The EEG signal has a bandwidth of approximately 1-50Hz, where anything <1Hz is usually voltage drift and anything >50Hz is often muscle or external noise (Malmivuo, Malmivuo et al. 1995, Niedermeyer and da Silva 2005). Over the years, studies have identified that neural oscillations at specific frequency ranges reflect different biological functionalities or components of information processing. These include delta (1-4Hz), theta (4-8Hz), alpha (8-12Hz), beta (12-30Hz) and gamma (30-50Hz) (Malmivuo, Malmivuo et al. 1995, Niedermeyer and da Silva 2005). Through power spectral density analysis of EEG, several studies have identified differences between patients with depression and healthy subjects. These differences are detailed below.

(1) **Alpha** oscillations are the most prominent oscillations seen in the resting-state, eyes-closed condition. An increase in alpha power is assumed to be associated with decreased neural activity and vice versa. Studies have shown that depression is associated with high alpha activity (i.e., decreased neural activation) (Grin-Yatsenko, Baas et al. 2010), mainly in occipital sites (Bruder, Sedoruk et al. 2008), parietal and frontal sites (Grin-Yatsenko, Baas et al. 2009, Jaworska, Blier et al. 2012). Some studies have shown increases in

absolute power (von Knorring, Perris et al. 1983, Roemer, Shagass et al. 1992, Begić, Popović-Knapić et al. 2011, Jaworska, Blier et al. 2012) while others have shown increases in relative power (John, Prichep et al. 1988, Prichep and John 1992). Some studies however, did not find any differences in alpha between patients and healthy subjects (Flor-Henry 1979, Knott and Lapierre 1987). Finally, a decrease in alpha activity was shown in patients with treatment-resistant depression compared to patients with major depressive disorder (Price, Lee et al. 2008). Interestingly, increased alpha oscillations have been associated with BDNF polymorphism (Gatt, Kuan et al. 2008, Zoon, Veth et al. 2013). Slow alpha (8-10Hz) was previously associated with thalamo-cortical activity and fast alpha (10-12Hz) was associated with cortico-cortical activity (Da Silva, Vos et al. 1980, Klimesch 1999).

(2) Depression was also associated with frontal **alpha asymmetry** (Schaffer, Davidson et al. 1983, Henriques and Davidson 1990, Henriques and Davidson 1991), where lower alpha was observed in the right prefrontal cortex (i.e., hyperactive) and higher alpha was observed in the left prefrontal cortex (i.e., hypoactive). However, not all studies have replicated this finding (Reid, Duke et al. 1998, Price, Lee et al. 2008, Carvalho, Moraes et al. 2011, Segrave, Cooper et al. 2011, Gold, Fachner et al. 2013). Asymmetry differences are suggested to be a result of abnormalities in functional connections in both the hemispheres (Fingelkurts, Fingelkurts et al. 2007).

(3) A few studies observed increased **delta** and frontal-midline **theta** oscillations in patients with depression (Nystrom, Matousek et al. 1986, Lieber and Prichep 1988, Roemer, Shagass et al. 1992, Kwon, Youn et al. 1996, Bjørk, Sand et al. 2008, Gatt, Kuan et al. 2008, Korb, Cook et al. 2008). Theta oscillations were previously associated with thalamo-cortical network activity (Klimesch 1999). Some studies have also reported decreased delta and occipital-parietal theta (Fingelkurts, Fingelkurts et al. 2006).

(4) Some studies have shown increased **beta** oscillations in patients with depression (Lieber and Prichep 1988, Knott, Mahoney et al. 2001).

Studies have also investigated differences between patients with depression and healthy subjects using frequency analysis in source space (i.e., current source density analysis). Current source density is an estimate on the current sources generating the electrical potentials measured by

EEG and is derived for each frequency band separately. Some studies associate depression with higher delta, theta, alpha and beta in bilateral frontal areas (Korb, Cook et al. 2008), higher theta in the anterior cingulate cortex (Pizzagalli, Pascual-Marqui et al. 2001), higher beta in the inferior and superior right frontal gyrus (Pizzagalli, Nitschke et al. 2002) and higher whole-brain delta (Pizzagalli, Oakes et al. 2003). Others show lower delta in the inferior and superior right temporal gyrus (Lubar, Congedo et al. 2003) and lower delta, theta and beta in the anterior cingulate cortex (Mientus, Gallinat et al. 2002).

Coherence is used to show functional links between different recording sites (areas of the brain) at each relevant frequency band on a topographical scalp map (Nunez, Srinivasan et al. 1997, Nunez, Silberstein et al. 1999). Depression in male subjects was previously associated with decreased coherence in the delta, theta, alpha and beta bands in several regions of the brain (Knott, Mahoney et al. 2001). Another study also showed a decrease in coherence in the theta, alpha and beta bands, however, these changes were not significant (Suhhova, Bachmann et al. 2009). Interhemispheric coherence was shown to be significantly lower in patients with depression in the delta, alpha, theta or beta bands during sleep (Armitage, Hoffmann et al. 1999, Knott, Mahoney et al. 2001).

1.6.4.2 Time-Domain Analysis

Variation in the pattern of neural oscillatory activity over time can be studied through linear or non-linear measures. Linear analysis methods for EEG data may include power spectral density analysis using the Fourier transform (outlined above). Although useful, linear measures may not be sufficient for the characterization of neural dynamics. Several lines of evidence suggest that the brain is a non-linear dynamical system and neural activity measured from the brain is complex, chaotic and unpredictable. Non-linear measures of the EEG signal may provide new insights into neural functions that are not seen with linear analysis methods. Studies that applied non-linear analysis to compare patients with depression with healthy controls used wavelet-chaos methodology (Ahmadlou, Adeli et al. 2012), wavelet entropy (Li, Li et al. 2007), Higuchi's fractal dimensions (Ahmadlou, Adeli et al. 2012, Bachmann, Lass et al. 2013, Cukic, Pokrajac et al. 2018), largest Lyapunov exponents (Hosseinifard, Moradi et al. 2013), Lempel-Ziv

complexity (Bachmann, Kalev et al. 2015), detrended fluctuation analysis (Lee, Yang et al. 2007), sample entropy (Cukic, Pokrajac et al. 2018), multiscale entropy (Méndez, Zuluaga et al. 2012). In general, these studies revealed higher complexity of neural activity in patients with depression.

1.6.4.3 Global Network Analysis

Global neural network dynamics can be studied with resting-state EEG data through microstate analysis. Microstate analysis postulates that spontaneous neural activity (i.e., resting-state) is not random and the topographic distributions of this spontaneous activity can be clustered into a set of brain state maps (i.e., microstates). Each microstate is stable for a short duration of time (50-120ms) before transitioning into another microstate. Microstates are thought to be generated by an underlying neuronal population. Therefore, the temporal characteristics of a microstate (such as rate of change or duration) can be considered as an expression of the dynamic stability of an underlying spatial network (Brodbeck, Kuhn et al. 2012). The duration of microstates is also consistent with the duration of high-level cognitive processes, as shown by evoked-potential studies (Kok 1997). Moreover, microstates were shown to be state-dependent, to vary across age, cognitive state (Koenig, Prichep et al. 2002, Brodbeck, Kuhn et al. 2012, Milz, Faber et al. 2016, Santarnecchi, Khanna et al. 2017) and in response to therapy (Kinoshita, Strik et al. 1995, Rodriguez, Vitali et al. 2002, Kikuchi, Koenig et al. 2007). Studies have also confirmed the reliability of microstates across repeated testing sessions (Khanna, Pascual-Leone et al. 2014). Microstates were previously linked with resting-state fMRI networks (Britz, Van De Ville et al. 2010, Musso, Brinkmeyer et al. 2010, Yuan, Zotev et al. 2012); specifically with networks suggested to be impaired in depression such as the salience network and the frontoparietal network (Veer, Beckmann et al. 2010, Whitfield-Gabrieli and Ford 2012, Kaiser, Andrews-Hanna et al. 2015, Fischer, Keller et al. 2016). To date, only one study examined and reported a decrease in duration of microstates in depression compared to healthy controls (Strik, Dierks et al. 1995).

1.6.5 Resting-State EEG Markers of Modulations in Neural Dynamics by Treatments of Depression

Neural dynamics, as measured by resting-state EEG, are potentially modulated by treatments of depression. These are summarized below.

1.6.5.1 Frequency Analysis

In the EEG frequency domain, several studies investigated response to antidepressants using pre-treatment neural dynamics.

(1) High **delta** oscillations were linked with response to SSRI (imipramine), tricyclic (paroxetine) antidepressants (Knott, Telner et al. 1996, Knott, Mahoney et al. 2000).

(2) High **theta** oscillations were linked with response to SSRI (paroxetine, venlafaxine, etc.) (Knott, Mahoney et al. 2000, Iosifescu, Greenwald et al. 2009) and tricyclic (imipramine) antidepressants (Knott, Telner et al. 1996). Specifically, high **frontal midline theta** was associated with response to several antidepressants (Spronk, Arns et al. 2011).

(3) **Theta cordance** combines absolute and relative power from EEG signal into a single measure (Leuchter, Cook et al. 1994) and is thought to represent regional cerebral perfusion (Leuchter, Uijtdehaage et al. 1999). An early decrease in frontal theta cordance (2 days to 1 week after treatment) was observed in responders of SSRI (fluoxetine) and SNRI (venlafaxine) (Cook and Leuchter 2001, Cook, Leuchter et al. 2002, Bares, Brunovsky et al. 2008).

(4) Studies have shown that high **alpha** oscillations were associated with tricyclic (clomipramine, imipramine) (Ulrich, Haug et al. 1988, Knott, Telner et al. 1996) and SSRI (paroxetine, fluoxetine) antidepressant response (Knott, Mahoney et al. 2000, Bruder, Stewart et al. 2001). High alpha in pre-treatment data is hypothesized to be a result of low levels of serotonin seen in depression (Bruder, Sedoruk et al. 2008). Antidepressants likely target this deficit by increasing levels of serotonin.

(5) Interhemispheric **alpha asymmetry** was also observed in responders of tricyclic (amitriptyline and pirindol) (Ulrich, Renfordt et al. 1984) and SSRI (fluoxetine) antidepressants (Bruder, Stewart et al. 2001, Bruder, Sedoruk et al. 2008). Higher alpha activity was observed in the left hemisphere and lower alpha activity was observed in the right hemisphere, mainly in occipital sites for responders. Non-responders reveal the

opposite activation in frontal and parietal regions. Causes for this asymmetry may rise from the genetic level (serotonin receptor polymorphisms) (Bismark, Moreno et al. 2010) or from the cellular neurotransmitter level (lateralized distribution of serotonin in the brain in patients with depression) (Bruder, Stewart et al. 2001).

Few studies have studied the long-term effects of antidepressant medications on resting-state EEG band power and provide conflicting results (Tarn, Edwards et al. 1993, Kwon, Youn et al. 1996, Knott, Mahoney et al. 2002, Bruder, Sedoruk et al. 2008). Several longitudinal studies however, have quantified changes in EEG band power following ECT. Studies have shown an increase in delta and theta oscillations following ECT (Kriss, Halliday et al. 1978, Sackeim, Luber et al. 1996).

Current source density analysis also revealed that increased **theta** activity in the rostral anterior cingulate cortex (Pizzagalli, Pascual-Marqui et al. 2001, Mulert, Juckel et al. 2007, Korb, Hunter et al. 2009, Korb, Hunter et al. 2011, Hunter, Korb et al. 2013) and medial orbitofrontal cortex (Korb, Hunter et al. 2009) at baseline was associated with response to medications such as SSRIs (Arns, Etkin et al. 2015) and tricyclic antidepressants (such as nortriptyline, citalopram, reboxetine, fluoxetine, or venlafaxine) (Korb, Hunter et al. 2009, Pizzagalli 2011) and also ECT (McCormick, Yamada et al. 2009). Specifically, increase in theta oscillations in the anterior cingulate cortex was shown to increase the activation of this region (Pizzagalli, Oakes et al. 2003).

1.6.5.2 Time-Domain Analysis

Very few studies have evaluated the association between non-linear EEG features (e.g., complexity) and response to antidepressants such as citalopram, clomipramine, escitalopram, bupropion and mirtazapine (Thomasson, Pezard et al. 2000, Méndez, Zuluaga et al. 2012, Okazaki, Takahashi et al. 2013, Farzan, Atluri et al. 2017, Jaworska, Wang et al. 2018). Multiscale entropy is one method of complexity analysis. It provides information on short and long temporal scales at various regions of the brain revealing information on the complexity of local and global neuronal processing (Vakorin, Lippé et al. 2011, McIntosh, Vakorin et al. 2013,

McDonough and Nashiro 2014). To our knowledge, only two studies have investigated whether pre-treatment neural complexity can predict response to antidepressant medication in depression (Méndez, Zuluaga et al. 2012, Jaworska, Wang et al. 2018). In *Mendez et al.,* a reduction in Lempel-Zev complexity of neural signals, as measured by magnetoencephalography, was shown to be associated with response to mirtazapine. Lower complexity at baseline was also associated with antidepressant response in their study. In comparison, *Jaworski et al.* used multiscale entropy on resting-state EEG data to predict response to escitalopram and/or bupropion in 36 patients with major depressive disorder. Results from the study revealed that increased baseline complexity in mid-coarse timescales (frontal, central, parietal) and decreased complexity in fine timescales (fronto-central) was associated with improvement in depressive symptoms. Another study investigated changes in complexity following ECT in 3 patients and showed reduction in complexity (Okazaki, Takahashi et al. 2013).

1.6.5.3 Global Network Analysis

To the best of my knowledge, the effect on global network dynamics by treatments of depression using resting-state EEG has not been studied previously and is a novel contribution of this thesis.

1.7 Review on the Performance of EEG Predictors for Antidepressant Response

The following table (**Table 1.1**) summarizes studies that have investigated the predictive performance of resting-state EEG features for antidepressant response. Accuracy was reported to be 87.9% when linear (EEG power) and non-linear features (mutual information) were combined with machine learning methods (Khodayari-Rostamabad, Reilly et al. 2013). However, this study was performed with a very low sample size (n=22) and responders were defined to have ≥30% improvement in HRSD-17 scores rather than the usual ≥50%. In addition, the sample size was a combination of patients on 4 different medications (sertraline, citalopram, fluvoxamine and paroxetine). The study also includes multiple epochs from the same individual as separate samples in the feature matrix and this can significantly impact the prediction accuracy of the model (Saeb, Lonini et al. 2016). Accuracy was also high (85-92%) in Rabinoff et al., (2011)

(Rabinoff, Kitchen et al. 2011) using spectral EEG features with classification and regression tree analysis. The study combined trials for 2 antidepressants (fluoxetine and venlafaxine) to predict response in 51 patients with unipolar depression. The high accuracy values however, may be due to overfitting to the data and this is suggested by the 100% specificity in all treatment groups. Remaining studies revealed accuracies between 60-77%. Many of these studies evaluated the prediction performance of single marker.

Table 1.1 - Summary of Studies Evaluating the Prediction Performance of Pre-treatment Resting-State EEG markers for Antidepressant Response

Citation	Sample Size of Patients	Treatment	Model	Features	Validation Method	Classification Performance
Bruder et al., (2008). *Biological Psychitatry* (Bruder, Sedoruk et al. 2008)	18 patients with major depressive disorder Response = "much improved" or "very much improved" on the Clinical Global Impression Improvement scale	12 weeks of fluoxetine treatment	ROC analysis/ Logistic regression	Alpha power and alpha asymmetry at occipital sites	none	Alpha power Accuracy = 65.1% Sensitivity = 72.7% Specificity = 57.5% Precision = 72.7% Alpha asymmetry Accuracy = 67.5% Sensitivity = 63.6% Specificity = 71.4% Precision = 77.8% Combined alpha power and alpha asymmetry Accuracy = 75.5% Sensitivity = 83.3% Specificity = 67.7% Precision = 71.4%
Iosifescu, et al., (2009). *European Neuropsycho-pharmacology* (Iosifescu, Greenwald et al. 2009)	82 patients with major depressive disorder Response: ≥50% reduction in HRSD-17 scores	8-week treatment trial escitalopram (n=53), fluoxetine (n=7), paroxetine (n=7), citalopram (n=5), sertraline (n=5), venlafaxine (n=5)	ROC analysis/ Logistic regression	Antidepressant Treatment Response (ATR) index using EEG parameters assessed at baseline and week 1. 4-channel EEGs (F7-Fpz, F8-Fpz, A1-Fpz, A2-Fpz)	none	Baseline relative theta power Accuracy = 63% Sensitivity = 64% Specificity = 62% Relative theta power at week 1 Accuracy = 60% Sensitivity = 57% Specificity = 61% ATR index Accuracy = 70% Sensitivity = 82% Specificity = 54%
Korb et al.,	72 patients with	Subjects randomized	ROC analysis/	Theta current source	none	High theta in rACC

(2009). *Clinical Neurophysiology* (Korb, Hunter et al. 2009)	major depressive disorder Response: ≥50% reduction in HRSD-17 scores	to receive 8-weeks of fluoxetine (n=13), venlafaxine (n=24), or placebo (n=35)	Logistic regression	density in rostral anterior cingulate cortex (rACC) and the medial orbitofrontal cortex (mOFC)		Accuracy = 65.5% Sensitivity = 64 %, Specificity = 67% High theta in mOFC Accuracy = 66.5% Sensitivity = 73% Specificity = 60%
Leuchter, et al., (2009). *Psychiatry Research*. (Leuchter, Cook et al. 2009)	220 patients with major depressive disorder Response: ≥50% reduction in HRSD-17 scores Remission: HRSD-17 score <= 7	Only escitalopram patients taken into model (n = 73) 7 weeks of treatment	ROC analysis/ Logistic regression	Antidepressant Treatment Response (ATR) index in alpha and theta bands of frontal brain activity integrated and scaled from 0 (low probability of response or remission to the medication) to 100 (high probability)	ROC curve	ATR Predicting Response Accuracy = 74% Sensitivity = 58% Specificity = 91% Precision = 88% ATR Predicting Remission Accuracy = 74% Sensitivity = 61% Specificity = 82% Precision = 68%
Bares, et al., (2010). *European Neuropsycho-pharmacology* (Bares, Brunovsky et al. 2010)	18 patients with major depressive disorder Response: ≥50% reduction in MADRS scores	4-week bupropion treatment for patients who had failed to respond to previous antidepressant treatments	ROC analysis/ Logistic regression	QEEG theta cordance computed at three frontal electrodes	none	Reduction of prefrontal theta cordance after one week of bupropion Positive predictive value = 0.9 Negative predictive value = 0.75
Rabinoff et al., (2011). *Open Medical Informatics Journal* (Rabinoff, Kitchen et al. 2011)	51 patients with unipolar depression clinical response was defined as reduction in final HRSD score to ≤10	8-weeks of fluoxetine or venlafaxine n=24: fluoxetine n=27: venlafaxine	classification and regression tree analysis with cost-complexity pruning	Absolute power, relative power and cordance values in four frequency bands (0.5–4 Hz, 4–8 Hz, 8–12 Hz, and 12–20 Hz)	10-fold cross-validation	Fluoxetine combined with Venlafaxine: Accuracy = 92.5% Sensitivity = 85% Specificity = 100% Precision = 100% Fluoxetine:

						Accuracy = 85.5% Sensitivity = 71% Specificity = 100% Precision = 100% Venlafaxine: Accuracy = 91.5% Sensitivity = 83% Specificity = 100% Precision = 100% Placebo: Accuracy = 82.5% Sensitivity = 90% Specificity = 75% Precision = 69%
Tenke et al., (2011). Biological Psychiatry (Tenke, Kayser et al. 2011)	41 patients with major depressive disorder Response = "much improved" or "very much improved" on the Clinical Global Impression Improvement scale	N=16 (SSRI only) N=15 (SSRI + NDRI) N=10 (SNRI) 8-12 weeks of treatment	ROC analysis/ Logistic regression	Current source density in whole brain	none	Prominent alpha activity Accuracy = 71.2% Sensitivity = 50% Specificity = 92.3% Precision = 93.3%
Khodayari-Rostamabad, et al., (2013). *Clinical Neurophysiology* (Khodayari-Rostamabad, Reilly et al. 2013)	22 patients with treatment-resistant depression Response: ≥30% reduction in HRSD-17 scores	6 weeks of SSRI treatment Mainly Sertraline hydrochloride but other SSRIs also used.	mixture of factor analysis (MFA) model	EEG power spectral density Spectral coherence Mutual information	''leave-n-out'' randomized permutation cross-validation	Specificity = 80.9% Sensitivity = 94.9% Accuracy = 87.9%
Baskaran et al.,	44 patients with	8-week escitalopram	ROC analysis/	EEG measures	none	Baseline whole-

(2017). *Journal of Affective Disorders* (Baskaran, Farzan et al. 2017)	major depressive disorder Response: ≥50% reduction in MADRS scores	treatment	Logistic regression	assessed at baseline, 2 weeks into treatment, and as an early change (from baseline to week 2)		brain absolute alpha asymmetry Accuracy = 70.7% Sensitivity = 72.2% Specificity = 69.2% Precision = 61.9% Baseline parietal alpha asymmetry Accuracy = 77.2% Sensitivity = 88.9% Specificity = 65.4% Precision = 64.0% Baseline whole-brain relative delta asymmetry Accuracy = 73.3% Sensitivity = 88.9% Specificity = 57.7% Precision = 59.3% 2 week whole-brain absolute delta asymmetry Accuracy = 67.8% Sensitivity = 77.8% Specificity = 57.7% Precision = 56.0%

1.8 Summary

In summary, the trial-and-error process of identifying treatments for patients with depression is ineffective and inefficient. As summarized in above sections, depression is characterized by dysfunctions at several levels of brain function. Given this understanding, neurophysiological markers of treatment response may improve the success rate of treatments for depression. In this thesis, we investigate changes in several characteristics of neural dynamics following seizure therapy (ECT and MST) and pharmacotherapy (escitalopram). We hypothesize both these treatments impact neural dynamics for therapeutic efficacy. Seizure therapy for example, may impact neural dynamics at the network level but may also affect other levels through top-down processes. In contrast, pharmacotherapy has an impact at the cellular level but, through bottom-up processes, it may also affect neural dynamics at the network level. We specifically focus on EEG markers of neural dynamics during the resting-state condition for two reasons: (1) EEG signal has the high temporal resolution needed to monitor changes in neural dynamics following treatments of depression, and (2) EEG has high potential for clinical translation. In addition to investigating the mechanism of treatments, we evaluated the predictive value of neural dynamic markers for pharmacotherapy response. Identifying early markers of response to medications may reduce the time spent in failed trials and avoid the debilitating impact of untreated depression.

This thesis is written as a multi-paper thesis. The results are split into three sections, each containing manuscript(s) that detail the work: Section II: Investigating the targets of seizure therapy using EEG markers of neural dynamics; Section III: Investigating the targets of pharmacotherapy using EEG measures of neural dynamics; and Section IV: Evaluating the predictive value of EEG measures of neural dynamics for response to pharmacotherapy. A general discussion and suggestions for future work are provided in Section V.

Chapter 2 – Rationale, Objectives and Hypothesis

There are two main objectives. The first is to use measures of neural dynamics to investigate neurophysiological targets of successful treatments for depression using measures of neural dynamics. The second is to investigate whether measures of neural dynamics can also predict response to treatment. We consider two types of treatments (1) seizure therapy for severely-ill, treatment-resistant patients with depression, and (2) pharmacotherapy for patients with major depressive disorder.

2.1 Study 1: Investigating Targets of Seizure Therapy

2.1.1 Rationale

Electroconvulsive therapy (ECT) is highly effective for treatment-resistant depression, yet its mechanism of action is still unclear. Magnetic seizure therapy (MST) (Lisanby, Luber et al.) also relies on the principles of seizure induction for therapeutic benefit but unlike ECT, the effect of MST is localized (Deng, Lisanby et al. 2011). Based on the few clinical trials conducted to date, MST improves depressive symptoms (Kayser, Bewernick et al. 2011, Cretaz, Brunoni et al. 2015, Kayser, Bewernick et al. 2015) and suicidal ideation (Sun, Farzan et al. 2016), without the cognitive side effects seen with ECT (Lisanby, Luber et al. 2003, Moscrip, Terrace et al. 2006, Spellman, McClintock et al. 2008, Deng, McClintock et al. 2015). However, in its early stage of development, its efficacy relative to ECT requires further study (Kayser, Bewernick et al. 2011). Although the mechanism of action of ECT is unknown, seizures induced by ECT are known to have a significant impact on neural dynamics (Farzan, Boutros et al. 2014). Identifying how neural dynamics (i.e., neural communication) are affected by treatment-resistant depression and modified by ECT may allow the optimization of MST as well as the development of non-invasive and non-seizure inducing treatments.

2.1.2 Objectives

The primary objective of the first study was to investigate the targets of successful seizure therapy using EEG measures of neural dynamics.

The secondary objective was to compare the targets of ECT and MST using EEG measures of neural dynamics.

2.1.3 Hypothesis

We hypothesize that seizure therapy will have a significant impact on neural dynamics for therapeutic efficacy. Further, we hypothesize that ECT will have a global effect on neural dynamics while MST will show specific effects on neural dynamics on networks and regions close to the area of stimulation (i.e., the dorsomedial prefrontal cortex).

2.2 Study 2 – Investigating the Targets of Pharmacotherapy

2.2.1 Rationale

Pharmacotherapy is often the first line of treatment for patients with major depressive disorder, yet remission rates are around 30% for the first medication, and decline progressively with subsequent medication trials (Rush, Trivedi et al. 2006, Trivedi, Rush et al. 2006). A trial-and-error process is implemented to identify the antidepressant medication best suited for each patient, but this process can take time and patients may spend months to years suffering from symptoms (Solomon, Keller et al. 1997). The average efficacy of antidepressants may be improved with better insight into the long-lasting neurophysiological changes that occur following successful treatment.

2.2.2 Objectives

The primary objective of the second study was to investigate the targets of pharmacotherapy (specifically escitalopram) for major depressive disorder using EEG measures of neuronal dynamics.

The secondary objective was to compare the targets of pharmacotherapy with seizure therapy.

2.2.3 Hypothesis

We hypothesize that pharmacotherapy will have specific effects on neural dynamics that will distinguish responders and non-responders. In responders, these effects will be seen in regions known to be impaired in depression such as the ACC, cingulate cortex, etc. (Pandya, Altinay et

al. 2012). Further, we hypothesize that the neural dynamics seen in non-responders of pharmacotherapy may be normalized by seizure therapy.

2.3 Study 3 – Predicting Response to Pharmacotherapy

2.3.1 Rationale

To reduce the time spent in failed trials and avoid the debilitating impact of untreated depression (i.e., poor quality of life, economic burden), early predictors of pharmacotherapy must be identified. Implementing a personalized tool for the prediction of response to antidepressant medications may lead to increase in the efficacy of current treatments and the faster relief of symptoms.

2.3.2 Objectives

The objective of the third study is to evaluate EEG measures of neural dynamics for the prediction of response to pharmacotherapy (escitalopram). In this study, we aim to develop a model for antidepressant response prediction.

2.3.3 Hypothesis

We hypothesize that markers of neural dynamics will have high predictive value for the prediction of response to escitalopram.

Chapter 3 – General Methods

In this chapter, we provide a general overview on data collection, data preprocessing and the data analysis methods used in this thesis. A brief overview of the machine learning method used to create a prediction model is also included. Detailed methods (including statistical analysis) for each project of the thesis are provided within each manuscript/chapter. In addition, details on the EEG measures of neural dynamics (frequency analysis, multiscale entropy analysis and microstate analysis) are also provided within each manuscript/chapter.

3.1 Data Collection

Two types of data were collected in the experiments: (i) clinical data and (ii) neurophysiological data. Clinical assessments on the severity of depression are performed by psychiatrists or medical professionals over the duration of the treatment. In some of the studies included in this thesis, behavioral data was also collected (i.e., cognitive outcomes). Clinical measures are used in the clinic to assess a patient's mental health, monitor the effect of treatment and to make clinical decisions such as diagnosis, treatment selection or modifications to treatment dosage. Neurophysiological data can be used to monitor changes in neural activity with treatment. Understanding how changes in the severity of depression can be linked to changes in neural activity may help improve existing treatments for depression or develop novel personalized treatments. In this thesis, neurophysiological data refers to resting-state, eyes-closed electroencephalography (EEG) data. Patients were instructed to close their eyes for a short period of time (between 5-10 minutes for the studies) without falling asleep. Details on the equipment and equipment settings for data collection are included within each manuscript.

Data for this thesis were provided by multiple studies. The following sections outline the research protocols for these studies as well as the specific data that was collected in each study. Details on the demographics and clinical data of the patients included in our analysis are provided within each manuscript and are therefore not discussed here.

3.1.1 Data for Study 1: Investigating the Targets of Seizure Therapy

EEG data was collected from two studies, each administering different seizure therapy treatments: electroconvulsive therapy and magnetic seizure therapy. Eyes-closed rest EEG data was collected within a week prior to the first treatment session and again within 2 weeks after the completion of the last treatment. Healthy subject data was pooled from several other studies. Details are provided below.

3.1.1.1 Electroconvulsive Therapy Trial (Open-Label)

Data used from this trial was collected from 2009 to 2014. A total of 60 patients with treatment-resistant depression/bipolar disorder were recruited for this trial. However, not all of these patients completed the treatment trial and neurophysiological data could not be collected from all of these patients. We also only included subjects with treatment-resistant depression in our analysis. The following information was taken directly from the research protocol of this trial.

3.1.1.1.1 Participants

Patients were included if they: (1) were voluntary and competent to consent to treatment, (2) had a Structured Clinical Interview for DSM-IV (SCID) (American Psychiatric Association 2000) confirmed DSM-IV diagnosis of major depressive disorder (3) were between the ages of 18 and 65, (4) had treatment-resistant depression (i.e., failed to achieve a clinical response, or did not tolerate, at least 2 separate antidepressant trials of sufficient dose for at least 6 weeks according to Stage II criteria outline by (Thase and Rush 1995), (5) had a score of greater than or equal to 20 on a HRSD-17 scale and (6) had no increase or initiation of new antidepressant (or other psychoactive) therapy in the 4 weeks prior to screening. A careful medical and neurological history was taken to ensure that subjects had no unstable conditions that would preclude them from entering into the study. This history focused on conditions such as seizures, stroke, hypertension, diabetes, coronary artery disease, thyroid problems, respiratory illness, allergies and presence of metal implants.

Patients were excluded if they: (1) had a history of DSM-IV substance dependence in the last 6 months, and have DSM-IV substance abuse in the last month, (2) had a concomitant major

unstable medical or neurologic illness or have had a history of seizures, (3) were acutely suicidal, (4) were pregnant, (5) had metal implants, (6) had a co-morbid borderline personality disorder and/or antisocial personality disorder as confirmed by the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II), which may prevent the patient from completing the procedures required for the study; (7) positive urine toxicology screen for drugs of abuse. With respect to concomitant medications patients were excluded if during the time of treatment (or in the last 4 weeks before treatment) they received: (1) more than 2 mg daily dose of lorazepam (or equivalent).

3.1.1.1.2 Clinical Measures

Demographic variables and potential covariates were recorded at baseline following a clinical interview. These included the duration of the current episode, years from first diagnosis, number of previous episodes, type and dose of current and previous treatment and family history of mood disorder. Clinical measures were collected at baseline and at the end of treatment. The primary outcome variable was the 17-item HRSD scale. Other outcome measures included the Montgomery-Asberg Depression Rating Scale (MADRS), Beck Depression Inventory (BDI).

3.1.1.1.3 Treatment

ECT was administered with a square-wave, constant-current, brief-pulse device (MECTA Corporation, Lake Oswego, OR). ECT was administered open-label three times per week on Mondays, Wednesdays and Fridays. First, seizure threshold was determined at the first treatment using the previously published titration procedure (Sackeim, Decina et al. 1987). For all subsequent treatments, stimulus intensity was delivered at either 1.5 times the seizure threshold (bilateral ECT) or 6 times the seizure threshold (right unilateral ECT). Patients either received right unilateral ECT or bilateral ECT, and the electrodes were placed according to guidelines outlined by the American Psychiatric Association (American Psychiatric Association 2001). Thiopental and succinylcholine were the typical anaesthetic medications used, at doses determined by the anesthetist. Treatment termination was based solely on clinical factors and/or if the patient expressed wish to discontinue. Subjects were free to withdraw consent for treatment at any time.

3.1.1.2 Magnetic Seizure Therapy Trial (Open-Label)

Data used from this trial was collected from 2012 to 2017. A total of 154 patients with major depressive disorder were recruited for this trial till date (we only included subjects with major depressive disorder in our analysis). However, not all of these patients completed the treatment trial or neurophysiological data could not be collected.

The following information was taken directly from the research protocol of this trial.

3.1.1.2.1 Participants

Participants were included if they: (1) had a DSM-IV diagnosis of a major depressive episode with or without psychotic features in the context of major depressive disorder, (2) were within the age range from 18-85, (3) had a 24-item HRSD score of $\geq$21 (depression patients, moderate–severe), and (4) were on a medically acceptable form of birth control (if a woman of child-bearing potential).

Participants were excluded from the trial if they: (1) had an unstable medical and/or neurological condition, (2) were pregnant or lactating, (3) were not considered sufficiently well to undergo general anesthesia, (4) had a cardiac pacemaker, cochlear implant, implanted electronic device or non-electric ferrometallic implant in the head only, (5) were taking a benzodiazepine at a dose greater than lorazepam 2mg or equivalent, (6) were taking any non-benzodiazepine anticonvulsant, (7) had active substance misuse or dependence within the past 3 months, (8) had a diagnosis of delirium, dementia or another cognitive disorder secondary to a general medical condition, (9) had other significant Axis I co-morbidity, (10) had a co-morbid borderline personality disorder and/or antisocial personality disorder as confirmed by the Structured Clinical Interview for DSM-IV Axis II Disorders (SCID-II), or (11) had a history of any suicide attempts in the past 6 months.

3.1.1.2.2 Clinical Measures

Demographic variables were recorded at baseline following a clinical interview. These included the duration of the current episode, years from first diagnosis, number of previous episodes, type and dose of current and previous treatment and family history of mood disorder. Diagnosis was

assessed with the SCID (DSM-IV). Clinical rating measures included the 24-item HRSD for consistency with most prior ECT studies. The Beck Scale for Suicide Ideation (BSS) (Beck, Kovacs et al. 1979) was used to evaluate suicidal ideation, which is a common symptom in depression.

Remission was defined as a 24-item HDRS score ≤ 10, and a greater than 60% decrease in scores from baseline. Response was defined as a ≥50% reduction in 24-item HDRS score from baseline. The Montreal Cognitive Assessment (MoCA) was administered at baseline and every 6 treatments during acute treatment. The MoCA was also administered before the last treatment.

3.1.1.2.3 Treatment

The frequency of stimulation was between 20 Hz and 100Hz, with a duration range between 2 and 20 seconds depending on the frequency used. Furthermore, the anatomical location of stimulation was either the frontal or vertex region of the brain. The MST determination of seizure threshold was done at 100% stimulator output applied at the selected treatment frequency with progressively escalating train durations until an adequate seizure was produced. During an ECT treatment an adequate seizure is described as generalized tonic-clonic activity ≥ 20 seconds on the EMG recording or ≥ 25 seconds of EEG seizure activity. However, little data is available on the characteristics of MST induced seizures therefore the adequacy of the seizure was determined at each session by the treating MST psychiatrist. During titration, a maximum of three stimulations were given at the same session, provided the coil temperature allowed for a third stimulation. If an adequate seizure was not produced by the third stimulation, titration was continued at the next treatment session until threshold was reached.

Six treatment sessions, at a frequency of two or three times per week were administered. Patients did not have treatments on consecutive days. Remission was assessed every 3rd treatment, and if the pre-defined remission rate was not met 3 additional treatments were provided. This was repeated a total of 5 times (i.e., maximum number of treatment was 24). Furthermore, response during the acute treatment course was monitored and the dose adjusted accordingly. That is, if the patient failed to achieve an equal to or greater than 30% decrease from baseline following treatment 3, the dose was increased on their 4th treatment. After treatment 6, if the patient failed

to achieve further response, that is an equal or greater than 30% decrease from the score after treatment 3, the dose was increased on their 7th treatment. Treatment continued in this manner up to a maximum of 24 total treatments. If the patient was already at maximum stimulation (20 seconds or 1000 pulses), the treatment continued with the dose unchanged.

3.1.1.3 Healthy Subjects

Healthy subject data was pooled from several trials conducted at the Temerty Centre for Therapeutic Brain Stimulation (Centre for Addiction and Mental Health, Toronto, ON). Healthy subjects in all these trials were screened and only included if they did not have a previous history of psychosis.

3.1.2 Data for Study 2: Investigating the Targets of Pharmacotherapy

Data was provided by phase 1 of the Canadian Biomarker Integration Network in Depression (CAN-BIND) study. The following information is taken from our manuscript in *Journal of Affective Disorders* (Baskaran, Farzan et al. 2017):

3.1.2.1 Participants

"Participants were outpatients aged 18–60 years of age, and met DSM-IV-TR criteria for major depressive episode, confirmed by the Mini International Neuropsychiatric Inventory (MINI) (Sheehan, Lecrubier et al. 1997). Study procedures were approved by research ethics institutional review boards at each centre and all participants signed written informed consent prior to participation. At study enrollment, all participants were experiencing a major depressive episode duration ≥ 3 months with a Montgomery Asberg Depression Rating Scale (MADRS) score ≥ 24; and were free of psychotropic medications for at least 5 half-lives before baseline visit. Participants were excluded if they had any Axis I diagnosis, other than major depressive disorder, that was considered the primary diagnosis or if they had a diagnosis of Bipolar Disorder Type I or II. Presence of a significant Axis II diagnosis (borderline, antisocial) was also exclusionary, along with high suicidal risk, substance dependence/ abuse in the past 6 months, and presence of significant neurological disorders, head trauma or other unstable medical

conditions. Female participants who were pregnant or breastfeeding were also excluded. Other exclusionary criteria included having failed four or more adequate pharmacological interventions, having started psychological treatment within the past 3 months with the intent of continuing the treatment, previously having failed escitalopram treatment or showing intolerance to escitalopram, and being at risk for hypomanic switch (i.e. with a history of antidepressant induced hypomania)."

3.1.2.2 Clinical Measures

"Participants were assessed every 2 weeks throughout the study period (8 weeks) including baseline (before administration of study medication). The primary outcome measure was the change in MADRS from baseline to week 8 of the study. Response was defined as a $\geq$ 50% decrease in MADRS score."

3.1.2.3 Treatment

"Escitalopram was administered in an open-label manner, starting at 10 mg daily, which was increased to 20 mg daily at week 2 or later if clinically necessary. For patients who were unable to tolerate the 20 mg dose, the dose could be reduced to 10 mg at the discretion of the treating psychiatrist."

3.1.2.4 EEG Recording

"Three compatible EEG acquisition systems were used across study sites. At UHN, recordings were performed using a Biosemi Active-Two amplifier system (Biosemi, Amsterdam, The Netherlands) from 64 channels using Ag/AgCl electrodes (active electrodes) mounted on an elastic cap. Eight additional electrodes were placed below the hairline (both mastoids, both pre-auricular points, outer canthus of each eye, and inferior orbit of each eye). Eye movements were recorded with the electrodes placed at the outer canthi (horizontal electrooculogram (EOG)) and at the inferior orbits (vertical EOG). Two further electrodes (Common Mode Sense [CMS] active electrode and Driven Right Leg [DRL] passive electrode) were used as reference and ground electrodes, respectively (cf. www.biosemi/faq/cms&drl.htm). Data were collected with a sampling rate of 512 Hz with a low-pass cut-off 102.4 Hz. At CAMH, EEG was recorded with a 64-channel electrode cap with Ag/AgCl electrodes using a Neuroscan Synamps RT amplifier

system (Compumedics Neuroscan USA, Ltd. Charlotte, North Carolina, USA). Data were digitized at 1000 Hz. Electrodes on the supra-orbital ridges and external eye canthi monitored EOG activity. The electrode posterior to Cz served as the reference electrode. Data were recorded with an online filter of 0.05–100 Hz. At UBC, EEG was recorded using a QuickAmp amplifier (Brain Products, Gilching, Germany) from a 64-channel electrode cap with Ag/ AgCl electrodes. Data were digitized at 1000 Hz. Electrodes on the supra-orbital ridges and external eye canthi monitored EOG activity. A common average of electrodes was used as the reference. Data were recorded with an online filter of 0.01–499 Hz." At QNS, EEG was recorded using a Geodesic sensor net (EGI, Eugene, USA) from a 128-channel electrode cap. Data were digitized at 1000 Hz. A common average of electrodes was used as the reference. For all sites, electrode placement was in accordance with the International 10–10 System. Impedance levels were set at less than 5 kOhm. When examining the electrode montages across data acquisition sites, 58 common electrodes were identified.

3.1.2.5 EEG Data Standardization

As mentioned, data from the CAN-BIND study was collected at multiple sites with different acquisition systems. Prior to analysis of this data, it was important to ensure that the data was standardized between all sites. An overview of the standardization process was included in our *Scientific Reports* manuscript (Farzan, Atluri et al. 2017):

"Preprocessing Module 1. The aim of the first data preprocessing module is to minimize raw data heterogeneity across sites and prepare the data for integration. If unique acquisitions systems are used between CAN-BIND sites it is crucial that the raw data files be re-configured into the same file format and composition. This can be done in MATLAB (The Mathworks, Inc., Natick, MA, USA) via the open-source EEGLAB toolbox. During this process, the data is also downsampled (e.g., sampling rate and electrode montages are reduced) and re-referenced such that data from all sites are converted to have equivalent sampling rate, bandwidth, electrodes, reference, and event matrix. In CAN-BIND, this preprocessing step has thus far been conducted by dedicated research personnel and the converted data files are in EEGLAB format (*.set) which are shared on the Brain-CODE EEG platform. An important step in this preprocessing stage is to address the variety of recording channels and their associated layouts. There are

currently no established guidelines in place for integrating data between acquisition systems. When standard systems are used across acquisition systems (e.g., 10-10 EEG system), one way to address this issue is to find the closest equivalent electrodes between layouts (approach currently adapted in CAN-BIND projects). This approach has a downside when a standard montage (e.g., 10-10 EEG system) is not used across all or most sites. In such a case, a large proportion of electrodes may not have equivalents thereby resulting in a loss of spatial resolution and potentially impacting the types of EEG analyses that can be performed. For added accuracy, investigators may also choose to digitize the three-dimensional representation of the electrode layout and skull shape for every EEG recording session through commercially available digitizers (e.g. Polhemus Patriot digitizer). Including this information with each EEG recording could also provide a means to account for issues related to human error in proper EEG cap placement and improve accuracy in the interpretation of EEG outcomes."

3.1.3 Data for Study 3: Predicting Response to Pharmacotherapy

Data from study 2 was used to create the prediction models.

3.2 Data Preprocessing (All Studies)

Prior to data analysis, EEG data was manually inspected and processed to extract sources of noise from the data. In all the studies, 58 or 60 EEG channels were used for data analysis (all other non-EEG channels or channels not common to all the data acquisition systems were deleted). Electrode names are listed in Appendix I. The data was downsampled to 512Hz or 1000Hz (for faster processing of data) and bandpass filtered between 1-80Hz to remove voltage drift (<1Hz) and high frequency noise (>80Hz) that is not brain electrical activity. A notch filter (55-60Hz) was also applied to remove power line noise at 60Hz. Data was then split into 2 second windows and independent component analysis was used to extract noise components including eye movements, muscle activity, bad electrode activity, etc. Channels were deleted if they contained noise for >40% of the recorded time and later interpolated using spherical interpolation. Finally, data was re-referenced to an average reference. To make this process more efficient, an in-house preprocessing pipeline (ERPEEG)

(**http://www.tmseeg.com/multisiteprojects/**) was modified from our original pipeline for preprocessing TMS-EEG data (Atluri, Frehlich et al. 2016)). A brief outline of ERPEEG was included in our *Scientific Reports* manuscript (Farzan, Atluri et al. 2017):

"Preprocessing Module 2. The aim of the second preprocessing module is to standardize and track EEG data through the noise removal procedure. For this purpose, to streamline the process of EEG data cleaning, we developed an open-source MATLAB application, ERPEEG toolbox, depicted in **Figure 3.1**. This toolbox is developed in MATLAB (R2013a) and built using the EEGLAB platform (v.12.0.2.6b). A copy of this toolbox can be downloaded from www.tmseeg.com/multisiteprojects. The ERPEEG toolbox is created following the same framework as TMSEEG toolbox. TMSEEG toolbox enables processing of EEG collected during Transcranial Magnetic Stimulation (TMS-EEG) (e.g., standardized processing of TMS evoked potentials or TEPs), while ERPEEG is intended for processing of resting-state EEG and ERPs (Event-related potentials). This toolbox has a main interactive graphical user interface (GUI) (**Figure 3.1**A) that allows users access to the dataset working folder, a sequential list of preprocessing procedures, and the settings menu. Clicking on each step opens another interactive GUI with several data visualization suites designed for each specific processing step. The order of data processing steps in ERPEEG is standardized, and is optimized towards improving performance of each processing step. For example, random and large amplitude artifacts are processed early in the pipeline in order to increase the performance of the independent component analysis (ICA) step later in the pipeline. Following the initial step to load data and segment it into trials (Step 1), the workflow provides a GUI for *removing* data segments that are contaminated with random noise and cannot be easily de-noised. This step is particularly targeted towards the removal of channels and trials contaminated with large-amplitude or random noise sources that cannot be extracted easily through blind source estimation technique or filtering. The GUI permits interactive deletions of trials, channels and specific trials in a channel (Step 2) and keeps a log of deleted segments. Filtering can then be applied to exclude low and high frequency noise (Step 3). After these initial cleaning steps, blind source separation techniques such as ICA can then be used (Step 4, 5) to *extract* eye movements, eye blinks, electromyography artifacts, electrode discontinuity, and cardiac signals in order to recover the desired brain signals. At the completion of this step, a further interactive GUI enables a final data

review and removal of trials and channels still contaminated with random noise (Step 6). Finally, the GUI provides users the options to interpolate the deleted channels and re-reference the data (Step 7).”

“The ERPEEG toolbox contains several important features for standardized EEG data preprocessing. First, it incorporates interactive data visualization capabilities (**Figure 3.1B**), allowing the user to visualize the data at each step of the workflow and verifying the effectiveness of the data cleaning procedure. Second, intermediate datasets are saved after each processing step along with other important meta-information such as the deleted trials and channels, and the removed artifacts (**Figure 3.1C**, File Directory). This enables the user to easily revert to a previous step in the workflow, check the output of each step, and create a database of selected artifacts. Third, ERPEEG is a flexible platform. It allows for basic customization through the settings menu while providing a modular structure for advanced users to modify the order of processing or incorporate additional steps to accommodate processing of different EEG projects. Finally, parameters selected through the setting menu are saved in a separate MATLAB file. This enables future replication or assessment of the data preprocessing steps.”

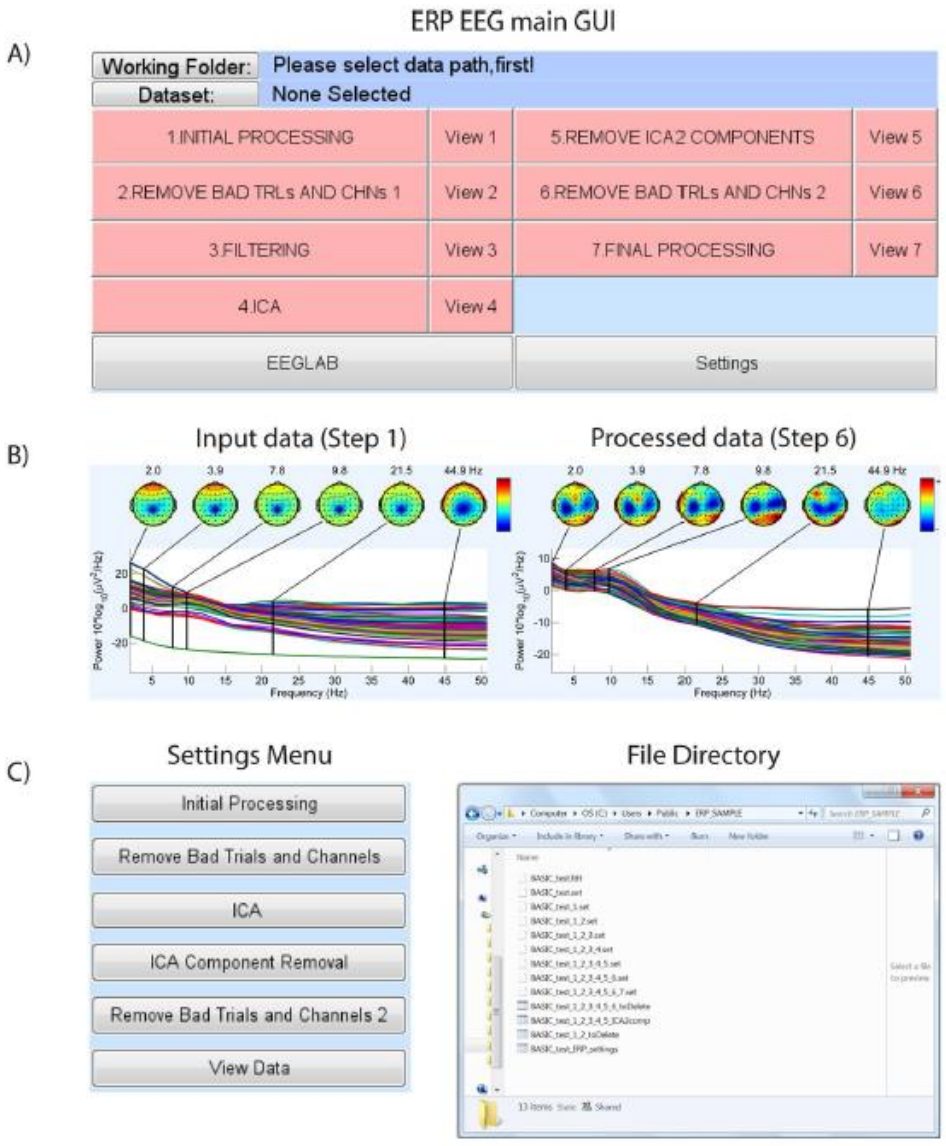


Figure 3.1 - A Streamlined Toolbox for Multi-site EEG Data Processing and Archiving (taken from (Farzan, Atluri et al. 2017)).
(A) The main graphical user interface (GUI) of the ERPEEG toolbox, with 7 preprocessing steps. Through this main interface, users select the data (by clicking on Working Folder, and Dataset), and navigate through each preprocessing step. Clicking on a processing step open a new GUI associated with that step, or runs that processing step. Steps that are completed turn green, and uncompleted turn green, and uncompleted steps remain red. (B) The view button (corresponding to each step) provides a visual summary of data cleaning processing (e.g., plots the power spectrum). This enables monitoring of data cleaning progress or detecting any major errors and data distortions. (C) The setting tab allows for selection of user-defined parameters for each step. (D) All intermediate steps (files created in completion of each step) are saved in the working folder following a standardized naming convention.

3.3 Statistical Analysis: Cluster-Based Permutation Tests

Neural signals, as measured by EEG, are characterized over several dimensions (e.g., frequency, time) across space (i.e., sensors, regions of interest). Statistical analysis of such multidimensional data must evaluate the significance of measured neural activations. Elements (voxels) of these large datasets are considered "active" if they meet a certain threshold of statistical significance (i.e., t-score or z-score) compared to other voxels in the dataset. Due to the large number of comparisons in this high-dimensional data, uncorrected voxel-wise comparisons can lead to Type I errors (false positives). Correction methods such as Bonferroni and others (Shaffer 1995) can be applied but they are highly conservative and result in Type II (false negative) errors. To address these issues, cluster-based thresholding frameworks (Poline and Mazoyer 1994, Bullmore, Suckling et al. 1999, Maris and Oostenveld 2007) were introduced and were shown to effectively control Type I errors while minimizing Type II errors (Pernet, Latinus et al. 2015). Cluster-based frameworks group active neighbouring voxels into clusters that represent neural activation patterns and rely on the continuity of the EEG signal across one or more dimensions for correction (Groppe, Urbach et al. 2011). In the spatial dimension for example, EEG data collected from sensors that are spatially close to each other are highly correlated because of volume conduction. Therefore, significant voxels that are spatially close are likely to represent significant neural activations. Cluster-based correction involves two steps: 1) setting a threshold statistic (i.e., $p<0.05$) and grouping neighbouring active voxels into clusters, and 2) calculating a *p*-value for each cluster, based on a measure of the size of activation (Cluster *P*). Cluster-mass is the measure of activation used in this thesis and is defined as the sum of original test statistics values within each cluster (Bullmore, Suckling et al. 1999, Maris and Oostenveld 2007). As such, it is considered to be a sensitive measure of neural activations since it accounts for cluster size and the intensity of the values contained within each cluster. For future studies, threshold-free methods of correction are recommended (Smith and Nichols 2009, Mensen and Khatami 2013, Frehlich, Dominguez et al. 2016)**.**

3.4 Machine Learning

For machine learning (study 3), it was important to ensure that features were extracted over the same length of time for each participant. Pre-processed EEG data was split into segments. The length of each segment was the maximum epoch length of clean data that could be derived from participants. Each feature was extracted at every segment and averaged over all segments to increase the signal-to-noise ratio of the measure. For more details on the features used to train the machine learning models, please refer to Chapter 8.

3.4.1 Feature Selection

To avoid overfitting and increase the speed of computations, filter methods were chosen to remove uninformative features from the dataset (Saeys, Inza et al. 2007). Three different filter methods were compared in this study: (1) t-test method for equal variances, (2) F-test method for equal variances, and (3) Spearman correlation method. The performance of the prediction model for each of the three methods was calculated using balanced accuracy (average of sensitivity and specificity).

3.4.2 Machine Learning Algorithm: Support Vector Machines

Support vector machines (SVMs) were used with the radial basis function (RBF) kernel using the LIBSVM toolbox (Chang and Lin 2011). Support vector machines select a hyperplane (linear) or hypersurface (non-linear) that best separates the input data space into the two (or more) pre-defined groups (i.e., responder and non-responder) (Hearst, Dumais et al. 1998). An RBF kernel uses nonlinear mapping to transform data into a higher dimension and determine an optimal separating hypersurface (Hsu, Chang et al. 2003). An optimal hypersurface that separates the two groups is one that has the largest margin (distance between the hypersurface and the closest data points). Two model parameters can be optimized for RBF kernel. The first is the cost parameter, associated with misclassification error. A high cost parameter leads to a hypersurface with a smaller margin therefore better classification accuracy. However, this may also lead to overfitting. The second parameter, gamma, is associated with the influence of a single training

example. A high gamma value suggests a close influence (Hsu, Chang et al. 2003). Default value for cost is 1 and gamma is 0. Therefore, the range for cost was specified to be around 1 and for gamma, the range was close to 0. Cost was specified as 2^C (where C = -3, -1, 1, 3), and gamma was specified as 2^G (where G = (-12, -10, -8, -6).

To optimize the parameters for the SVM model, a grid search (Chang and Lin 2011) with 10-fold cross validation was performed. Model performance is evaluated for all possible combinations of the specified values for cost and gamma and the combination that provides the best performance is returned. In this study, values of parameters were chosen to maximize balanced accuracy. It was important to use balanced accuracy since the number of responders and non-responders was not equal in our study.

3.4.3 Cross Validation

Cross validation was implemented to: (i) optimize model parameters for support vector machines, and (ii) evaluate the performance of the classifier. In the k-fold cross-validation method (**Figure 3.2**), the dataset is divided into k equal-sized segments or folds and k-repetitions of training and testing is performed. For each repetition, a 1/k parts of the data is held-out for testing while the remaining (k-1)/k parts are used for training the classifier model. In this thesis, a 10-fold randomized permutation cross validation technique was used (similar to the Monte-Carlo cross validation procedure). At each permutation, data was randomly split into 80% training set and 20% test set (ratio of responders and non-responders was ensured to equal in both sets). A feature selection method was applied to the training set and using the selected features, a 10-fold cross validation was performed to optimize model parameters for balanced accuracy. Finally, features identified in the feature selection process and the parameters identified using cross-validation were used to evaluate model performance on the separate independent test set. In this study, 100 permutations were used and overall classification performance was evaluated as an average over these 100 permutations. Please refer to Chapter 8 for more details.

1st fold

Training	Training	Training	Training	Testing

2nd fold

Training	Training	Training	Testing	Training

3rd fold

Training	Training	Testing	Training	Training

4th fold

Training	Testing	Training	Training	Training

5th fold

Testing	Training	Training	Training	Training

Figure 3.2 - K-fold cross-validation structure (for k=5)

3.4.4 Model Evaluation Metrics

The performance metrics reported in this study are accuracy, balanced accuracy, sensitivity (or recall), specificity and precision. A true positive (TP) is a responder who was correctly predicted to be a responder and a true negative (TN) is a non-responder who was correctly predicted to be a non-responder. False positive (FP) means non-responders were incorrectly identified as responders and false negative (FN) means responders were incorrectly identified as non-responders. The metrics are therefore defined as below:

$$Accuracy = \frac{TP + TN}{Total\ Number\ of\ Samples}$$

$$Balanced\ Accuracy = \frac{\frac{TP}{TP + FN} + \frac{TN}{TN + FP}}{2}$$

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Precision = \frac{TP}{TP + FP}$$

Section II: Investigating the Targets of Seizure Therapy using EEG Measures of Neural Dynamics

Chapter 4 – Brain Temporal Complexity in Explaining the Therapeutic and Cognitive Effects of Seizure Therapy

In this chapter, we apply EEG frequency analysis and multiscale entropy analysis to extract power and complexity measures of neural oscillations. We aim to identify whether these measures can provide insight into mechanism of action of ECT and MST.

Faranak Farzan, Sravya Atluri, Ye Mei, Sylvain Moreno, Andrea J. Levinson, Daniel M. Blumberger, and Zafiris J. Daskalakis. "Brain temporal complexity in explaining the therapeutic and cognitive effects of seizure therapy." *Brain* 140, no. 4 (2017): 1011-1025.

4.1 Abstract

Over 350 million people worldwide suffer from depression, a third of whom are medication resistant. Seizure therapy remains the most effective treatment in depression, even when many treatments fail. The utility of seizure therapy is limited due to its cognitive side effects and stigma. The biological targets of seizure therapy remain unknown, hindering design of new treatments with comparable efficacy. Seizures impact the brains temporal dynamicity observed through electroencephalography. This dynamicity reflects richness of information processing across distributed brain networks subserving affective and cognitive processes. We investigated the hypothesis that seizure therapy impacts mood (depressive symptoms) and cognition by modulating brain temporal dynamicity. We obtained resting-state EEG from thirty-four patients (age = 46.0 ± 14.0, 21 females) receiving two types of seizure treatments - electroconvulsive therapy or magnetic seizure therapy. We used multiscale entropy to quantify the complexity of brain's temporal dynamics before and after seizure therapy. We discovered that reduction of complexity in fine time scales underlined successful therapeutic response to both seizure treatments. Greater reduction in complexity of fine time scales in parieto-occipital and central brain regions was significantly linked with greater improvement in depressive symptoms. Greater increase in complexity of coarse time scales was associated with greater decline in cognition including the autobiographical memory. These findings were *region*- and *time-scale specific*. That is, change in complexity in occipital regions (e.g., O2 electrode or right occipital pole) at fine time-scales was only associated with change in depressive symptoms, and not change in cognition, and change in complexity in parieto-central regions (e.g., Pz electrode or intra and transparietal sulcus) at coarser time-scale was only associated with change in cognition, and not depressive symptoms. Finally, region and time-scale specific changes in complexity classified both antidepressant and cognitive response to seizure therapy with good (80%) and excellent (95%) accuracy, respectively. In this study, we discovered a novel biological target of seizure therapy: complexity of the brain resting-state dynamics. Region and time-scale dependent changes in complexity of the brain resting-state dynamics is a novel mechanistic marker of response to seizure therapy that explains both the antidepressant response and cognitive changes associated with this treatment. This marker has tremendous potential to guide design of the new generation of antidepressant treatments.

4.2 Introduction

Major depression is a leading cause of disability affecting over 350 million people globally (Murray and Lopez 1996). Over a third of these patients fail responding to medications. Dating back to 1700s, the induction of seizures has been used to treat severe psychiatric conditions such as depression. Introduced in 1930s, seizure therapy administered through electroconvulsive therapy (ECT) still remains the most effective treatment for depression (Carney, Cowen et al. 2003) even when many other antidepressant treatments have failed. However, the cognitive side effects of ECT (Lisanby, Maddox et al. 2000, McClintock, Choi et al. 2014) limit its widespread use. Magnetic seizure therapy (MST) is an emerging antidepressant treatment that involves the induction of seizure through the administration of transcranial magnetic stimulation (TMS) (Moscrip, Terrace et al. 2006, Hoy and Fitzgerald 2010, McClintock, DeWind et al. 2013). This approach to seizure induction causes less memory impairment than ECT (McClintock, DeWind et al. 2013) and early treatment studies report efficacy in depression (Kayser, Bewernick et al. 2015). Despite decades of research, the biological targets of seizure therapy for depression remain unclear. This has hindered the progress in development of new antidepressant interventions that have comparable efficacy to ECT without the cognitive side effects. Here, we propose a novel approach in examining the biological target of seizure therapy by assessing the impact of seizure on the temporal fluctuations (i.e., dynamics) of brain signals.

Seizure is a biological phenomenon that significantly impacts brain dynamicity visualized through electroencephalography (EEG). It is increasingly evident that temporal fluctuations and variability observed in biological systems such as brain signals have a fundamental role in shaping the brain's capacity for information processing (Tononi, Sporns et al. 1994, Tononi and Edelman 1998, Sporns, Tononi et al. 2000, Costa, Goldberger et al. 2005). This temporal fluctuation, occasionally referred to as biological "noise", is distinct from random noise and structurally rich (Costa, Goldberger et al. 2005) exhibiting varying degree of recurring patterns (Costa, Goldberger et al. 2005). The less recurring temporal patterns, the more complex and unpredictable the signal is. In the brain, the complexity of signals at fine (smaller time increment) and coarse (larger time increments) time-scales is proposed to arise from transient increases and decreases in correlated activity among local and distributed brain regions,

subserving integration and segregation of information at different *spatiotemporal* scales (Sporns, Tononi et al. 2000, McIntosh, Vakorin et al. 2014). While majority of existing experiments have quantified the strength of functional coupling between brain regions and its disturbance in disorders of mood and consciousness (Fox, Buckner et al. 2012, Kaiser, Andrews-Hanna et al. 2015, Sale, Mattingley et al. 2015), emerging evidence points to the abnormalities in the temporal complexity of brain signals in disorders of affect and cognition (McIntosh, Vakorin et al. 2014).

We hypothesized that seizures impact both mood and cognition by modifying the temporal complexity of brain signals in a time-scale dependent manner. We obtained resting-state EEG from two independent cohorts of patients undergoing either MST (n = 15) or ECT (n = 19). Depressive symptoms were rated through the Hamilton Rating Scale for Depression (HAMD). General cognition and autobiographical memory were obtained through the Montreal cognitive assessment scale (MoCA) and autobiographical memory interview (AMI) (**Table 4.1**).

4.3 Methods

Patients. A total of 34 subjects (age = 46.0 ± 14.0, 21 females) diagnosed with treatment-resistant depression patients participated in either of two parallel open-label seizure therapy research protocols at Centre for Addiction and Mental Health (19 ECT and 15 MST). The demographic and clinical characteristics are in (**Table 4.1**).

Seizure Therapy. ECT was administered with MECTA spECTrum 5000Q (Corporation, Lake Oswego, OR) according to standards of practice (Sackeim, Prudic et al. 2008). Sixteen patients received right unilateral ultra brief (RUL-UB) pulse width ECT, one received bitemporal (BT) brief pulse width ECT, and two started on RUL-UB and were switched to BT due to lack of efficacy (**Table 4.1**). Treatment sessions occurred twice or three times per week. Seizure threshold titration was used to determine stimulus intensity: RUL-UB was delivered at 6xthreshold with a pulse width of 0.3 to 0.37msec and BT was delivered at 1.5x threshold with a pulse width of 1.0msec. ECT treatments were continued until depressive symptoms was in remission or improvement had plateaued (refer to **Table 4.1** for number of treatments).

Finally, methohexital was administered for sedation and succinylcholine as neuromuscular blocker. In general, the target dosage was 0.75 mg/kg of methohexital and 0.5mg/kg of succinylcholine. MST was administered with Magpro MST using a Twin Coil (Magventure, Denmark). The centre of each circular coil was placed over F3 and F4 respectively, using the EEG international 10-20 system. This induces the highest electric field strength between the two coils roughly corresponding to Fz (Deng, Lisanby et al. 2013). The orientation of the magnetic fields was posterior-anterior. Subjects underwent a dose titration procedure to establish convulsive stimulation threshold. At 100Hz and 50Hz an initial train of 200 pulses was used followed by increments of 200 pulses with a maximum train of 1000 pulses. At 25Hz an initial train of 100 pulses was used with increments of 100 pulses up to a maximum of 500 pulses. Twelve subjects received 100Hz, two subjects received 50Hz and one subject received 25Hz (**Table 4.1**). All stimulations occurred at the maximum stimulator output of 100%. Threshold seizure was defined as a generalized tonic-clonic activity $\geq$ 20s of visual motor activity or $\geq$ 25s of EEG seizure activity. Subsequent treatments occurred three times per week, and were initially delivered with a train 400 pulses longer in the 100Hz and 50Hz group and 200 pulses longer in the 25Hz group. In subjects that had not achieved a 50% reduction in HAMD after three treatments, the dose was increased by 100 pulses (25Hz), or 200 pulses (50Hz, 100Hz) up to a maximum of 500 or 1000 pulses, respectively. A maximum of 24 sessions were allowed in the acute course. Methohexital (n = 9), methohexital with remifentanil (n = 5), and ketamine (n = 1) were administered for sedation and succinylcholine was used as the neuromuscular blocker.

EEG. Ten minutes of resting-state eyes closed EEG data were recorded within one week prior to the start and within 48 hours after the completion of a course of seizure therapy in both ECT and MST protocols. Subjects were instructed to sit in an armchair with eyes closed. EEG recording was through a 64-channel NeuroScan EEG system. The reference electrode was behind CZ electrode, and ground was behind FZ. The sampling rate was 10 kHz. The online filter setting was 0.05 to 1 kHz. The skin/electrode impedance was kept below 5 kOhm.

Mood. Changes in depressive symptoms were assessed by HAMD within one week prior to the start and within 48 hours after the completion of a course of seizure therapy in both ECT and MST protocols. Response to treatment was defined as 50% change in HAMD from baseline.

Cognition. 19 patients (6 ECT and 13 MST) completed the MoCA within 48 hours prior to and within a week after a course of seizure therapy in both protocols. In addition, the autobiographical memory interview short form (AMI-SF) was completed in 12 MST patients before and after a course of seizure therapy.

EEG Preprocessing. Data were imported into MATLAB (The MathWorks. Inc. Natick, MA, USA) for preprocessing. The open source signal processing functions in EEGLAB toolbox version 12.0 (Delorme and Makeig 2004) were used for data import and preprocessing. The EEG signals were epoched into segments of two seconds duration and down sampled to 1 kHz. A notch filter (band-stop: 55–65 Hz) was used to remove the 60 Hz noise. EEG signals were band passed filtered 1–50 Hz to further minimize contamination by high frequency artifact. The infinite impulse response (IIR) Butterworth filter of second order and forward and backward filtering were applied to maintain a zero phase shift. All epochs were manually reviewed and trials and channels containing eye movements, muscle or any other non-physiological artifact were discarded. The data was average re-referenced.

Power. The EEGLAB function *spectopo* was used to obtain the power spectrum for each electrode. The relative power was obtained for 1 to 50 Hz frequencies. Relative power was calculated as the ratio in the power of each frequency relative to the sum of power across all frequencies.

Multiscale Entropy. MSE was examined across all electrodes using two steps (Costa, Goldberger et al. 2005): The *coarse-graining* process and the calculation of the sample entropy (*SampEn*) for each coarse-grained time series. First, for a given time series $\{x_1, x_2, \dots x_N\}$, the multiple coarse-grained time series $\left\{y_1^{(\tau)}, y_2^{(\tau)}, \dots, y_N^{(\tau)}\right\}$ at scale factor τ (in this paper referred to as time scale)were calculated by averaging the data points within non-overlapping windows of increasing length τ. Each element of the coarse-grained time series $y_j^{(\tau)}$, was calculated according to the equation:

$$y_j^{(\tau)} = \frac{1}{\tau}\sum_{\mathrm{i}=(\mathrm{j}-1)\tau-1}^{j\tau} x_i \quad \textbf{(1)}$$

where τ represents the scale factor (i.e., time scale) and $j\ \left(1 \leq \mathrm{j} \leq \frac{N}{\tau}\right)$ represents the time index of the element. The length of each coarse-grained time series was M, where M = floor ($\frac{N}{\tau}$). At scale factor (or time scale) τ =1, the coarse-grained time series was the original time series. Second, the degree of predictability was measured for each of the multiple coarse-grained time series $\left\{y_1^{(\tau)}, y_2^{(\tau)}, \ldots, y_N^{(\tau)}\right\}$ using SampleEn. SampleEn was calculated according to the equation:

$$\mathrm{SampleEn}(\mathrm{r}, \mathrm{m}, \mathrm{M}) = -\ln(C(m+1)/C(m)) \quad \textbf{(2)}$$

where C(m) is the total number of pairs of *m* consecutive similar data points, C(m+1) is the total number of pairs of m+1 consecutive similar data points in the multiple coarse-grained time series. SampleEn quantifies the variability of time series by estimating the predictability of amplitude patterns across a time series. In our experiments, two consecutive data points were used for data matching (i.e. m = 2) and data points were considered to match if their absolute amplitude difference was less than 15% (i.e., r = 0. 15) of standard deviation of time series. MSE was calculated for a 30 second continuous epoch.

EEG Source Localization. EEG source localization was performed using an open-source application, Brainstorm **(**Tadel *et al.*, 2011**)**. First, the electrode locations of our 68-channel Neuroscan Quik Cap EEG electrode sites were co-registered to the ICBM152 MRI template in Brainstorm. The forward solution was then calculated using the OpenMEEG BEM head model (Gramfort *et al.*, 2010) and the inverse solution was derived using sLORETA (Pascual-Marqui 2002), with the solution space constrained to the cortex surface. To localize the dynamics of neural activity, we used the Destrieux Atlas, which provides 148 regions of interest (ROIs) in the MNI co-ordinate space (Destrieux *et al.*, 2010). After the EEG data was mapped to the 148 ROIs, MSE and power spectrum measures were calculated for all subjects at these sources.

***Statistics*.** In addition to two intervention groups of ECT and MST, subjects were grouped into two groups of antidepressant responders and non-responders: subjects were grouped as responders if there was a 50% or higher change in HAMD relative to baseline, and non-responders otherwise. Analysis of variance was used to 1) examine the effect of seizure therapy on MSE (1-70 time-scales) and relative power (1-50 Hz frequencies) for the main effect of Seizure Therapy Intervention (ECT, MST) and Time (Pre, Post), as well as 2) Antidepressant

Response (Responder, Non-Responder) and Time (Pre, Post) across 60 electrodes in sensor space and 148 ROIs in source space. Bootstrapping was used to correct for multiple comparisons in the analysis of variance. For the post-hoc t-test comparisons, cluster-based non-parametric permutation test (Maris and Oostenveld 2007) was used to correct for the multiple comparisons in this multi-dimensional dataset (60 channels (or 148 ROIs) x 50 frequencies, 60 channels (or 148 ROIs) x 70 scales) by assigning significance statistics to the probability of size of clusters formed by pooling adjacent pixels with original test statistics $p<0.05$. The significance of original clusters was defined against probably distribution of clusters obtained through 1000 permutations of the shuffled data labels. Identical parameters were used across the cluster-based permutations: threshold statistics of $p<0.05$, identical neighborhood, 1000 permutation using Monte Carlo approach with cluster test statistics computed as the maximum of the cluster-level summed values. Analysis of variance, and post-hoc paired t-test and independent sample t-test analyses were used to calculate the original test statistics. Spearman correlation coefficient was used to examine the association between change in complexity and symptom severity or cognitive score. Similarly, cluster-based non-parametric permutation test was applied to the behavioral scores to correct for the multiple comparisons in the correlation analyses.

In addition to correlation analysis, it was examined if change in complexity classified patients based on antidepressant and cognitive response. Subjects were grouped to have had cognitive decline if the percent change in MoCA was negative. For AMI-SF, median performance was used to divide the patients into two groups. The level of prediction was quantified by the receiver operating characteristic (ROC) curve, plotting the sensitivity and specificity of the predictor (change in complexity) across all possible threshold values. To determine the significance of the prediction, the area under the curve (AUC), standard error of the AUC and confidence intervals were quantified for each electrode and source.

Throughout the paper, except otherwise noted, reported statistics are corrected p values, and descriptive values indicate mean and standard deviation unless otherwise stated. Percent change (i.e., %Δ) in outcome variables is calculated as: (post treatment score - baseline score/baseline score) x 100, except for HAMD which is calculated as (baseline score - post treatment/baseline score) x 100.

4.4 Results

4.4.1 The Impact of Seizure on Neural Oscillations

There was a significant ($p < 0.05$) main effect of Intervention (*df* = 72, *mean F* = 14.7 (4.8 to 53.2)), Time (*df* = 72, mean *F* = 7.3 (4.8 to 19.9)) and Intervention x Time interaction effect (*df* = 72, mean *F* = 7.1 (4.3 to 18.4)) across several frequencies and electrodes. There was also a significant main effect of Antidepressant Response (mean *F* = 5.10 (3.98 to 6.83)), Time (mean *F* = 24.63 (4.07 to 80.16)), and Antidepressant Response x Time interaction effect (mean *F* = 5.98 (4.01 to 11.97)). In source space, there was a significant main effect of Time (mean F = 13.79 (4.02 to 61.86)) across multiple scales and ROIs, however the main effect of Intervention or Intervention x Time interaction effect were not significant. Finally, there was a main effect of Antidepressant Response (mean *F* = 5.14 (3.61 to 9.58)), Time (mean *F* = 27.42 (4.00 to 111.14)) and an interaction effect of Antidepressant Response x Time (mean *F* = 6.00 (3.96 to 14.53)) across multiple scales and ROIs.

Post-hoc analyses replicated the findings of prior studies that ECT induces an increase in relative power of slow cortical oscillations (Nobler and Sackeim 2008). This effect was spatially global and present regardless of the ECT therapeutic outcome. It was significant for frequencies less than 8Hz in responders (**Figure S 4.1A)** and was between 2 to 7Hz in non-responders (**Figure S 4.1B).** However, the slowing of oscillations was not significant in MST (**Figure S 4.1C-D**). Consistently, we replicated the previous finding (Nobler and Sackeim 2008) that the spatially global increase of slow oscillations (e.g., 1Hz) is associated with decline in general cognition (**Figure S 4.3A**). We found no association between change in slow oscillations and change in depressive symptoms (**Figure S 4.2A**).

We discovered that common to ECT and MST responders there was a global reduction in relative power of oscillations above 18Hz (**Figure S 4.1A, C**). ECT non-responders also had a global decrease in oscillations between 10 to 35 Hz (**Figure S 4.1B**). No changes were observed in MST non-responders (**Figure S 4.1D**). Comparing ECT with MST intervention group, we identified that ECT treatment led to higher increases in slow oscillations and higher decreases in

high frequency oscillations (**Figure S 4.4A**). This finding was also spatially global. Furthermore, comparing antidepressant responders with non-responders revealed that responders exhibited higher reduction in the power of 22 Hz oscillations and also higher frequency oscillations (**Figure S 4.4B**). This finding was spatially global at ~22Hz, but more local in higher frequencies (30-50Hz). Specifically in 30-50Hz, the reduction in power is observed in regions such as the inferior frontal sulcus, left orbital part of the frontal inferior gyrus, bilateral preocciptial notch, orbital gyri, lateral orbital sulcus, lateral occi-temporal sulcus, medial orbital sulcus, bilateral parieto-occiptial sulcus, or bilateral superior parietal lobule.

The results of correlation analysis revealed that the reduction in high frequency oscillations (gamma, e.g., 45Hz) correlated with improvement in depressive symptoms (**Figure S 4.2A**). This effect was localized to fronto-central (e.g., AF4, F1, FZ, F2, F4, FC2) and parieto-occipital (e.g., P7, P5, PO7, PO5, PO4, PO6, PO8, O1, OZ) brain regions in sensor space. In source space, there were significant negative clusters in brain areas including the orbital sulci and gyri, bilateral posterior-dorsal part of the cingulate gyrus (dPCC), ventral PCC (vPCC), precuneus, parieto-occipital sulcus, occipital pole, inferior temporal gyrus, and lateral occi-temporal sulcus in frequencies higher than 30Hz **(Figure S 4.5A).**

Finally, a spatially widespread decrease in low frequency oscillation (< 9 Hz) correlated with a change in cognition (**Figure S 4.3A**). Source analysis also revealed that this effect was spatially global. Finally, reduction in high frequency oscillations (e.g., > 40Hz) in parieto-central regions (e.g., C1, C3, CZ, CP3, P1, PZ, P2, P4, POZ) correlated to change in cognition. In source space, this effect was identified primarily in brain regions including the central sulcus, angular gyrus, and subparietal sulcus **(Figure S 4.5B).**

4.4.2 The Impact of Seizure on Temporal Complexity

We then employed multiscale entropy (MSE) (Costa, Goldberger et al. 2005) to quantify the change in complexity of dynamics across multiple time-scales. In sensor space, there was a significant (p <0.05) main effect of Intervention (df = 72, *mean* F = 10.5 (4.7 to 27.2)), Time (df = 72, mean F = 6.7 (4.6 to 14.4)) and Intervention x Time interaction effect (df = 72, *mean* F =

7.2 (4.5 to 18.6)) across multiple time-scales and electrodes. There was also significant main effect of Antidepressant Response (mean F = 5.1 (4.5 to 6.8)), Time (mean F = 13.7 (4.1 to 46.5)), and Antidepressant Response x Time interaction effect (mean F = 6.3 (4.3 to 11.2)). Similarly, in source space, we found a significant main effect of Intervention (mean F = 14.84 (4.33 to 46.52)), Time (mean F = 6.74 (4.20 to 17.4), and Intervention x Time interaction effect (mean F = 5.76 (4.16 to 11.20)) across multiple scales and ROIs. Finally, there was a main effect of Antidepressant Response (mean F = 5.59 (3.87 to 10.43)), Time (mean F = 16.07 (4.02 to 61.19)), and interaction effect of Antidepressant Response x Time (mean F = 5.92 (4.04 to 17.09)) across multiple scales and ROIs.

Post-hoc analysis revealed that, change in temporal complexity was only significantly modified in responders of both seizure therapies. Common to ECT and MST responders, there was a decrease in time-scales finer than 20 factors (**Figure 4.1A, C**). ECT responders showed a significant (*cluster p* = 0.003) global decrease in time-scales less than 30 and a significant (*cluster p* = 0.002) global increase in coarser time-scales (spatially global changes are seen in time scales > 50). Source-space analysis (**Figure 4.5B**) confirmed the spatially global extent of this finding. By contrast, in MST responders, a wide spread reduction in time-scales less than 20 was observed (*cluster p* = 0.033). In MST responders, the reduction of MSE in fine time-scales (e.g., scale factor 4) was found in the parieto-occipital (P1, P3, P2, P4, POZ, PO3, PO5, PO7, PO4, PO6, PO8, O1, OZ) and fronto-central regions (F4, FC1, FC2, FCZ, CZ, C1, CZ) in sensor space. Similarly, in source space, this change was observed across several tempro-parieto-occipital regions (e.g., cuneous, precuneus, posterior-dorsal part of the cingulate gyrus (dPCC), parieto-occipital sulcus, occipital pole, etc) and fronto-central regions (e.g., opercular part of the inferior frontal gyrus, central sulcus, pre and post central gyrus, etc) (**Figure 4.5A**). No significant changes were observed in either ECT or MST non-responders (**Figure 4.1B, D**).

Finally, comparing ECT with MST intervention group, we identified MSE in fine time scales (e.g., < 10) was significantly lower post treatment in ECT compared to MST group, and increases in coarse time scales (e.g, > 28) were significantly higher in ECT compared to MST intervention group (**Figure 4.2A**). This effect was spatially global. Comparing antidepressant

responders with non-responders (**Figure 4.2B**) identified significant differences between groups; however this finding did not survive the cluster-based correction for multiple comparisons. The comparison revealed that responders may have a larger reduction in MSE post treatment at finer time scales in brain regions such as the precuneus, bilateral cuneus, bilateral parieto-occiptial sulcus, bilateral occipital pole, bilateral lateral occi-temporal gyrus, calcarine sulcus, and bilateral posterior transverse collateral sulcus. Responders also appeared to have increased MSE post treatment in coarser time scales (e.g., > 40) mainly in the left inferior, middle and superior frontal sulcus, middle and superior frontal gyrus, and orbital part of inferior frontal gyrus. This latter observation is likely related to the higher number of ECT responders (compared to MST responders) who exhibited significant increases in complexity of coarse time scales (e.g., **Figure 4.1A**).

4.4.3 The Impact of Change in Temporal Complexity on Mood and Cognition

We then determined whether change in complexity was linked with the impact of seizure on mood and cognition. We found a negative association between percent changes in MSE (%ΔMSE) and percent change in HAMD (%ΔHAMD). This effect was selective to fine time-scales in parieto-occipital and fronto-central regions (**Figure 4.3A**). Specifically, a negative association ($p<0.01$) was identified in tempro-parieto-occipital (TP7, P7, P5, P8, PO7, PO5, PO6, PO8, O1, O2, Oz) and fronto-central regions (AF4, F1, FZ, F2, F4, FC1, FC2, FC4, FCZ, C1, C4, CZ) in time-scales less than 30. Source space analysis localized this effect to several regions including the dPCC, cuneus, precuneus, parieto-occipital sulcus, occipital pole, temporal sulci, and lateral occi-temporal sulcus as depicted in **Figure 4.5C**. This association illustrated that a spatially specific decrease in complexity of fine time-scales was linked with a greater improvement in depressive symptoms.

Moreover, we found a negative association between %ΔMSE and percent change in general cognition (%ΔMoCA). This effect was spatially global in coarse time-scales (e.g., > 66) and included a wide range of time-scales in parieto-central regions (e.g., PZ, POZ, P1, P2) (**Figure 4.3B**). Source-space analysis (**Figure 4.5D**) confirmed that this effect was spatially global across

coarse time scales and included brain regions such as intraparietal sulcus and transverse parietal sulci. This negative association was replicated for change in autobiographical memory (%ΔAMI) (**Figure S 4.6A**) and was prominent in bilateral frontoparietal brain regions. Likewise, source space analysis revealed a spatially global effect in coarse time scales with many brain areas involved including the bilateral superior parietal sulcus, and superior temporal sulcus (**Figure S 4.7A**). Collectively, this negative association illustrated that an increase in MSE, in particular globally in coarser time-scales, was linked with a greater decline in cognition.

These findings were *region*- and *time-scale specific*. That is, change in complexity in occipital regions and fine time-scales was only associated with change in HAMD (e.g., O2 electrode, time-scale 4, r = -0.52, p = 0.0017), and not change in MoCA (**Figure 4.3C**), and change in complexity in parieto-central regions at coarser time-scale (e.g., PZ electrode, time-scale 70, r = -0.63, p = 0.0038) was only associated with change in cognition, and not HAMD (**Figure 4.3D**).

4.4.4 Classifying Antidepressant and Cognitive Response to Seizure Therapy

Finally, we examined whether change in temporal complexity could classify patients based on cognitive and antidepressant response to seizure therapy. We found that %ΔMSE classified the antidepressant response to seizure therapy with good performance and cognitive response with excellent performance as illustrated with the area under the curve (AUC) property of the receiver operating characteristic (ROC) curve (**Figure 4.4**). Specifically, change in complexity of low time-scales (e.g., 4-6,8) in right parieto-occipital brain regions (OZ, O2, PO8) offered good (AUC ≥ 0.8) prediction performance of antidepressant response ((e.g., AUC (OZ electrode, time-scale 5) = 0.83, p<0.0001; **Figure 4.4A, B**)) and a fair (0.7 < AUC < 0.8) prediction performance was observed across low time-scales (e.g., 1-22) in bilateral fronto-central (e.g., FC1, FC2, FCZ, F1) and bilateral parieto-occipital (e.g., O1, PO3, PO5, PO7, PO4, PO6, P7, P8) brain regions. In source space, similar prediction accuracy was identified for the right occipital pole at similar time scale (AUC (right occipital pole, time-scale 5) = 0.79, p<0.0001; **Figure 4.6A, B).** Moreover, change in complexity of time-scales 14 and higher in parieto-central (e.g., PZ) and then globally in coarser time-scales provided excellent (e.g., AUC ≥ 0.9) prediction

performance for change in cognition (e.g., AUC (P2 electrode, time-scale 23) = 0.98, $p<0.0001$; **Figure 4.4C, D**). In source space, similar prediction accuracy was identified for the intraparietal sulcus and transverse parietal sulci at similar time scale (AUC (intraparietal sulcus transverse parietal sulci, time-scale 22) = 0.97, $p<0.00001$; **Figure 4.6C, D).**

These findings were *region*- and *time-scale specific*. That is, change in complexity in occipital regions and fine time-scales (e.g., OZ, time-scale 5; **Figure 4.4B**) only classified antidepressant response, and did poorly in classifying cognitive response (AUC (OZ, time-scale 5) = 0.55, p = 0.35), and change in complexity in parieto-central regions at coarser time-scale (e.g., P2, scale 23; **Figure 4.4D**) only classified cognitive response and did poorly in classifying antidepressant response (AUC(P2, time-scale 23) = 0.47, p = 0.63).

Moreover, the seizure therapy induced changes in autobiographical memory (%ΔAMI) could also be accurately (e.g., AUC range: 0.9 to 1.00; **Figure S 4.6B**) classified by change in complexity in coarse time-scales (e.g.,>47) in frontoparietal regions (**Figure S 4.6B**). Likewise, source space analysis revealed a spatially global effect in coarse time scales with many in frontal parietal brain regions (**Figure S 4.7B**). Finally, our results showed that change in complexity had better accuracy than neural oscillations in predicting antidepressant or cognitive response to seizure therapy (**Figure S 4.2B, Figure S 4.3B**).

4.5 Discussion

This study presented a novel biological target – i.e., complexity of the brain resting-state dynamics - whose modulation in specific brain regions explained the antidepressant efficacy and cognitive consequences of seizure therapy in depression. In contrast to neural oscillations, significant changes in the complexity of brain dynamics were only present in responders of seizure therapy. Specifically, complexity of fine time-scales was significantly reduced following successful ECT and MST. Across groups, the greater reduction in complexity of fine time-scales in fronto-central and parieto-occipital regions (e.g., right occipital pole) was associated with greater improvement of depressive symptoms. In ECT, the complexity of coarse time-scales was

also significantly increased. Across groups, the greater global increase in complexity of coarse time-scales was linked with the greater decline in general cognition. Finally, region- and time-scale dependent changes in complexity classified patients based on antidepressant efficacy (e.g., in right occipital pole, scale 5) and cognitive consequences (e.g., intraparietal sulcus and transverse parietal sulci, scale >22), of seizure therapy with good (≥80%) and excellent (≥90%) accuracy, respectively.

ECT remains the most effective treatment in depression. Several hypotheses have attempted to explain the mechanism of action of ECT (reviewed in (Farzan, Boutros et al. 2014)). We recently proposed a unifying *connectivity-resetting* hypothesis, stating that ECT resets aberrant neural connectivity by activating the brain's major oscillatory pacemaker, thalamus and subsequently multiple thalamic loops (Farzan, Boutros et al. 2014). The significant impact of seizure therapy on neural oscillations has been quantified since early studies of ECT (reviewed in (Farzan, Boutros et al. 2014)). The most replicated finding is the general slowing of oscillations in ECT (Small, Small et al. 1978) linked with improvement in mood (Fink and Kahn 1957, Sackeim, Luber et al. 1996). While we replicated previous findings demonstrating that ECT induces increase in power of slow oscillations, this effect was present in both ECT responders and non-responders. Moreover, we found no correlation between change in slow oscillations and improvement in symptoms. Most previous studies focused on limited and predefined frequency bands, using a few electrodes placed near the site of stimulation, or utilized statistical approaches that limited multidimensional analysis. The present study used non-parametric statistical approaches and two distinct modalities of seizure induction to comprehensively assess changes common to seizure therapy without *a priori* hypothesis or limiting analysis to regions and frequencies of interest. Our comprehensive analysis revealed that it is the reduction in relative power of frequencies 18Hz and above, particularly higher than 35Hz, rather than increase in slow oscillations, that is linked with response to seizure therapy across ECT and MST. The observation that successful MST modulated high frequencies without significantly impacting slow oscillations further confirms that successful seizure therapy may be achieved without impacting the slow oscillations that are linked with the adverse effects of ECT as reported previously (Sackeim, Luber et al. 2000, Nobler and Sackeim 2008) and replicated in our study.

The peri-ictal characteristic of seizure reflects a rapid modification of the brain dynamicity. It seems intuitive that modulation of the brain dynamics would be a mechanism by which seizure exerts its therapeutic action. Yet this has been only minimally investigated. In an ECT case study in three patients with depression, reduction in MSE in fine time-scales was reported (Okazaki, Takahashi et al. 2013). Our results are also in line with a previous study that showed an abnormal enhancement in complexity of frontal brain regions in depression which was normalized by antidepressant medication (Mendez, Zuluaga et al. 2012). Complexity, as indexed by Lempel-Ziv Complexity, was increased as a function of age in healthy subjects, a relationship not found in depression. Furthermore, six months of treatment with the antidepressant mirtazapine normalized the excess complexity in depression specifically in younger adults (Mendez, Zuluaga et al. 2012). Such findings may suggest that both medications and seizure therapy act on reducing complexity in depression, while the higher efficacy of seizure therapy may be linked to direct stimulation of oscillatory pacemakers. We found that antidepressant efficacy of seizure therapy was linked with local changes in complexity. Our findings and these previous studies encourage design of non-seizure interventions that target the same biological targets as seizure therapy toward eliminating the risk and complications of seizure induction.

Complexity of time series in biological systems is suggested to reflect plasticity to an ever changing environment and adaptability to stressors (McIntosh, Vakorin et al. 2014). When examined across brain regions and time-scales, the complexity of brain dynamics can arise from transient increases and decreases in correlated activity across brain regions reflecting rate of information generation (McIntosh, Vakorin et al. 2014). Induction of seizure could reset integration and synchronization of information across brain regions, through activation of thalamus and multiple thalamic loops and interconnected brain regions, significantly impacting rate of information generation across distributed brain networks. The association between reduction in complexity and improvement in symptoms is in line with imaging findings that have shown that depression is associated with states of hyperconnectivity between frontoparietal and default mode network (Kaiser and Pizzagalli 2015). The clinically relevant reduction of complexity in fronto-central and parieto-occipital regions adds to the resting-state functional connectivity findings in fMRI literature.

A recent study using fMRI data from Human Connectome Project showed differential association between functional connectivity of resting-state networks and complexity of fMRI time signals in fine versus coarse time-scales (McDonough and Nashiro 2014). The time-scales in this fMRI study are coarse in comparison to the present high resolution EEG study, hindering direct interpretation. Yet, it provides evidence that there may be a link between seizure-induced changes in complexity and aberrant neural connectivity in depression. We suggest that a change in dynamics of functional connectivity between distributed brain regions may be a mechanism by which seizure therapy exerts its impact on behavior. Design of non-invasive interventions that can selectively modify the complexity of the brain dynamics will enable careful examination of the consequence of region- and network-specific modification of MSE on human behavior.

Moreover, the finding that seizure induced changes in the occipital lobe (e.g., occipital pole) were linked to mood improvement and predicted therapeutic response is also in line with several lines of emerging evidence that have linked depression with impairment in this brain region (reviewed in (Koch and Schultz 2014)). For example, as reviewed by Koch et al., a recent meta-analysis reported the right occipital lobe, with the inferior fronto-occipital fibre tract, to be among the most consistently reported site of decreased white matter integrity in this population. Furthermore, in addition to the changes in white matter structure, changes in resting-state connectivity and gray matter volume have been previously shown in this brain region in depression (e.g., (Grieve, Korgaonkar et al. 2013, Meng, Brandl et al. 2014)). Moreover, a recent study has reported that occipital bending is more common in depression (Maller, Thomson et al. 2015). Finally, a prior study in post-stroke depression have identified that post-stroke depression was closely linked with the right hemisphere lesion volume and its proximity to the occipital pole (Shimoda and Robinson 1999). Therefore, our finding that seizure therapy may exert its antidepressant efficacy by impacting the dynamics of the occipital region, particularly source localized to the occipital pole, not only complements these prior findings, but also provides a direction for development of novel antidepressant treatments.

This study also adds new insight about the link between region- and time-scale dependent changes of complexity and human behavior. We illustrated a region-specific reduction of MSE in fine time-scales that was linked with improvement in mood, and a more spatially-distributed

(e.g., bilateral frontoparietal) increase of MSE in coarse time-scales that was linked with cognitive decline. Previous studies have reported both global and region-specific modulation in complexity, such as during development (Misic, Mills et al. 2010) and aging (McIntosh, Vakorin et al. 2014), respectively. Indeed, the observed link between increase in MSE in coarse time-scales and cognitive decline is consistent with findings in Alzheimer's disease (Mizuno, Takahashi et al. 2010). Our findings also extend previous studies that revealed significant modifications in this marker during adolescence (Vakorin, McIntosh et al. 2013), when the prevalence rate of depression peaks, and in disorders of cognition and affect with overlapping symptoms with depression including autism spectrum disorder (Bosl, Tierney et al. 2011) or schizophrenia (Takahashi, Cho et al. 2010) in which seizure therapy is also indicated.

MST treatment frequency may be an important dimension involved in production of a seizure. The majority of prior MST trials have applied MST at 100Hz frequency to achieve seizure induction. However, it was proposed that the optimal frequency for seizure induction may be in the vicinity of 22 Hz (Peterchev, Rosa et al. 2010). In our sample, the most common MST frequency used was 100Hz (in 12/15 subjects), while a few patients who also took part in the resting-state EEG assessments received lower frequency of stimulation to induce seizure. Nevertheless, the present EEG study was not designed to evaluate the impact of different frequency of stimulation on therapeutic outcome. We propose that the markers presented in this study have the potential to be used to protect against any potential cognitive adverse effects through neurophysiological monitoring that may predate any cognitive deterioration.

4.6 Conclusions

Our findings support a focal antidepressant target for seizure therapy. First, the association between change in MSE and depressive symptoms was identified in fronto-central and parieto-occipital electrodes and source localized to several parieto-occipital brain regions including the occipital pole. Second, the reduction in MSE was observed more localized to these brain regions in fine time-scales in responders in MST which is a more focal method of seizure induction. Consistently, the association between change in neural oscillations and depressive symptoms was also localized to fronto-central and parieto-occipital brain regions and high frequency

oscillations that correspond to fine time-scales. Fourth, the classification performance of the change in complexity was region- and time-scale specific. Brain regions at which change in complexity classified antidepressant response with good accuracy failed to classify cognitive response, and brain regions at which change in complexity classified cognitive response failed to classify antidepressant response. Recent evidence indicates the possibility of modulating the temporal complexity of brain signals by network guided rTMS (Farzan, Pascual-Leone et al. 2016). Therefore, treatment of depression may benefit from design of more localized seizure induction strategies or non-seizure treatments (e.g., rTMS) that could focally modulate complexity.

4.7 Tables

Table 4.1 - Demographics and Clinical Characteristics

	ECT Responders [n=12]	**ECT Non-responders [n=7]**	**MST Responders [n=5]**	**MST Non-responders [n=10]**
Age (years)	43.3 ± 16.3	57.7 ± 9.2	45.8 ± 8.01	40.1 ± 15.5
Sex, M/F	4/8	2/5	2/3	5/5
Illness Duration, Years [n]	19.9 ± 11.5 [11]	19.3 ± 13.6	25.4 ± 14.6	18.3 ± 14.2
Number of Treatments	12.5 ± 4.0	14.4 ± 3.2	18.6 ± 7.5	21.5 ± 5.8
Site of Treatment,[n]	right unilateral ultra brief pulse [12]	right unilateral ultra brief pulse [4] bitemporal standard pulse (2) unilateral followed by bitemporal (1)	Midline Frontal [5]	Midline Frontal [10]
Stimulation Frequency, [n]	NA	NA	100Hz [4] 50Hz [0] 25Hz [1]	100Hz [8] 50Hz [2] 25Hz [0]
% Change in HAMD	65.20 ± 7.8	13.5 ± 23.4	70.5 ± 16.0	13.6 ± 21.4
% Change in MoCA, [n]	-13.0 ±11.9 [4]	-7.69 ± 10.9 [2]	12.3 ± 24.2 [5]	0.9 ± 11.8 [8]

4.8 Figures

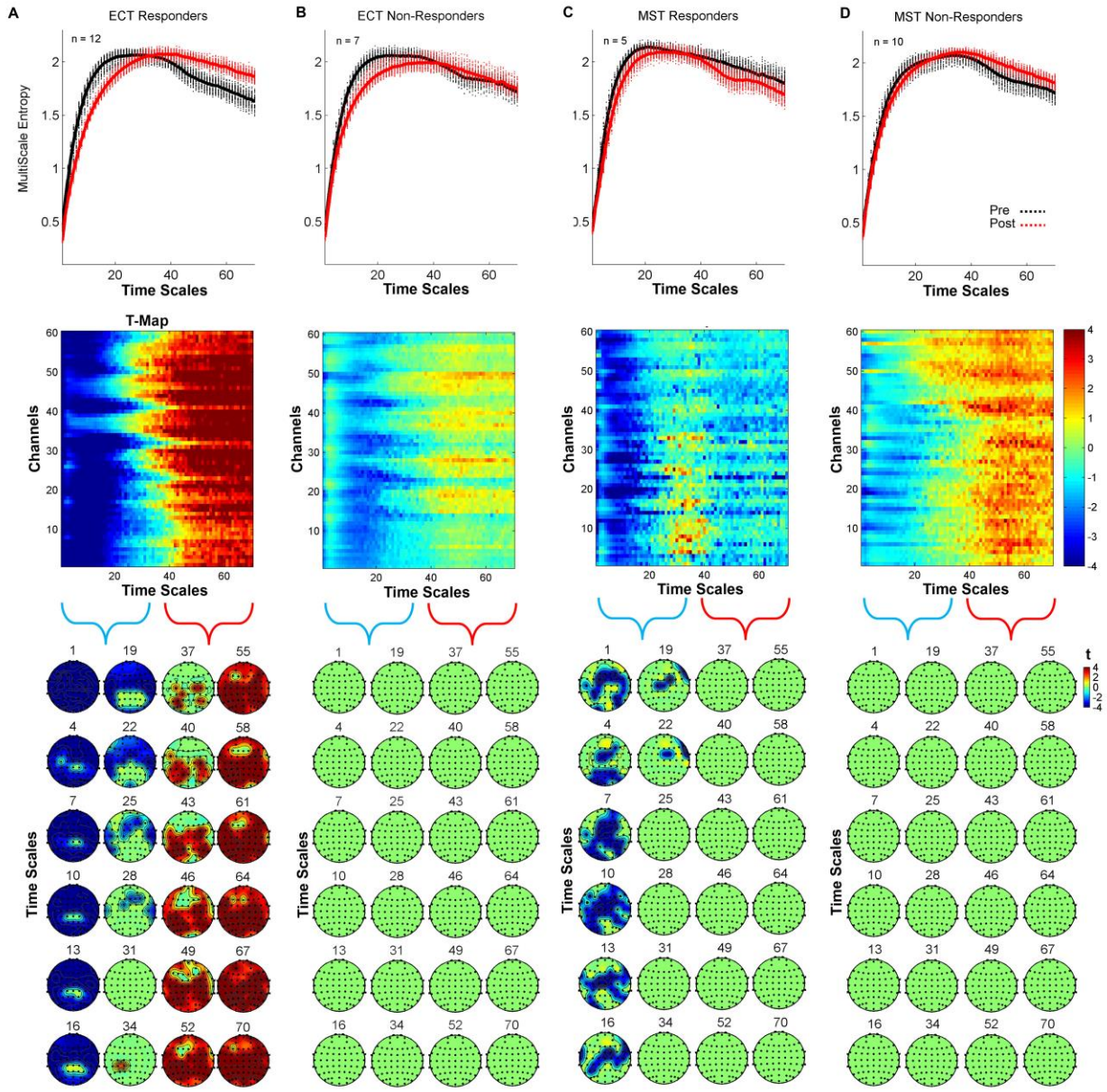


Figure 4.1 - Effect of Seizure Therapy on Complexity of Temporal Dynamics
Top. Waveforms depict multiscale entropy (MSE) pre (black line) and post (red line) electroconvulsive therapy (ECT) and magnetic seizure therapy (MST) in responders (**A, C**) and non-responders (**B, D**). The lines represent the average MSE (y-axes) across electrodes (dots) for time-scales 1 to 70 (x-axes). **Middle.** Images show the original post-hoc test statistics comparing MSE post to pre-treatment across all electrodes (1 to 60) and all time-scales (1 to 70) (blue: decreases; red: increases following treatment) for responders and non-responders to ECT (**A,B**) and MST(**C,D**). **Bottom.** Each topography reflects the significant t-maps following correction for multiple comparison, using cluster-based non-parametric permutation test, depicting only the

significant clusters $p<0.05$ and setting to 0 non-significant pixels. Topographies highlight the spatial characteristics of the reduction of MSE in fine time-scales common to both ECT and MST responders (**A, C**) and the increase in MSE in coarse time-scales following ECT alone (**A**). In ECT responders, there was a significant (*cluster p* = 0.003) global decrease in time-scales less than 30 and a significant (*cluster p* = 0.002) global increase in coarser time-scales. By contrast, in MST responders, only a wide spread reduction in time-scales less than 20 was observed (*cluster p* = 0.033). In MST responders, the reduction of MSE in fine time-scales (e.g., scale factor 4) was localized to parieto-occipital (P1, P3, P2, P4, POZ, PO3, PO5, PO7, PO4, PO6, PO8, O1, OZ) and fronto-central regions (F4, FC1, FC2, FCZ, CZ, C1, CZ). No significant changes were observed in either ECT or MST non-responders.

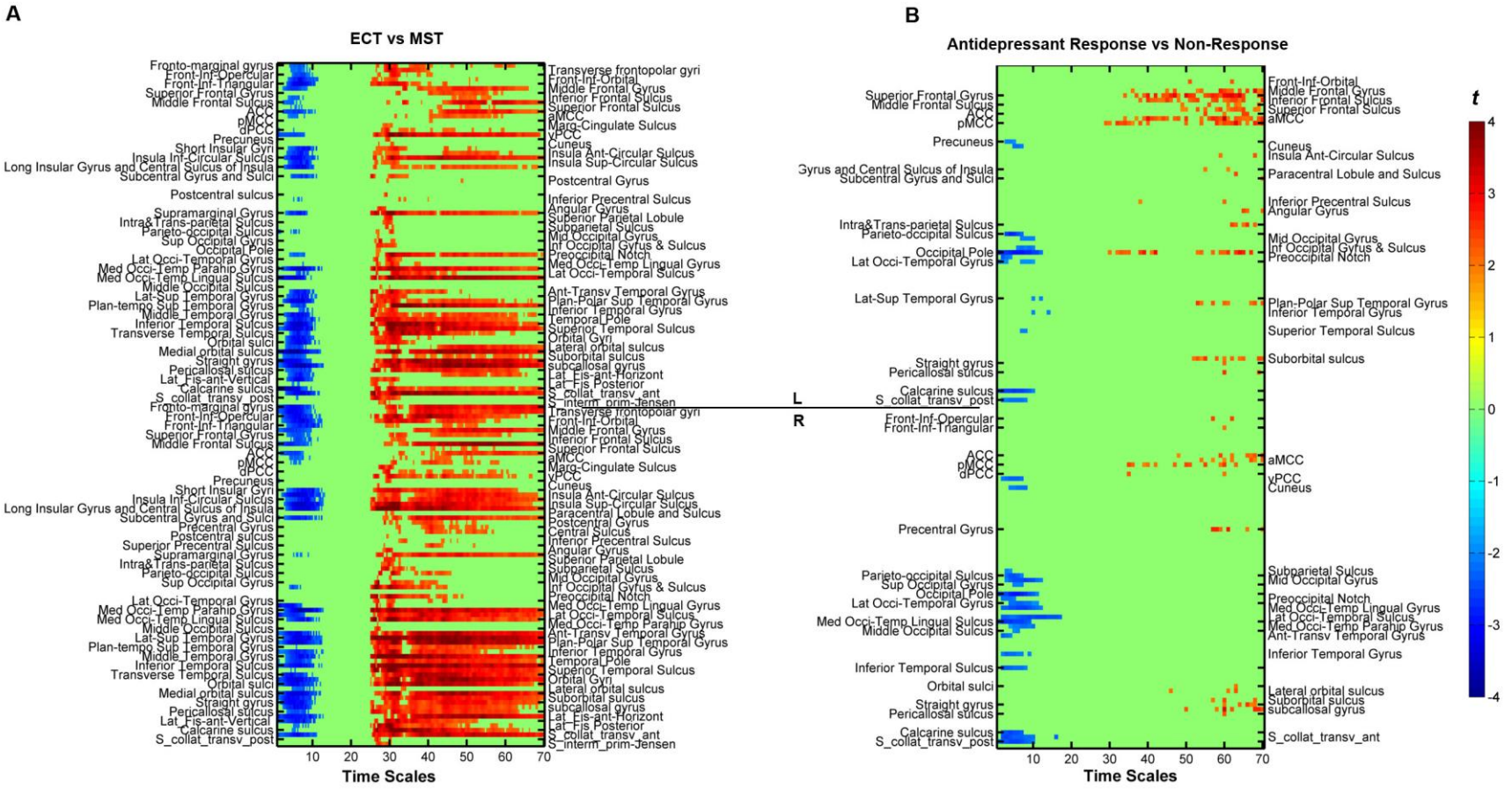


Figure 4.2 - Effect of Seizure Therapy on Complexity in the Source Space.
In all images, X-axis represents the time scale (1 to 70) and y-axis represents Regions of Interest (ROIs) of the Destrieux Atlas (1 to 148). The ROIs are grouped into brain regions in the left (L: the upper half the images) and right (R: the lower half of the images) hemisphere separated by the horizontal black line. **A.** Image show the post-hoc independent sample t-test statistics following cluster-based permutation test correction for multiple comparison, depicting only the significant clusters $p<0.05$, labeling only the significant corresponding ROIs, and setting to 0 non-significant pixels. Image shows the t-test statistics comparing the change in MSE (Post-Pre/Pre) between participants who received ECT and MST interventions (red: higher increases in ECT; blue: higher decreases in ECT). This image depict that MSE in fine time scales (e.g., < 10) was significantly lower post treatment in ECT compared to MST

group, and increases in coarse time scales (e.g, > 28) were significantly higher in ECT compared to MST intervention group. **B.** Image shows the independent sample t-test statistics comparing the change in MSE (Post-Pre/Pre) between participants who were considered responders to seizure therapy (>= 50% reduction in HAMD from baseline) and non-responders (red: higher increases in responders; blue: higher decreases in responders). The regions of significance did not survive the cluster-based correction for multiple comparisons at cluster p <0.05, thereby, this image depicts the outcome of bootstraping statistics only. Responders may have more reduction in MSE post treatment in fine time scales in brain regions such as precuneus, bilateral cuneus, bilateral parieto-occiptial sulcus, bilateral occipital pole, bilateral lateral occi-temporal gyrus, calcarine sulcus, and bilateral posterior transverse collateral sulcus. Responders may also have more increases post treatment in coarser time scales (e.g., > 40) mainly in the left inferior, middle and superior frontal sulcus, middle and superior frontal gyrus, and orbital part of inferior frontal gyrus.

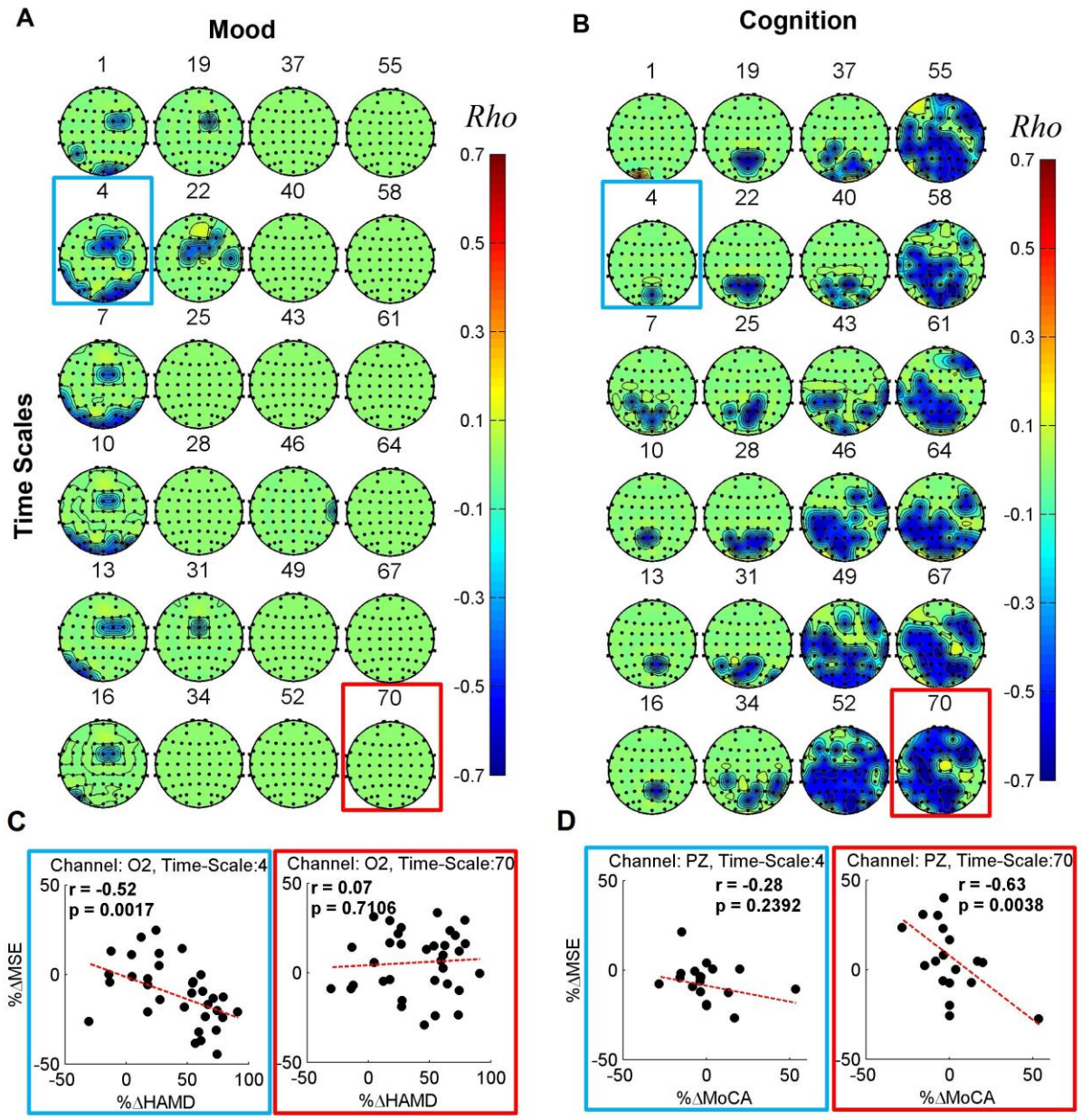


Figure 4.3 - Association between Modulation of Temporal Complexity and Mood and Cognition.

A. Topographies illustrate all the significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in HAMD and MSE in 34 patients receiving seizure therapy. Cluster-based correction for multiple comparison resulted in significant negative clusters ($p<0.01$) in parieto-occipital (TP7, P7, P5, P8, PO7, PO5, PO6, PO8, O1, O2, Oz) and fronto-central regions (AF4, F1, FZ, F2, F4, FC1, FC2, FC4, FCZ, C1, C4, CZ) in time-scale less than 30 factors. **B.** Topographies illustrate all the significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MoCA and MSE across time-scales in 19 patients receiving seizure therapy. Cluster-based correction for multiple comparison revealed a significant negative cluster ($p<0.01$) in parieto-central region (e.g., PZ, POZ, P1, P2) across

time-scales and globally in coarser (higher) time-scales. **C, D.** Scatter plots highlight the time-scale and region-specific association between percent change in MSE (y-axes) and percent change in HAMD (x-axis in **C**), and percent change in MoCA (x-axis in **D**). **C.** Scatter plots show that change in MSE was significantly associated with change in HAMD in the occipital region in fine time-scale (O2, time-scale 4, r = - 0.52, p = 0.0017) but not coarse time-scale (O2, time-scale 70, r =0.07, p = 0.71). **D.** Scatter plots show that change in MSE was significantly associated with change in MoCA in the parieto-central region in coarse time-scale (PZ, time-scale 70, r = -0.63, p = 0.0038) but not fine time-scales (PZ, time-scale 4, r = -0.28, p = 0.24).

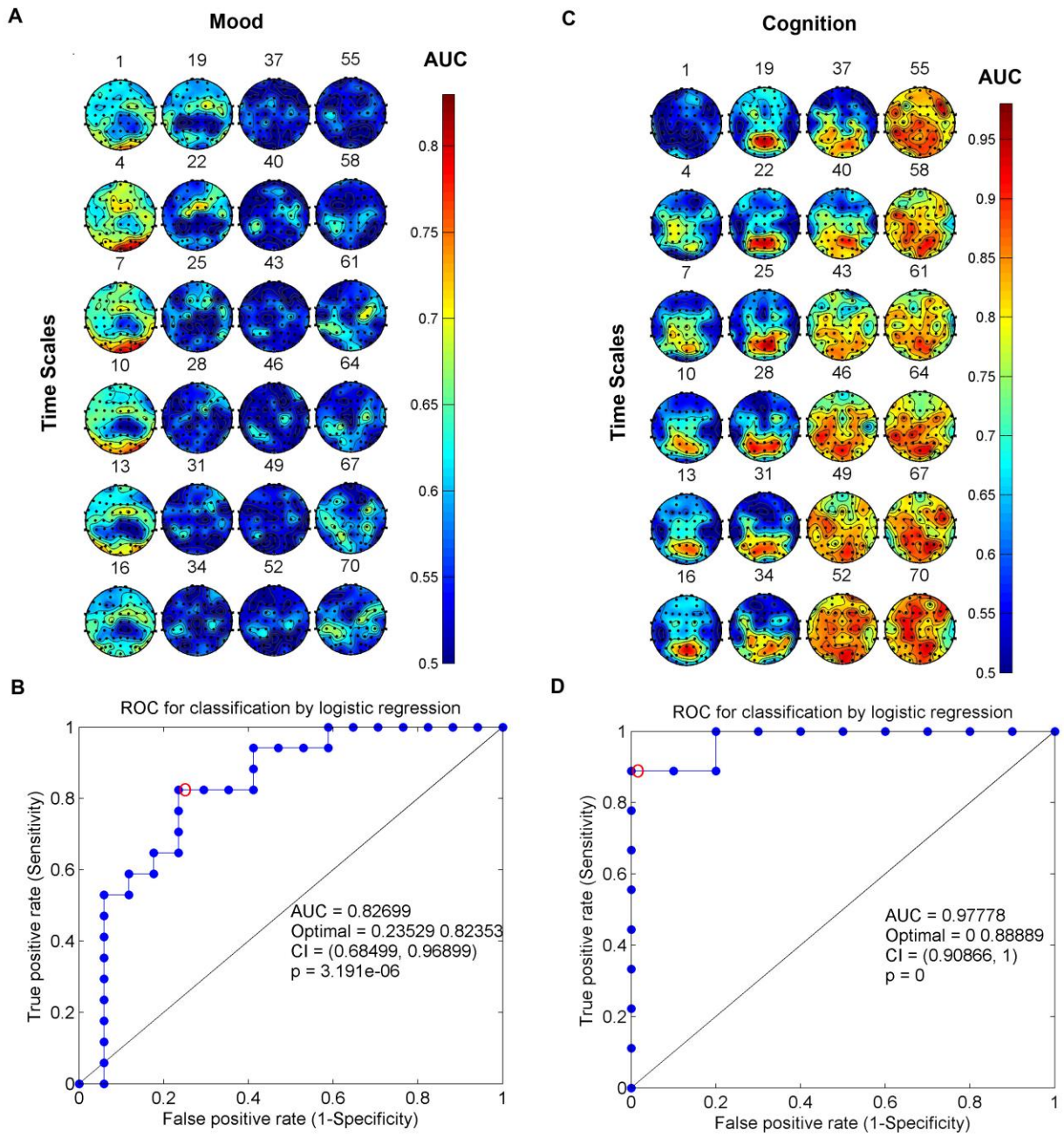


Figure 4.4 - Region-Specific Change in Temporal Complexity Predicts Change in Mood and Cognition

A, C. Topographies depict area under the curve (AUC) of the receiver operating characteristic (ROC) curve of change in multiscale entropy (MSE) in predicting antidepressant (**A**), and cognitive change (**C**) in response to seizure therapy at every electrode and time-scale. The hot colors illustrate higher AUC and better prediction. Change in complexity of low time-scales (e.g., 4-6,8) in right parieto-occipital brain regions (OZ, O2, PO8) offered good (AUC $\geq$ 0.8) prediction performance of antidepressant response and a fair (0.7 < AUC < 0.8) prediction

performance was observed across low time-scales (e.g., 1-22) in bilateral fronto-central (e.g., FC1, FC2, FCZ, F1) and bilateral parieto-occipital (e.g., O1, PO3, PO5, PO7, PO4, PO6, P7, P8) brain regions (**A**). Change in complexity of time-scales 14 and higher in parieto-central (e.g., PZ) and then globally in coarser time-scales provided excellent (e.g., AUC $\geq$ 0.9) prediction performance for change in cognition. **B, D.** Figures depict the ROC curve across all possible threshold values of the predictor for an electrode and time-scale with best prediction performance for antidepressant response (OZ, scale 5) (**B**) and change in cognition (e.g., AUC (P2 electrode, time-scale 23) = 0.98, p <0.0001). (**D**). X-axes represent false positive rates (1-specificity), y-axes the true positive values (sensitivity). The red circle shows the optimum operating point of the ROC curve. **B.** At optimum point, this electrode and scale has 82% sensitivity and 77% specificity (good classification). **D.** At optimum point, this electrode and scale has 89% sensitivity and 100% specificity (excellent classification).

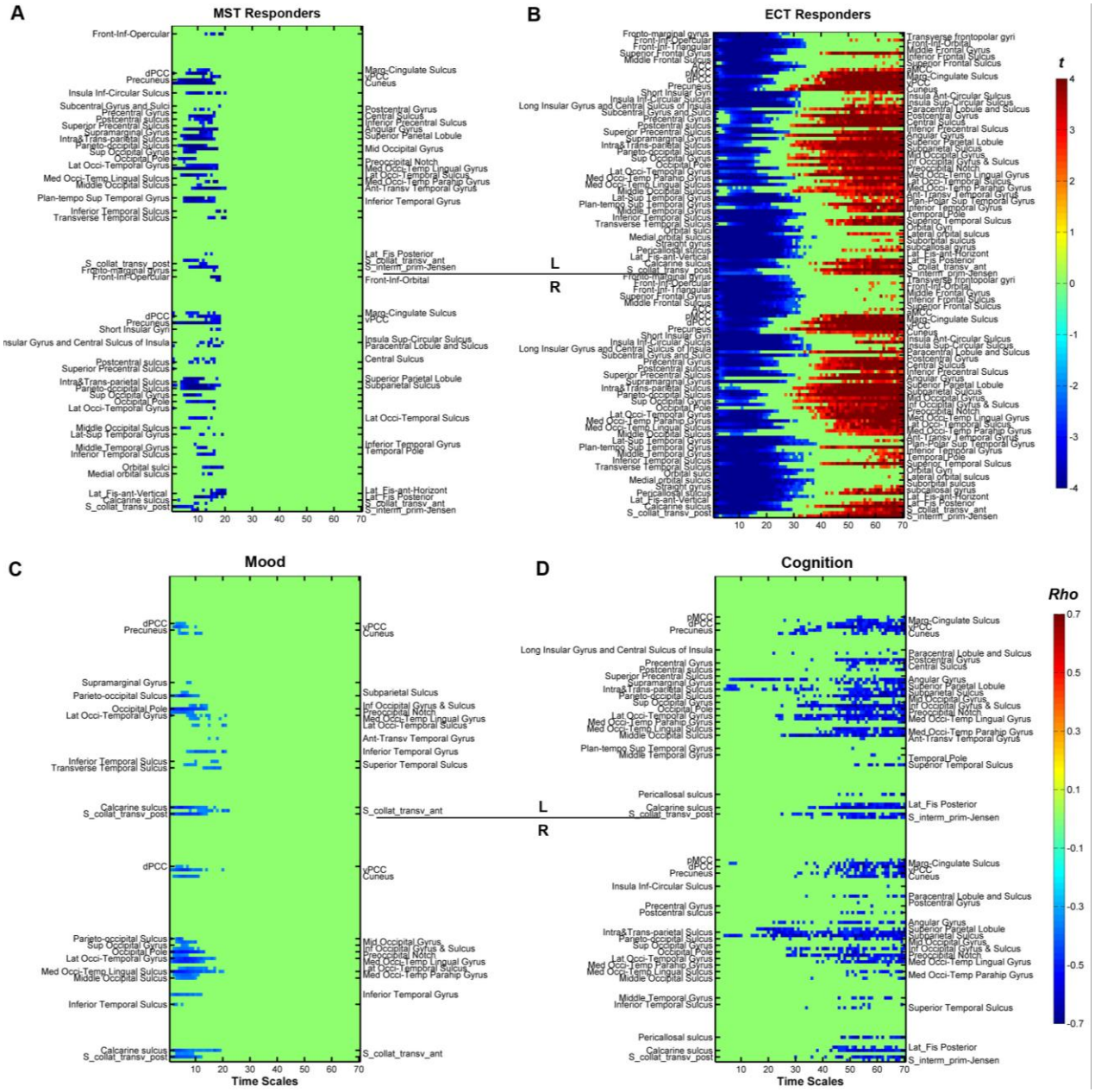

A
MST Responders
B
ECT Responders
C
Mood
D
Cognition
L
R
t
Rho
Time Scales
Time Scales

Figure 4.5 - Seizure Induced Modulation of Complexity and Its Association with Mood and Cognition in the Source Space
In all images, X-axis represents the time scales (1 to 70) and y-axis represents Regions of Interest (ROIs) of the Destrieux Atlas (1 to 148). The ROIs are grouped into left (L: the upper half the images) and right (R: the lower half of the images) hemisphere brain regions separated by the horizontal red line in each figure. Images show the post-hoc test statistics following cluster-based permutation test correction for multiple comparison, depicting only the significant clusters $p<0.05$, labeling only the significant corresponding ROIs and setting to 0 non-significant pixels. **Top:** Images show the t-test statistics comparing MSE post to pre-treatment (blue: decreases; red: increases following treatment). **A.** In MST responders, a wide spread reduction in time-scales less than 20 was observed (*cluster p* <0.05). **B.** By contrast, in ECT responders, there was significant (*cluster p* <0.01) global decrease in time-scales less than 30 and significant (*cluster p <0.01*) global increase in coarser time-scales. In MST responders, the reduction of MSE in fine time-scales was found in several tempro-parieto-occipital (e.g., cuneous, precuneus, posterior-dorsal part of the cingulate gyrus (dPCC), parieto-occipital sulcus, occipital pole, etc) and fronto-central brain regions (e.g., opercular part of the inferior frontal gyrus, central sulcus, pre and post central gyrus, etc). No significant changes were observed in either ECT or MST non-responders. **Bottom: C.** Image illustrate the significant (*p<0.05*) spearman correlation coefficients (rho) between percent change in HAMD and MSE in 34 patients receiving seizure therapy. There was significant negative clusters (*p<0.01*) in time-scale less than 20 factors in tempro-parieto-occipital regions including the bilateral dPCC, bilateral vPCC, bilateral cuneus, precuneus, parieto-occipital sulcus, occipital pole, temporal sulci, bilateral inferior temporal sulcus, bilateral lateral occi-temporal sulcus, bilateral calcarine sulcus, bilateral anterior and posterior transverse collateral sulcus. **D.** Image illustrates spearman correlation coefficients (rho) between percent change in MoCA and MSE across time-scales in 19 patients receiving seizure therapy. There was significant negative clusters (*p<0.005*) in several central, parieto-central, parieto-occipital, occi-temporal, and temporal brain regions (as labeled on the image) across primarily coarser (>30) time-scales.

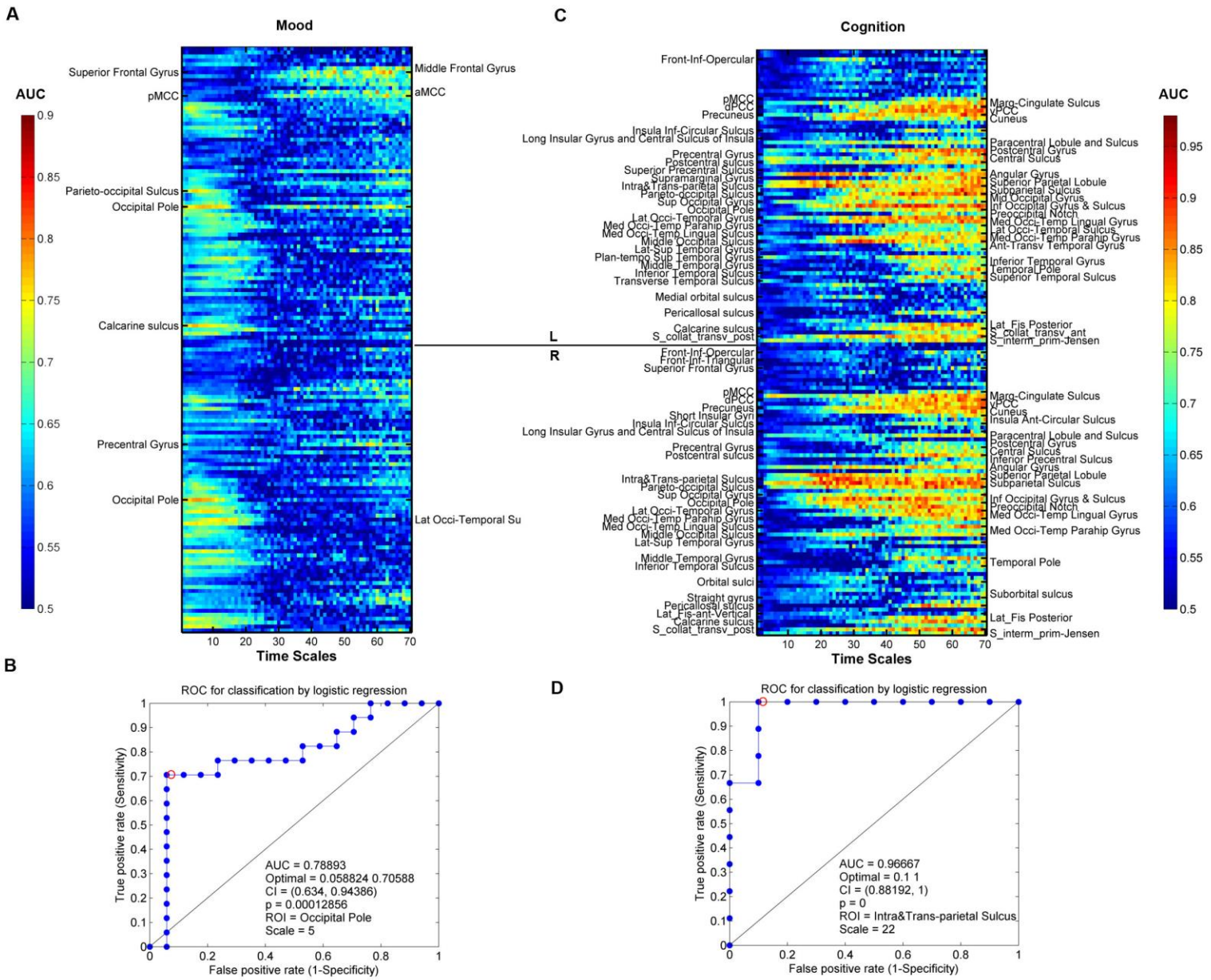

A
Mood
AUC
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
Superior Frontal Gyrus
pMCC
Parieto-occipital Sulcus
Occipital Pole
Calcarine sulcus
Precentral Gyrus
Occipital Pole
Middle Frontal Gyrus
aMCC
Lat Occi-Temporal Su
10 20 30 40 50 60 70
Time Scales
C
Cognition
Front-Inf-Opercular
pMCC
dPCC
Precuneus
Insula Inf-Circular Sulcus
Long Insular Gyrus and Central Sulcus of Insula
Precentral Gyrus
Postcentral sulcus
Superior Precentral Sulcus
Supramarginal Gyrus
Intra&Trans-parietal Sulcus
Parieto-occipital Sulcus
Sup Occipital Gyrus
Occipital Pole
Lat Occi-Temporal Gyrus
Med Occi-Temp Parahip Gyrus
Med Occi-Temp Lingual Sulcus
Middle Occipital Sulcus
Lat-Sup Temporal Gyrus
Plan-tempo Sub Temporal Gyrus
Middle Temporal Gyrus
Inferior Temporal Sulcus
Transverse Temporal Sulcus
Medial orbital sulcus
Pericallosal sulcus
Calcarine sulcus
S_collat_transv_post
L
R
Front-Inf-Opercular
Front-Inf-Triangular
Superior Frontal Gyrus
pMCC
dPCC
Precuneus
Short Insular Gyri
Insula Inf-Circular Sulcus
Long Insular Gyrus and Central Sulcus of Insula
Precentral Gyrus
Postcentral sulcus
Intra&Trans-parietal Sulcus
Pareto-occipital Sulcus
Sup Occipital Gyrus
Occipital Pole
Lat Occi-Temporal Gyrus
Med Occi-Temp Parahip Gyrus
Med Occi-Temp Lingual Sulcus
Middle Occipital Sulcus
Lat-Sup Temporal Gyrus
Middle Temporal Gyrus
Inferior Temporal Sulcus
Orbital sulci
Straight gyrus
Pericallosal sulcus
Lat_Fis-ant-Vertical
Calcarine sulcus
S_collat_transv_post
Marg-Cingulate Sulcus
vPCC
Cuneus
Paracentral Lobule and Sulcus
Postcentral Gyrus
Central Sulcus
Angular Gyrus
Superior Parietal Lobule
Subparietal Sulcus
Mid Occipital Gyrus
Inf Occipital Gyrus & Sulcus
Preoccipital Notch
Med Occi-Temp Lingual Gyrus
Lat Occi-Temporal Sulcus
Med Occi-Temp Parahip Gyrus
Ant-Transv Temporal Gyrus
Inferior Temporal Gyrus
Temporal Pole
Superior Temporal Sulcus
Lat_Fis Posterior
S_collat_transv_ant
S_interm_prim-Jensen
Marg-Cingulate Sulcus
vPCC
Cuneus
Insula Ant-Circular Sulcus
Paracentral Lobule and Sulcus
Postcentral Gyrus
Central Sulcus
Inferior Precentral Sulcus
Angular Gyrus
Superior Parietal Lobule
Subparietal Sulcus
Inf Occipital Gyrus & Sulcus
Preoccipital Notch
Med Occi-Temp Lingual Gyrus
Med Occi-Temp Parahip Gyrus
Temporal Pole
Suborbital sulcus
Lat_Fis Posterior
S_interm_prim-Jensen
AUC
0.95
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0.55
0.5
10 20 30 40 50 60 70
Time Scales
B
ROC for classification by logistic regression
True positive rate (Sensitivity)
False positive rate (1-Specificity)
AUC = 0.78893
Optimal = 0.058824 0.70588
CI = (0.634, 0.94386)
p = 0.00012856
ROI = Occipital Pole
Scale = 5
D
ROC for classification by logistic regression
True positive rate (Sensitivity)
False positive rate (1-Specificity)
AUC = 0.96667
Optimal = 0.1 1
CI = (0.88192, 1)
p = 0
ROI = Intra&Trans-parietal Sulcus
Scale = 22

Figure 4.6 - Prediction of Change in Mood and Cognition in the Source Space

Images depict area under the curve (AUC) of the receiver operating characteristic (ROC) curve of change in multiscale entropy (MSE) in predicting antidepressant (**A**), and cognitive change (**C**) in response to seizure therapy at every Region of Interest (ROI) of the Destrieux Atlas (1 to 148) and each time-scale (1 to 70). Hot colors illustrate higher AUC and better prediction. Change in complexity of low time-scales (1 to 20) in parieto-occipital regions (e.g., parieto-occipital sulcus, occipital pole, calcarine sulcus) offered moderate to good (e.g., AUC of 0.75 to 0.80) prediction performance for change in antidepressant response (**A**). Change in complexity of higher time-scales in parietal brain regions and then spatially globally across time-scales provided excellent (e.g., AUC > 0.9) prediction performance for change in cognition. **B, D.** Figures depict the ROC curve across all possible threshold values of the predictor for an ROI and time-scale for antidepressant response (AUC (right occipital pole, time scale 5) = 0. 79, p <0.0001) (**B**) and change in cognition (e.g., AUC (intra and trans-parietal sulcus, time-scale 22) = 0.97, p <0.00001). (**D**). X-axes represent false positive rates (1-specificity), y-axes the true positive values (sensitivity). The red circle shows the optimum operating point of the ROC curve. **B.** At optimum point, this brain region and scale has 70% sensitivity and 94% specificity. **D.** At optimum point, this ROI and scale has 100% sensitivity and 90% specificity.

4.9 Supplementary Material

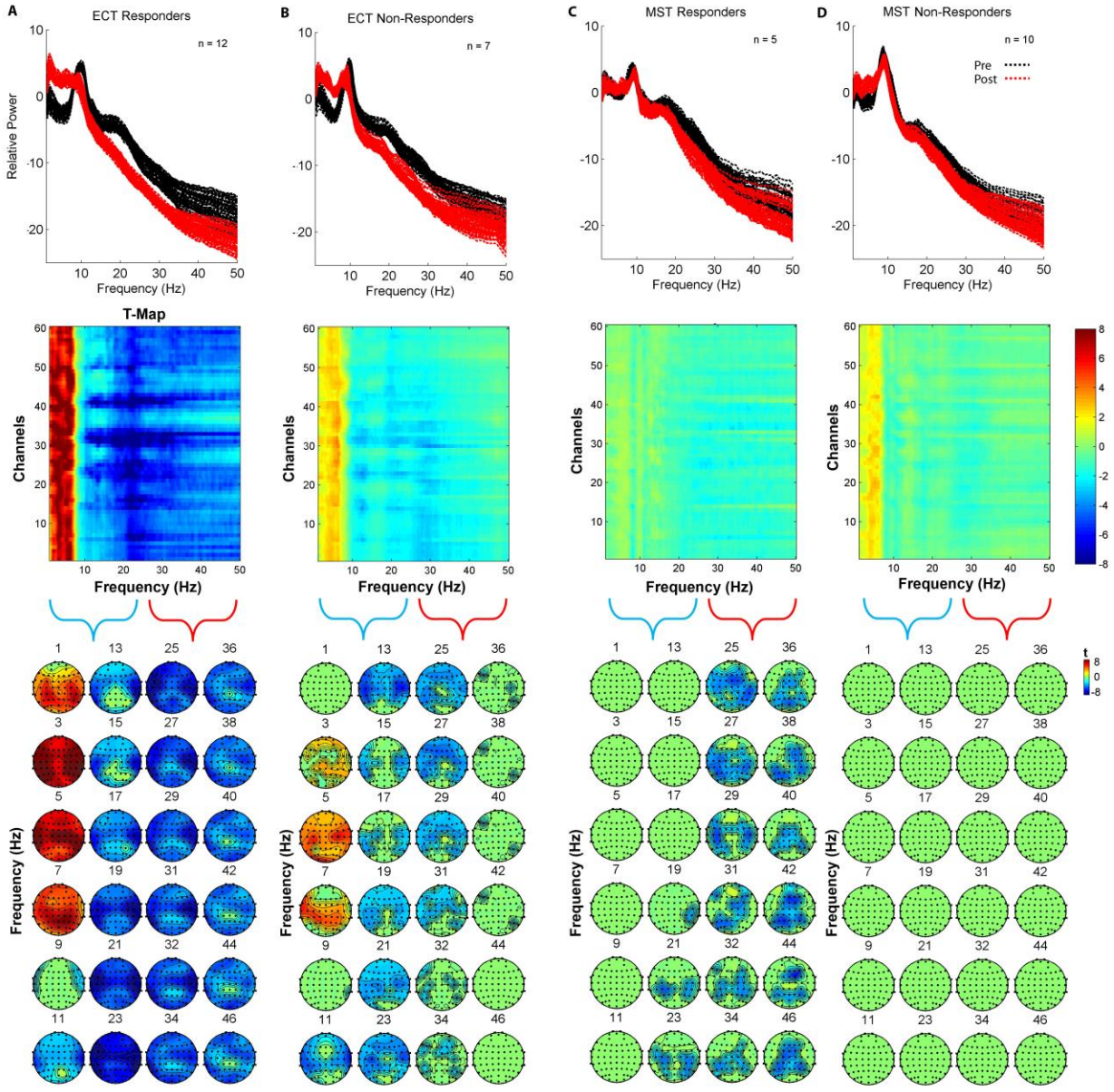


Figure S 4.1 - Effect of Seizure Therapy on Cortical Oscillations
Top. Waveforms depict the relative power spectrum of resting-state eyes-closed EEG pre (black waveforms) and post (red waveforms) electroconvulsive therapy (ECT) and magnetic seizure therapy (MST) in responders (**A, C**) and non-responders (**B, D**). The x-axes are frequency in Hz and the y-axes the relative power in dB. Middle. Images show the original post-hoc test statistics maps comparing the relative power across frequency bands (x-axes) and channels (y-axes) post compared to pre-treatment (blue: decreases; red: increases following treatment) for responders and non-responders. **Bottom.** Each topography reflects the significant t-map depicting only the significant clusters $p<0.05$, setting to 0 non-significant pixels. Topographies highlight the spatial characteristics of a global increase in relative power of frequencies < 8Hz (*cluster p* = 0.018*)* and

a significant (*cluster p<0.001*) global decrease in frequencies > 9 Hz in ECT responders, but a wide spread reduction in relative power of frequencies > 18 Hz (*cluster p*<0.001) in MST Responders. Significant (*cluster p* = 0.042) but less pronounced wide spread increase of 2 to 7Hz and decrease (*cluster p*=0.017) of 10 to 35Hz were observed in ECT non-responders. No significant changes were observed in MST non-responders.

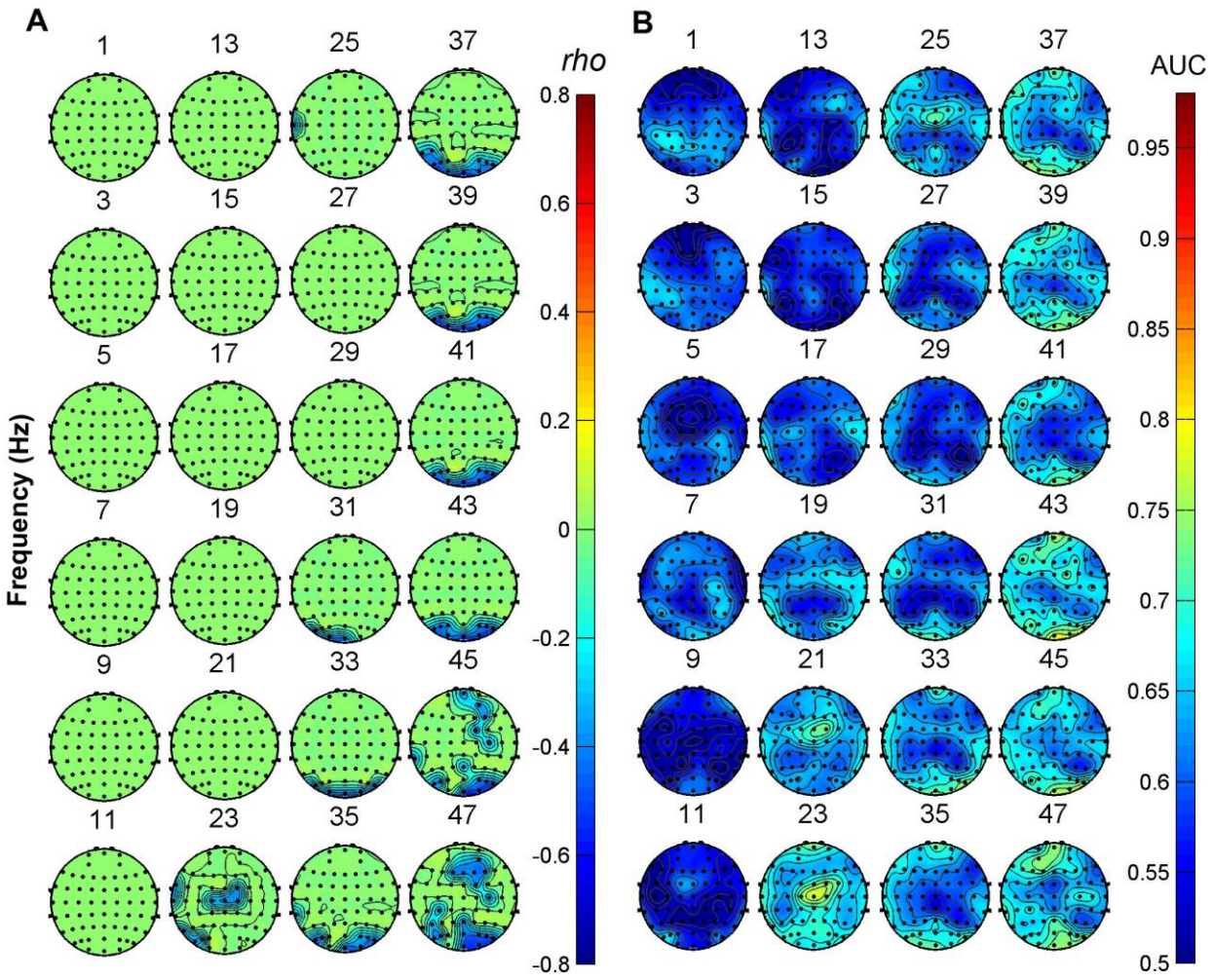


Figure S 4.2 - The Association between Cortical Oscillations and Mood

A. Topographies illustrate all the significant (original $p<0.05$) spearman correlation coefficients (rho) between percent change in HAMD and change in power. All electrodes and frequencies that did not survive the correction for multiple comparisons were set to 0 (green colors). Cluster-based permutation test correction for multiple comparison revealed significant negative clusters ($p<0.01$) in high frequencies (e.g., > 30Hz) in parieto-occipital regions (e.g., P7, P5, PO7, PO5, PO4, PO6, PO8, O1, OZ) and fronto-central regions (e.g., AF4, F1, FZ, F2, F4, FC2) **B.** Topographies depict area under the curve (AUC) of the receiver operating characteristic (ROC) curve of change in relative power of cortical oscillations in predicting change in depressive symptoms in response to seizure therapy at every electrode and frequency. The hot colors illustrate higher AUC and better prediction. Change in cortical oscillations did not provide good accuracy (i.e., AUC > 0.8) in predicting change in depressive symptoms.

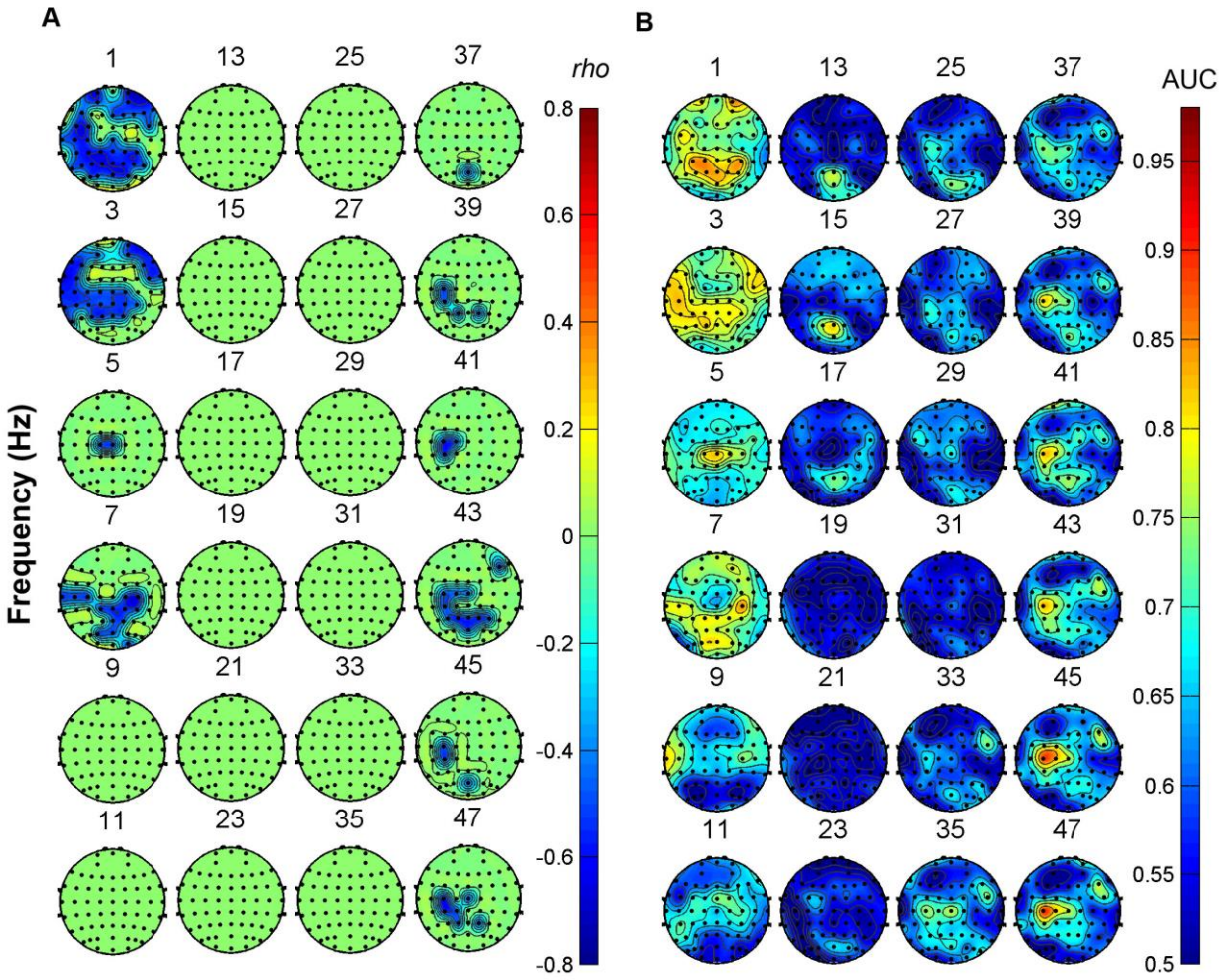


Figure S 4.3 - The Association between Cortical Oscillations and Cognition
A. Topographies illustrate all the significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MoCA and change in power. All electrodes and frequencies that did not survive the correction for multiple comparisons were set to 0 (green colors). Cluster-based non-parametric correction for multiple comparison revealed a significant global negative cluster ($p<0.01$) in slow oscillations (e.g., 1 and 3Hz) and in parieto-central regions (e.g., C1, C3, CZ, CP3, P1, PZ, P2, P4, POZ) in high frequencies (e.g., > 40Hz). **B.** Topographies depict area under the curve (AUC) of the receiver operating characteristic (ROC) curve of change in relative power of cortical oscillations in predicting cognitive change in response to seizure therapy at every electrode and frequency. The hot colors illustrate higher AUC and better prediction. Change in power of low frequency oscillations (e.g., 1-3Hz) a provided good prediction value ($0.8 < AUC < 0.9$) such as in parieto-central regions (e.g., PZ, P2). Power of high frequency oscillations in the left motor cortex (i.e., C3 electrode, 47Hz) provided the best prediction value (AUC = 0.9).

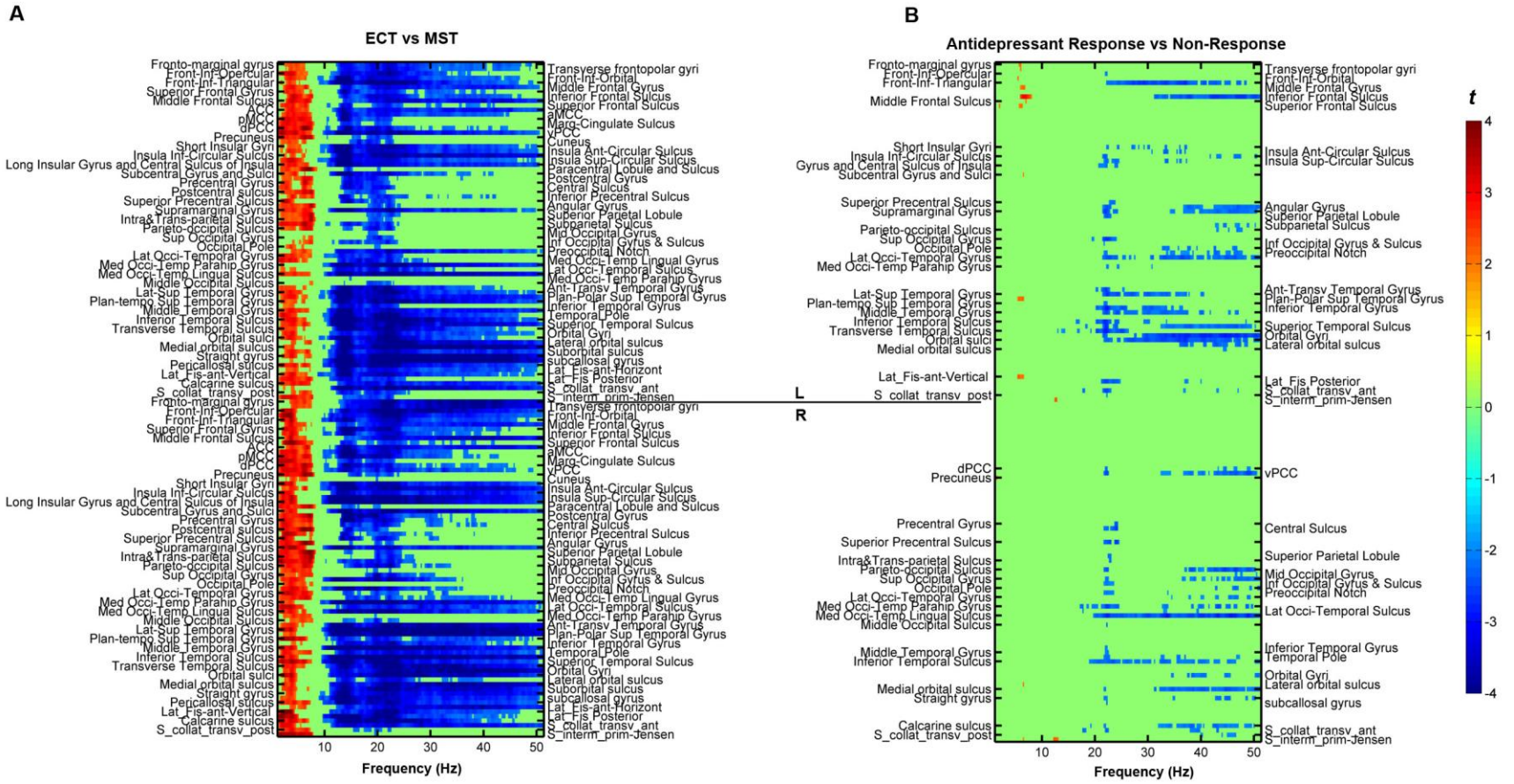


Figure S 4.4 - Effect of Seizure Therapy on Cortical Oscillations in Source Space

In all images, x-axis represents the frequency (1 to 50) in Hertz and y-axis represents Regions of Interest (ROIs) of the Destrieux Atlas (1 to 148). The ROIs are grouped into brain regions in the left (L: the upper half the images) and right (R: the lower half of the images) hemisphere separated by the horizontal black line. Images show the post-hoc independent sample t-test statistics following cluster-based permutation test correction for multiple comparison, depicting only the significant clusters $p<0.05$, labeling only the significant corresponding ROIs, and setting to 0 non-significant pixels. **A.** Image shows the t-test statistics comparing the change in power between participants who received ECT and MST interventions (red: more increase in ECT; blue: more reduction in ECT). This image depicts a significantly greater increase in slow oscillations (<10 Hz) and greater decrease in power of frequencies 20-50 Hz in the ECT group. This effect is spatially global. **B.** Image shows the independent sample t-test statistics comparing the change in power

between participants who were considered responders to seizure therapy (>=50% reduction in HAMD from baseline) and non-responders. This image depicts a greater reduction in power of frequencies 20-50Hz in responders. This finding is spatially global at ~22Hz, but more local in higher frequencies (30-50Hz). Specifically in 30-50Hz, the reduction in power is observed in regions such as the inferior frontal sulcus, left orbital part of the frontal inferior gyrus, bilateral preocciptial notch, orbital gyri, lateral orbital sulcus, lateral occi-temporal sulcus, medial orbital sulcus, bilateral parieto-occiptial sulcus, or bilateral superior parietal lobule.

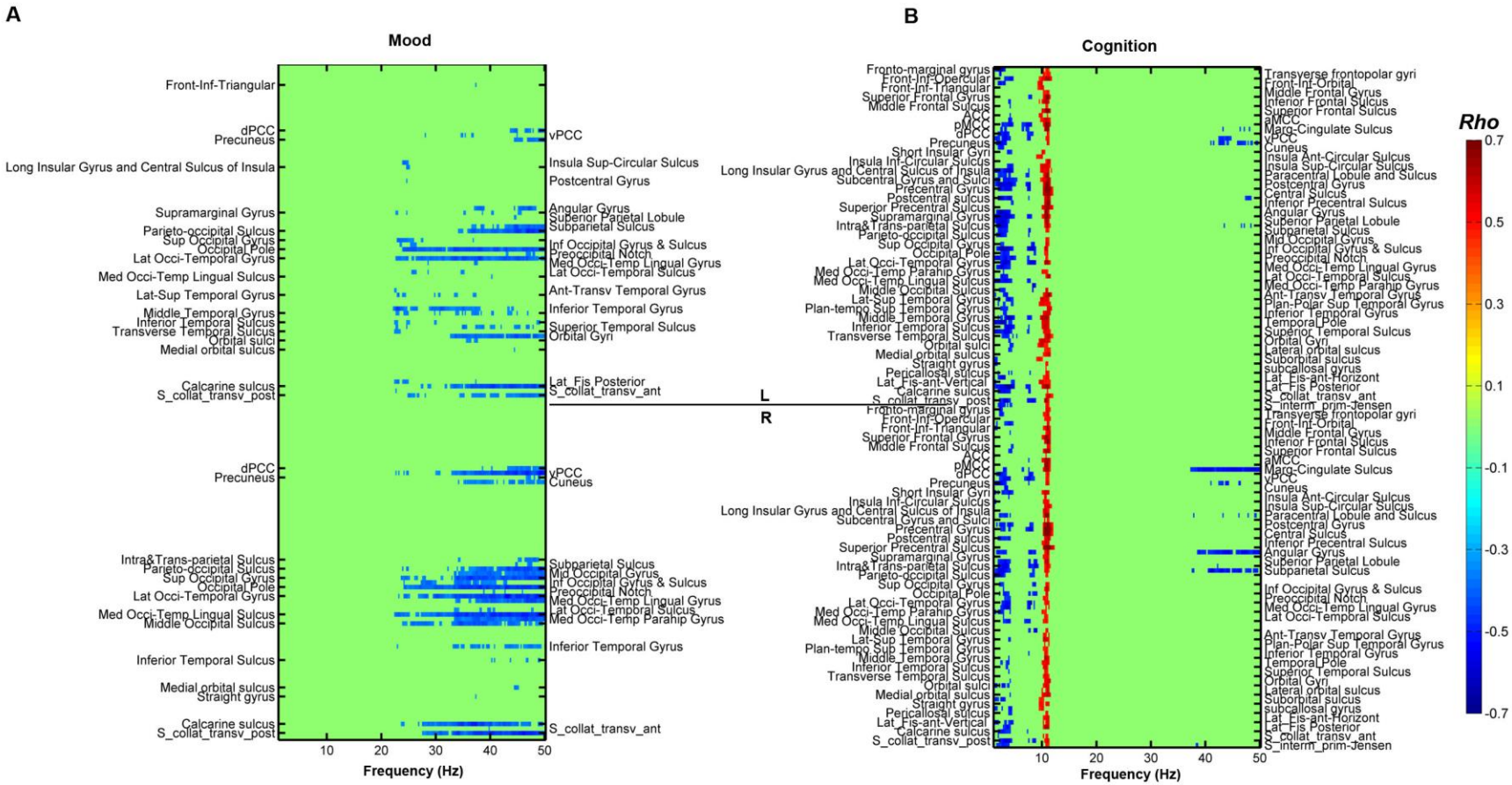


Figure S 4.5 - The Association between Cortical Oscillations and Mood and Cognition in Source Space

A. Image illustrate the significant (*p <0.05*) spearman correlation coefficients (rho) between percent change in HAMD and power in 34 patients receiving seizure therapy. All sources and frequencies that did not survive the correction for multiple comparisons were set to 0 (green colors). Only sources that are significant have been listed. There were significant negative clusters in tempro-parieto-occipital regions (e.g., orbital sulci and gyri, bilateral dPCC, vPCC, precuneus, parieto-occipital sulcus, occipital pole, inferior temporal gyrus, lateral occi-temporal sulcus, etc.) in frequencies higher than 30Hz. **B.** Image illustrates spearman correlation coefficients (rho) between percent change in MoCA and power across time-scales in 19 patients receiving seizure therapy. There was a global negative cluster in slow oscillations and a global positive association at 10 Hz frequency.

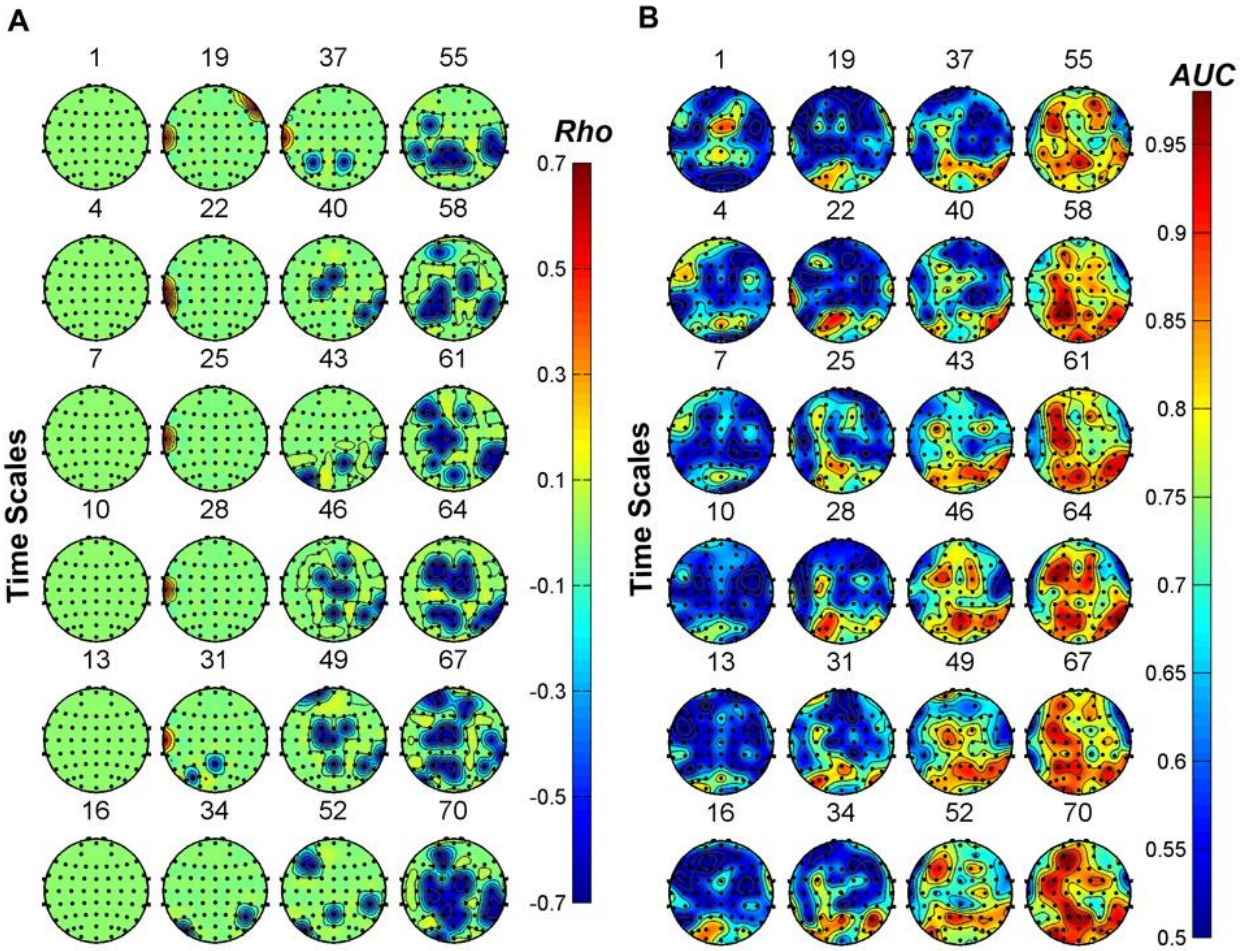


Figure S 4.6 - The Association between Change in Complexity and Autobiographical Memory

A. Topographies illustrate all the significant (original $p<0.05$) spearman correlation coefficients (rho) between percent change in autobiographical memory interview (AMI) and multiscale entropy (MSE) across all time-scales for each electrode. Cluster-based permutation test correction for multiple comparison revealed a significant negative cluster in time-scales higher than 40 across brain regions including the frontoparietal regions. **B.** Topographies depict area under the curve (AUC) of the receiver operating characteristic (ROC) curve of change in MSE in predicting change in AMI in response to seizure therapy at every electrode and time-scale. The hot colors illustrate higher AUC and better prediction. Change in complexity of coarse time-scales (e.g.,>47) in frontoparietal regions had excellent (AUC range: 0.9 to 1.00) prediction performance.

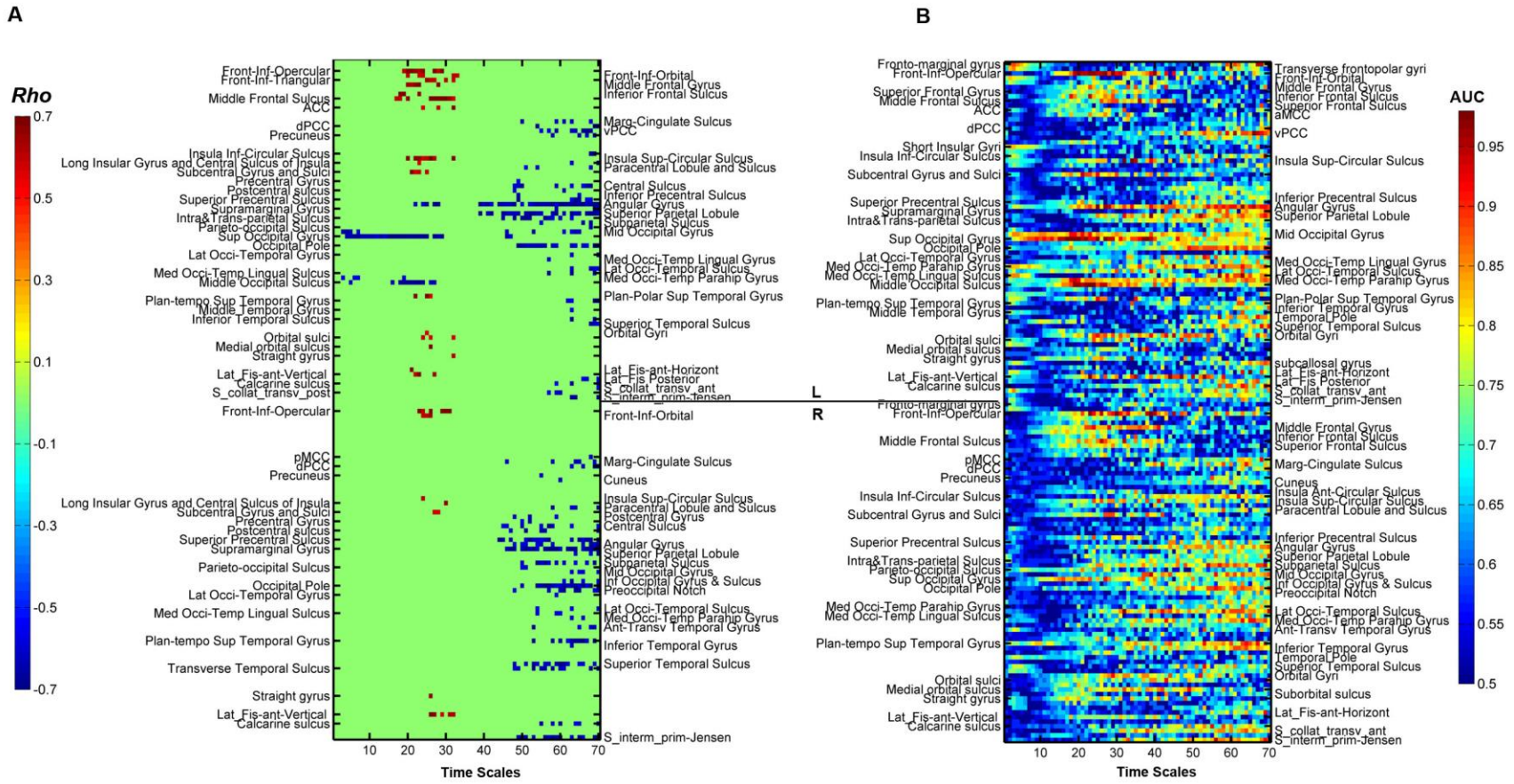


Figure S 4.7 - The Association between Change in Complexity and Autobiographical Memory in Source Space

A. Image illustrates the significant (*p <0.05*) spearman correlation coefficients (rho) between percent change in autobiographical memory interview (AMI) and multiscale entropy (MSE) at every Region of Interest (ROI) of the Destrieux Atlas (1 to 148) and each time-scale (1 to 70). All sources and scales that did not survive the correction for multiple comparisons were set to 0 (green colors). Only sources that are significant have been listed. **B.** Image depicts the area under the curve (AUC) of the receiver operating characteristic (ROC) curve of change MSE in predicting change in AMI in response to seizure therapy at every ROI and each time-scale (1 to 70). Hot colors illustrate higher AUC and better prediction. Change in complexity of higher time-scales in several bilateral frontal and parietal regions provided excellent (AUC range: 0.9 to 1.00) prediction performance for change in AMI.

Chapter 5 – Selective Modulation of Brain Network Dynamics by Seizure Therapy in Treatment-Resistant Depression

In this chapter, we apply EEG frequency analysis and microstate analysis to extract power and global brain-network measures of neural oscillations. We aim to identify whether these measures can provide insight into mechanism of action of ECT and MST.

Sravya Atluri, Willy Wong, Sylvain Moreno, Daniel M. Blumberger, Zafiris J. Daskalakis, and Faranak Farzan. "Selective Modulation of Brain Network Dynamics by Seizure Therapy in Treatment-Resistant Depression." *NeuroImage Clinical.* Accepted (Oct 2018).

5.1 Abstract

Electroconvulsive therapy (ECT) is highly effective for treatment-resistant depression, yet its mechanism of action is still unclear. Understanding the mechanism of action of ECT can advance the optimization of magnetic seizure therapy (MST) towards higher efficacy and less cognitive impairment. Given the neuroimaging evidence for disrupted resting-state network dynamics in depression, we investigated whether seizure therapy (ECT and MST) selectively modifies brain network dynamics for therapeutic efficacy. EEG microstate analysis was used to evaluate resting-state network dynamics in patients at baseline and following seizure therapy, and in healthy controls. Microstate analysis defined four classes of brain states (labelled A, B, C, D). Source localization identified the brain regions associated with these states. An increase in duration and decrease in frequency of microstates was specific to responders of seizure therapy. Significant changes in the dynamics of States A, C and D were observed and predicted seizure therapy outcome (specifically ECT). Relative change in the duration of States C and D was shown to be a strong predictor of ECT response. Source localization partly associated C and D to the salience and frontoparietal networks, argued to be impaired in depression. An increase in duration and decrease in frequency of microstates was also observed following MST, however it was not specific to responders. This study presents the first evidence for the modulation of global brain-network dynamics by seizure therapy. Successful seizure therapy was shown to selectively modulate network dynamics for therapeutic efficacy.

5.2 Introduction

Over one third of patients with major depressive disorder are treatment-resistant and fail to respond to two or more antidepressant medications or psychotherapy (Fava 2003, Berlim and Turecki 2007). This trial-and-error approach can be overbearing for patients in terms of cost, as well as the emotional trauma associated with the prolonged treatment process (Fekadu, Wooderson et al. 2009). Such cases of treatment-resistant depression are also challenging for clinicians. Through the process of identifying an optimal treatment, patients may receive several courses of medication and each treatment course can significantly impact brain circuitry,

regardless of clinical outcome (Mayberg, Brannan et al. 1997, Kennedy, Evans et al. 2001, Mayberg 2003, Fu, Steiner et al. 2013). The complexity of these individual differences can further complicate the process of identifying an appropriate treatment for these patients. To date, electroconvulsive therapy (ECT) remains the most effective treatment for patients with treatment-resistant depression (Kho, van Vreeswijk et al. 2003, Heijnen, Birkenhäger et al. 2010). Yet its mechanism of action is still not known.

It is hypothesized that the brief, generalized seizure triggered by ECT impacts the dynamics of brain networks disrupted in depression (Farzan, Boutros et al. 2014) but also networks involved in cognition, leading to its most common adverse effect: memory impairment (Devanand, Sobin et al. 1995). Magnetic seizure therapy (MST) (Lisanby, Luber et al.) also relies on the principles of seizure induction for therapeutic benefit but unlike ECT, the effect of MST is localized (Deng, Lisanby et al. 2011). Based on the few clinical trials conducted to date, MST improves depressive symptoms (Kayser, Bewernick et al. 2011, Cretaz, Brunoni et al. 2015, Kayser, Bewernick et al. 2015) and suicidal ideation (Sun, Farzan et al. 2016), without the cognitive side effects seen with ECT (Lisanby, Luber et al. 2003, Moscrip, Terrace et al. 2006, Spellman, McClintock et al. 2008, Deng, McClintock et al. 2015). However, in its early stage of development, its efficacy relative to ECT requires further study (Kayser, Bewernick et al. 2011). Identifying brain networks affected by treatment-resistant depression and modified by ECT may allow the optimization of MST as well as the development of non-invasive and non-seizure inducing treatments.

Evidence from neuroimaging studies suggests that specific patterns of brain network dysfunction at rest may contribute to core deficits in cognitive and affective functions underlying neuropsychiatric disorders (Bassett and Bullmore 2006, Garrity, Pearlson et al. 2007, Greicius 2008, Buckner, Sepulcre et al. 2009, Wang, Zhu et al. 2009, Zhang, Wang et al. 2011). It has also been shown that the resting-state of the brain can predict physiological consequences of brain stimulation and treatment outcome (Mayberg 2003, Greicius, Flores et al. 2007, Fox, Buckner et al. 2014). Collectively, the temporal variation in resting-state brain network dynamics may be a significant marker of illness and therapeutic outcome (Honey, Kötter et al. 2007, Chang and Glover 2010, Hutchison, Womelsdorf et al. 2013). To advance the development of novel and

targeted treatments, it is critical to characterize the disruption and changes in brain network dynamics in treatment-resistant depression and by successful treatments such as ECT (Nobler, Oquendo et al. 2001, Perrin, Merz et al. 2012, Abbott, Gallegos et al. 2014).

Using electroencephalography (EEG), functional brain networks and their dynamics can be examined through microstate analysis, a data-driven approach used to measure the spatial stability of brain network dynamics over time (Lehmann, Ozaki et al. 1987, Pascual-Marqui, Michel et al. 1995, Michel and Koenig 2017). Microstate analysis clusters the topographical distributions of spontaneous EEG activity into a set of four classes (A, B, C and D as described in the method) of brain states (i.e., microstates) that each remain stable over a short period of time before transiting into another state (50-120ms) (Strik, Dierks et al. 1995). An increase in the duration of a microstate implies an increase in the probability of that microstate to be followed by itself. Since each microstate is generated by an underlying neuronal population, the temporal characteristics of a microstate (such as rate of change or duration) may be considered as an expression of the dynamic stability of underlying spatial networks (Brodbeck, Kuhn et al. 2012). The duration of microstates is also consistent with the duration of high-level cognitive processes, as shown by evoked-potential studies (Kok 1997). Moreover, microstates were shown to be state-dependent, to vary across age, cognitive state (Koenig, Prichep et al. 2002, Brodbeck, Kuhn et al. 2012, Milz, Faber et al. 2016, Santarnecchi, Khanna et al. 2017) and in response to therapy (Kinoshita, Strik et al. 1995, Rodriguez, Vitali et al. 2002, Kikuchi, Koenig et al. 2007). Studies have also confirmed the reliability of microstates across repeated testing sessions (Khanna, Pascual-Leone et al. 2014).

Microstates were previously linked with brain networks identified through resting-state functional magnetic resonance imaging (fMRI) (Britz, Van De Ville et al. 2010, Musso, Brinkmeyer et al. 2010, Yuan, Zotev et al. 2012) some suggested to be impaired in depression (Veer, Beckmann et al. 2010, Whitfield-Gabrieli and Ford 2012, Kaiser, Andrews-Hanna et al. 2015, Fischer, Keller et al. 2016). For example, microstates C and D were linked to the salience and frontoparietal networks (Britz, Van De Ville et al. 2010) and the relative activation of these networks is hypothesized to be impaired in depression (Hamilton, Furman et al. 2011, Mulders, van Eijndhoven et al. 2015). However, to date, there has been only one study that examined and

reported a decrease in duration of microstates in depression compared to healthy controls (Strik, Dierks et al. 1995). This study was not in treatment-resistant patients and may not extend to treatment-resistant depression (Guo, Sun et al. 2011, Wu, Li et al. 2011, de Kwaasteniet, Rive et al. 2015, Yamamura, Okamoto et al. 2016).

Collectively, the evidence in support of resting-state network abnormalities in depression (Veer, Beckmann et al. 2010, Whitfield-Gabrieli and Ford 2012, Kaiser, Andrews-Hanna et al. 2015, Fischer, Keller et al. 2016), the sensitivity of microstates in detecting intervention-related changes in resting-state networks (Kinoshita, Strik et al. 1995, Rodriguez, Vitali et al. 2002, Kikuchi, Koenig et al. 2007), and the consistency of microstates across repeated testing sessions (Khanna, Pascual-Leone et al. 2014), motivated the utility of EEG microstates in this study. Using the high temporal resolution of EEG to an advantage, we aimed to investigate the therapeutic impact of seizure therapy on network dynamics and the temporal stability of brain network dynamics between patients with treatment-resistant depression and healthy subjects.

Our primary hypotheses were two-fold: (a) microstates C and D, previously associated with the salience and frontoparietal networks implicated in depression, will be modulated by successful seizure therapy; (b) baseline and seizure therapy-induced changes in microstates will be associated with therapeutic outcome, and could explain changes in cognition and suicidal ideation. Our secondary hypothesis was that patients with treatment-resistant depression will present different microstate dynamics compared to healthy controls. These dynamics will be modulated by seizure therapy towards the healthy group dynamics.

5.3 Methods

***Subjects*.** Data was collected from 75 patients (Age: μ=45.7, σ=14.4; 44 females) with a Structured Clinical Interview for the Diagnostic and Statistical Manual of mental disorders (DSM-IV) diagnosis of Major Depressive Disorder who previously did not respond to 2 or more antidepressants (i.e., treatment-resistant depression), and 55 healthy controls (Age: μ=39.2, σ=17.4; 29 females) with written informed consent. Of the 75 patients, follow-up assessments

were conducted for 22 patients receiving ECT and 24 receiving MST. The remaining 29 patients either did not provide their consent for follow-up or withdrew from the study. There were no significant differences in clinical scores, age or sex between the group of patients that did the follow-up assessment and the group of patients that did not.

Seizure Therapy. ECT was administered with spectrum 500Q (MECTA Corporation) according to standards of practice (Sackeim, Prudic et al. 2008). Of the 22 patients who completed ECT treatment, 14 patients received right unilateral ultra-brief (RUL-UB) pulse width ECT, 2 received bitemporal brief pulse (BL) ECT and the rest (6 patients) started on RUL-UB and then switched to BL ECT due to the lack of efficacy. Treatments were administered 2-3 times a week and continued until patients were in remission or improvement plateaued. MST was administered using the MagPro MST using a Twin Coil (MagVenture). The centre of each circular coil was placed over F3 and F4 respectively, using the international 10-20 system for EEG electrode placement. The highest electric field strength roughly corresponds to Fz (Deng, Lisanby et al. 2013) or the dorsomedial prefrontal cortex. Of the 24 patients who received MST, 12 patients received 100Hz MST, 1 patient received 60Hz, 2 received 50Hz and 9 received 25Hz. Treatments were administered 2-3 times per week until remission or up to a maximum of 24 sessions. Please see **Table 5.1** for additional details.

Clinical Assessments. Prior to and following a course of ECT, the 17-scale Hamilton Rating Scale for Depression (HRSD) and Montreal Cognitive Assessment (MoCA) v7.1-7.3 were used to clinically assess severity of depression and global cognition. Montgomery-Åsberg Depression Rating Scale (MADRS) and Beck's Depression Inventory (BDI-II) scale were also used to assess clinical severity of depression and self-rated depression symptoms, respectively. Prior to and following a course of MST, the 24-scale HRSD and Montreal Cognitive Assessment (MoCA) v7.1-7.3 were used to clinically assess severity of depression and global cognition. Severity of suicidal thoughts and overall risk for suicide was assessed using the Scale for Suicidal Ideation (SSI). For all SSI-related analyses, only participants who showed suicidal ideation at baseline were included. The criterion for treatment response was a minimum of 50% improvement in HRSD (final scores were less than 17). Response for BDI, MADRS and SSI was also defined as

a minimum of 50% improvement in score. Demographic and clinical characteristics are presented in **Table 5.1**.

Data Recording and Preprocessing. Eyes-closed rest EEG data was collected within a week prior to the first treatment session and again within 2 weeks after the completion of the last treatment. Data was recorded with the Compumedics (Charlotte, NC, USA) Neuroscan SynAmps 2/RT 64-channel EEG system at 10kHz. During preprocessing, EEG data was downsampled to 1000Hz, divided into 2-second epochs, bandpass-filtered between 1-80Hz, and notch-filtered at 60Hz. With the removal of eye electrodes and other unused channels, the total number of EEG channels used for analysis was 60. Using EEGLAB (Delorme and Makeig 2004), independent component analysis was used to extract eye, muscle and electrode artifacts. Deleted EEG channels were interpolated using spherical spline interpolation (Perrin, Pernier et al. 1989) and data was re-referenced to an average reference. This preprocessing pipeline is currently made available as ERPEEG (http://www.tmseeg.com/multisiteprojects/). Channels were deleted if: (1) they were disconnected during collection for a significant amount of the data collection time (>40%), or (2) heavily contaminated with noise (muscle or spurious artifacts) for a significant part of the data collection time (>40%). On average, 6±3 independent components were removed and 3±1 channels were deleted and interpolated in the data collected from healthy subjects. In the data collected from patients at baseline, 9±5 independent components were removed and 3±1 channels were deleted and interpolated. In the data collected from patients following treatment, 10±5 independent components were removed and 3±2 channels were deleted and interpolated.

Microstate Analysis. Microstate analysis followed the standard procedure outlined in seminal work (**Figure S 5.1**) (Lehmann, Ozaki et al. 1987, Pascual-Marqui, Michel et al. 1995) and was implemented using CARTOOL (Brunet, Murray et al. 2011). Prior to the application of microstate analysis, four minutes of the pre-processed EEG data was bandpass-filtered from 1-30 Hz.

Global field power is a measure of the electric field strength over the scalp and is defined as the variance in electrical activity across EEG electrodes at each time point (**Eq. 1**). The topographical maps at the local maxima peaks of the global field power curve are clustered to

derive the four prototypical microstate classes (Koenig, Prichep et al. 2002). Using a data-driven approach, the optimal number of clusters for the data used in this study was found to be four (Figure S 5.2).

Equation 1:

$$GFP(t) = \sqrt{\frac{\left[\sum_i^N (V_i(t) - V_{mean}(t))\right]^2}{N}},$$

where N represents the number of EEG electrodes (i = 1:60 electrodes) and *v* represents the electrical potential measured over the scalp. In addition, v_i represents the electrical potential measured at electrode (*i*) and time (*t*), and v_{mean} represents the average electrical potential over all electrodes at time (*t*).

In this study, the topographical atomize–agglomerate hierarchical clustering algorithm (Tibshirani and Walther 2005) was applied across all subjects and conditions (global approach). By recalculating microstate classes for each subject or condition, minor differences may be introduced in the microstate topographies. The global clustering approach provides low within-subject error and high test-retest reliability in resting-state microstate analysis (Khanna, Pascual-Leone et al. 2014). Clustered microstates were labelled A, B, C and D as seen in seminal work (Koenig, Prichep et al. 2002) and explained 83% of variance in our data. In the final step, topographical maps at each local maxima point of the global field power curve were assigned to the microstate class of highest correlation using spatial Pearson's product-moment correlation coefficient (**Eq. 2**) (Brandeis, Naylor et al. 1992). Three features were calculated for each of the four microstate classes: (i) average duration, (ii) frequency, and (iii) coverage. Average duration is the amount of time a microstate class remains stable when it appears, in milliseconds; frequency refers to the occurrence of each microstate class per second; and coverage is the percent of recording covered by each microstate class.

Equation 2:

$$r_{spatial} = 1 - \frac{1}{2}\left[\frac{1}{N}\sum_{i=1}^{N}\left\{\frac{u_i - u_{mean}}{\sqrt{\sum_{i=1}^{N}(u_i - u_{mean})^2/N}} - \frac{v_i - v_{mean}}{\sqrt{\sum_{i=1}^{N}(v_i - v_{mean})^2/N}}\right\}^2\right],$$

where N represents the number of EEG electrodes (i=1:60 electrodes) and *u* and *v* represent the 2 different spatial topographies (maps) being correlated. In addition, u_i, v_i represent the electrical potential measured at electrode *i* of the 2 maps and u_{mean}, v_{mean} represent the average electric potential over all electrodes.

Statistical Analysis. To examine the effect of treatment response (≥50% improvement in HRSD) on microstate characteristics following seizure therapy (ECT and MST), a 2x2x4 repeated-measures ANOVA (Lehmann, Faber et al. 2005, Tomescu, Rihs et al. 2014) was conducted for each microstate feature (Duration, Frequency and Coverage) with *RESPONSE* (Responder, Non-responder) as a categorical factor, and *TIME* (Pre, Post) and *MICROSTATE CLASS* (A, B, C, D) as the repeated-measures factors. These ANOVAs were performed on the ECT and MST group data separately as well.

For each microstate feature (Duration, Frequency and Coverage), a 2x4 repeated-measured ANCOVA (Lehmann, Faber et al. 2005, Tomescu, Rihs et al. 2014) was conducted between (2 *GROUPS*) healthy controls and patients with treatment-resistant depression prior to seizure therapy using *MICROSTATE CLASS* (A, B, C and D) as the repeated-measure. For each of the three microstate features, an ANCOVA was performed again between (2 *GROUPS)* healthy controls and patients following seizure therapy using *MICROSTATE CLASS* (A, B, C and D) as the repeated-measure. Age was used as a covariate. There were no significant effects of gender.

Based on our hypothesis that seizure therapy modulates global neural dynamics, planned comparisons were performed to determine whether changes in microstate characteristics were associated with treatment response. For each of the three microstate characteristics (duration, frequency, coverage), paired t-tests were performed to compare the characteristic before and after treatment for each of the four states. The results were corrected for multiple comparisons using the Bonferroni correction method (for the 4 microstates).

A significance level of $\alpha<0.05$ was used for all statistical tests. Pairwise post-hoc comparisons were performed using Tukey-HSD. All planned comparisons were corrected using the Bonferroni method (4 comparisons for the 4 microstate classes).

Correlation and Predictive Analysis. Associations between microstate characteristics and clinical assessments (HRSD, MADRS, BDI, MoCA, SSI) were evaluated with a non-parametric spearman rank-order correlation test and corrected for multiple-comparisons using permutation tests. Receiver operating characteristic (ROC) curves were used to assess predictive value of significant spearman rank-order correlations (i.e., correlations with $p<0.05$). Significance of prediction for the ROC curves was quantified through area under the curve (AUC). Only AUC values greater than or equal to 0.7 (i.e., fair, good, or excellent predictors) are reported in this manuscript. For ROC curves with HRSD, MADRS or BDI, subjects were grouped to be responders (>= 50% improvement in symptoms) or non-responders. For ROC curves with MoCA, subjects were grouped to have cognitive decline if the percent change in MoCA was negative.

Microstate features were correlated with the percent change in clinical scores following therapy. Raw clinical scores (e.g., HRSD) were not used for correlation analysis. Change in microstate features was calculated as (Post-Pre)/Pre*100 where a higher percentage represents an increase in the feature value; change in HRSD, MADRS and BDI were calculated as (Pre-Post)/Pre*100 where a higher percentage represents improvement in depressive symptoms; and lastly change in MoCA and SSI were calculated as (Post-Pre) where a higher value represents improvement in cognition or suicidal ideation symptoms.

Source Localization. The eLORETA algorithm was used to localize the four global-clustered microstates in the source domain. Using LORETA-KEY (Pascual-Marqui, Lehmann et al. 1999), the co-ordinates of the 60 electrodes were identified according to the 10-10 system. A transformation matrix (60x6239) was then derived with a relative regularization parameter of 1.

Power Spectral Density Analysis. The EEGLAB function *spectopo* was used to obtain the power spectrum for each electrode. Relative power was obtained for 1 to 30 Hz (to be consistent with microstate analysis) and was calculated as the ratio of the power at each frequency relative to the sum of power across all frequencies. In addition, relative power was calculated as an average over the following bands: Delta: 1-4Hz; Theta: 4-8Hz; Alpha: 8-14Hz; Low Beta: 14-20Hz; High Beta: 20-30Hz. Spearman rank-order correlations were performed between power in each

band and the characteristics of microstate analysis (duration, frequency and coverage). The results are reported as a correlation matrix. Correlation p-values were also calculated and were Bonferroni-corrected for multiple comparisons (12 comparisons: 4 microstate classes by 3 features).

5.4 Results

A significant improvement in HRSD score was observed following ECT (paired t=−6.6; *df*=21; *p*<0.0001; Cohen's *d*=2.2), following MST (paired t=−4.62, *df*=23; *p*=0.0001; Cohen's *d* = 1.3) and when both the groups were combined as seizure therapy (paired t=−7.8, *df*=45; *p*<0.0001; Cohen's *d*=1.6). In addition, there was a significant improvement in BDI score (paired t=−5.8, *df*=16; *p*<0.0001; Cohen's *d*=1.9) and a decrease in MoCA score (cognition) approaching significance (paired t=−2.5; *df*=5; *p*=0.05; Cohen's *d*=1.2) following ECT. Suicidal ideation (paired t=−4.5, *df*=19; *p*=0.0002; Cohen's *d*=1.3) and cognition scores (paired t=2.7, *df*=20; *p*=0.01; Cohen's *d*=0.48) significantly changed following MST.

5.4.1 Effect of Seizure Therapy on EEG Microstate Dynamics

5.4.1.1 Seizure Therapy (ECT and MST)

A main effect of *Time* (F=15.9; *df*=1,44; *p*=0.0003; η_p^2=0.27) and *Microstate Class* (F=13.2; *df*=3,132; *p*<0.0001; η_p^2=0.23) were observed in the **duration** of microstates. The interaction of *Time* x *Microstate Class* was not significant (F=1.2; *df*=3,132; *p*=0.3; η_p^2=0.026). An effect of *response* (≥50% improvement in HRSD) was not observed (F=0.57; *df*=1,44; *p*=0.5; η_p^2=0.013). Since all the states increased in duration following seizure therapy (ranging between 3.3 to 8.2 ms), paired t-tests were performed to identify which states revealed a statistically significant increase in duration. These were corrected using the Bonferroni method for 4 comparisons (4 microstates). State A (paired t=5.0; *df*=45; *Bonferroni-corrected p*<0.0001; Cohen's *d*=0.77) showed a significant increase in duration following seizure therapy (left panel of Figure 5.1**A**). There was no significant change in the duration of State B (paired t=2.4; *df*=45; *Bonferroni-corrected p*=0.08; Cohen's *d*=0.33), State C (paired t=2.2; *df*=45; *Bonferroni-corrected p*=0.1;

Cohen's *d*=0.32), or State D (paired t=1.4; *df*=45; *Bonferroni-corrected p*=0.7; Cohen's *d*=0.15) (left panels of **Figure 5.1B-D**).

A main effect of *Time* (F=12.4; *df*=1,44; *p*=0.001; η_p^2=0.22) and *Microstate Class* (F=3.5; *df*=3,132; *p*=0.02; η_p^2=0.07) were observed in the **frequency** of microstates. The interaction of *Time* x *Microstate Class* approached significance (F=2.4; *df*=3,132; *p*=0.06; η_p^2=0.052). Post-hoc Tukey-HSD tests revealed that State B (HSD=4.7; *df*=132; *p*=0.03; Cohen's *d*=0.65) (left panel of **Figure 5.2B**), State C (HSD=5.6; *df*=132; *p*=0.004; Cohen's *d*=0.78) (left panel of **Figure 5.2C**) and State D (HSD=6.2; *df*=132; *p*=0.0008; Cohen's *d*=0.87) (left panel of **Figure 5.2D**) *significantly* decreased in frequency following seizure therapy. There was no significant change in the frequency of State A (HSD=1.3; *df*=132; *p*=0.99; Cohen's *d*=0.19) (left panel of **Figure 5.2A**). An effect of *response* was not observed (F=0.54; *df*=1,44; *p*=0.5; η_p^2=0.012).

Apart from the main effect of *Microstate Class* (F=8.6; *df*=3,132; *p*<0.0001; η_p^2=0.16), no significant effects were observed in the **coverage** of microstates (**Figure S 5.3**).

Planned comparisons were conducted to investigate the effect of response based on our hypotheses. The increase in State A duration was specific to responders of seizure therapy (t=6.3; *df*=19; *p*<0.0001; Bonferroni-corrected *p*<0.0001; Cohen's *d*=1.1) (middle panel of **Figure 5.1A**). This effect was not observed in non-responders for State A (t=2.2; *df*=25; *p*=0.03; Bonferroni-corrected *p*=0.1; Cohen's *d*=0.48) (right panel of **Figure 5.1A**). The decrease in frequency of State B (t=−3.3; *df*=19; *p*=0.004; Bonferroni-corrected *p*=0.01; Cohen's *d*=0.56), State C (t=−4.9; *df*=19; *p*=0.0001; Bonferroni-corrected *p*=0.0004; Cohen's *d*=0.80) and State D (t=−2.8; *df*=19; *p*=0.01; Bonferroni-corrected *p*=0.04; Cohen's *d*=0.53) was specific to responders of seizure therapy (middle panels of **Figure 5.2B-D**). This effect was not observed in non-responders for State B (t=−1.1; *df*=25; *p*=0.3; Cohen's *d*=0.15), State C (t=−1.3; *df*=25; *p*=0.19; Cohen's *d*=0.31) or State D (t=−1.5; *df*=25; *p*=0.2; Cohen's *d*=0.28) (right panels of **Figure 5.2B-D**).

5.4.1.2 Electroconvulsive Therapy

A main effect of *Time* (F=8.3; *df*=1,20; *p*=0.009; η_p^2=0.29) and *Microstate Class* (F=5.9; *df*=3,60; *p*=0.001; η_p^2=0.23) were observed in the **duration** of microstates. The interaction of *Time* x *Microstate Class* was not significant (F=0.71; *df*=3,60; *p*=0.6; η_p^2=0.034). An effect of *response* was not observed (F=0.002; *df*=1,20; *p*=0.97; η_p^2=0.00008). Since all the states increased in duration following ECT (ranging between 6.1 to 11.1ms), paired t-tests were performed to identify which states revealed a statistically significant increase in duration. These were corrected using the Bonferroni method for 4 comparisons (4 microstates). State A (paired t=5.7; *df*=21; *Bonferroni-corrected p*<0.0001; Cohen's *d*=1.08) showed a significant increase in duration following ECT (left panel of **Figure 5.3A**). There was no significant change in the duration of State B (paired t=1.7; *df*=21; *Bonferroni-corrected p*=0.4; Cohen's *d*=0.36), State C (paired t=1.7; *df*=21; *Bonferroni-corrected p*=0.4; Cohen's *d*=0.43) or State D (paired t=2.3; *df*=21; *Bonferroni-corrected p*=0.1; Cohen's *d*=0.35) (left panels of **Figure 5.3B-D**).

A main effect of *Time* (F=7.3; *df*=1,20; *p*=0.01; η_p^2=0.27), and an interaction effect of *Time x Microstate Class* (F=3.4; *df*=3,60; *p*=0.02; η_p^2=0.15) were observed in the **frequency** of microstates. The main effect of *Microstate Class* was not significant (F=1.3; *df*=3,60; *p*=0.3; η_p^2=0.06). All states decreased in frequency following seizure therapy. Post-hoc Tukey-HSD tests revealed that State B (HSD=4.4; *df*=60; *p*=0.03; Cohen's *d*=1.0) (left panel of **Figure 5.4B**), State C (HSD=5.7; *df*=60; *p*=0.002; Cohen's *d*=1.24) (left panel of **Figure 5.4C**) and State D (HSD=6.6; *df*=60; *p*=0.0003; Cohen's *d*=1.46) (left panel of **Figure 5.4D**) *significantly* decreased in frequency following ECT. There was no significant change in the frequency of State A (HSD=0.54; *df*=60; *p*=0.99; Cohen's *d*=0.11) (left panel of **Figure 5.4A**). An effect of *response* was not observed (F=0.04; *df*=1,20; *p*=0.8; η_p^2=0.002).

Apart from the main effect of *Microstate Class* (F=3.1; *df*=3,60; *p*=0.03; η_p^2=0.13), no significant effects were observed in the **coverage** of microstates (**Figure S 5.4**).

Planned comparisons revealed that the increase in State A duration was specific to responders of ECT (t=8.5; *df*=12; *p*<0.0001; Bonferroni-corrected *p*<0.0001; Cohen's *d*=1.60) (middle panel of **Figure 5.3A**). This effect was not observed in non-responders for State A (t=1.7; *df*=8; *p*=0.1;

Cohen's *d*=0.52) (right panel of **Figure 5.3A**). A decrease in the frequency of State B (t=−3.1; *df*=12; *p*=0.008; Bonferroni-corrected *p*=0.03; Cohen's *d*=0.67) (middle panel of **Figure 5.4B**), State C (t=−4.0; *df*=12; *p*=0.002; Bonferroni-corrected *p*=0.008; Cohen's *d*=0.90) (middle panel of **Figure 5.4C**) and State D (t=−2.9; *df*=12; *p*=0.01; Bonferroni-corrected *p*=0.04; Cohen's *d*=0.71) (middle panel of **Figure 5.4D**) was specific to responders of ECT. This effect was not observed in non-responders for State B (t=−0.78; *df*=8; *p*=0.5; Cohen's *d*=0.24), State C (t=−2.0; *df*=8; *p*=0.08; Cohen's *d*=0.67) or State D (t=−1.6; *df*=8; *p*=0.2; Cohen's *d*=0.71) (right panels of **Figure 5.4B-D**).

5.4.1.3 Magnetic Seizure Therapy

Similar to ECT, there was an increase in duration and decrease in frequency of microstates following MST. A main effect of *Time* (F=6.8; *df*=1,22; *p*=0.01; η_p^2=0.24) and *Microstate Class* (F=6.4; *df*=3,66; *p*=0.0007; η_p^2=0.23) were observed in the **duration** feature. The interaction effect of *Time x Microstate Class* was not significant (F=0.11; *df*=3,66; *p*=0.96; η_p^2=0.005). An effect of *response* was not observed (F=2.3; *df*=1,22; *p*=0.2; η_p^2=0.094). In addition, a main effect of *Time* (F=4.4; *df*=1,22; *p*=0.04; η_p^2=0.16) and a main effect of *Microstate Class* (F=4.0; *df*=3,66; *p*=0.01; η_p^2=0.15) were observed in the **frequency** feature. The interaction effect of *Time x Microstate Class* was not significant (F=0.075; *df*=3,66; *p*=0.97; η_p^2=0.003). An effect of *response* was not observed (F=2.4; *df*=1,22; *p*=0.1; η_p^2=0.10). Apart from the main effect of *Microstate Class* (F=5.8; *df*=3,66; *p*=0.001; η_p^2=0.21), no significant effects were observed in the **coverage** of microstates. See supplementary figures (**Figure S 5.5**, **Figure S 5.6** and **Figure S 5.7**). Planned comparisons were not significant (see supplementary **Table 5.2**).

5.4.2 Correlation and Prediction Analysis Results

5.4.2.1 Seizure Therapy (ECT and MST)

An increase in the duration of State A correlated with improvement in depressive symptoms (HRSD) (r = 0.33, 95% CI 0.044 to 0.57, *p-corrected*=0.02). The increase was also a fair predictor of improvement in depressive symptoms (HRSD) (AUC=0.71, *p*=0.003) (**Figure 5.5**).

5.4.2.2 Electroconvulsive Therapy

A decrease in State D duration following ECT treatment significantly correlated with improvement in self-rated depressive symptoms (BDI) (r = −0.55, 95% CI −0.82 to −0.09, *p-corrected*=0.02) (left panel of **Figure 5.6B**). The decrease was also a good predictor of improvement in self-rated depressive symptoms (BDI) (AUC=0.83, *p*=0.0007) (right panel of **Figure 5.6B**). Based on the association of State C to the salience and State D to the frontoparietal network (Britz, Van De Ville et al. 2010) and based on research indicating that the salience network facilitates the activation of the frontoparietal network (Menon and Uddin 2010), we hypothesized that the change in State D duration relative to the change in State C duration will also significantly correlate with clinical outcome. The log ratio between change in State D duration and change in State C duration correlated with improvement in self-rated depressive symptoms (BDI) (r = −0.67, 95% CI −0.87 to −0.28, *p*=0.003) (left panel of **Figure 5.6C**) and clinical depression scores (MADRS) (r = −0.50, 95% CI −0.78 to −0.04, *p*=0.03). The ratio was also an excellent predictor of improvement in self-rated depressive symptoms (BDI) (AUC=0.97, *p*<0.0001) (right panel of **Figure 5.6C**).

Furthermore, an increase in State A coverage correlated with the improvement in self-rated depressive symptoms (BDI) (r = 0.57, 95% CI 0.12 to 0.82, *p-corrected*=0.02) (left panel of **Figure 5.6A**). The increase was also a fair predictor of improvement in self-rated depressive symptoms (BDI) (AUC=0.79, *p*=0.005) (right panel of **Figure 5.6A**).

5.4.2.3 Magnetic Seizure Therapy

Changes in microstate dynamics following MST were not associated with change in depressive symptoms (see supplementary **Table 5.3**). Baseline characteristics of all microstates were also shown to predict suicidal ideation response (right panels of **Figure 5.7A**). At baseline, a shorter duration of State A (r = −0.57, 95% CI −0.81 to −0.17, *p-corrected*=0.01), State B (r = −0.49, 95% CI −0.77 to −0.06, *p-corrected*=0.03), State C (r = −0.49, 95% CI −0.77 to −0.06, *p-corrected*=0.03) and a higher frequency of State D (r = 0.52, 95% CI 0.10 to 0.78, *p-corrected*=0.02) predicted suicidal ideation response. Furthermore, a decrease in State B frequency was shown to be correlated with improvement in cognition (r = −0.50, 95% CI −0.77

to −0.09, *p-corrected*=0.01) (middle panel of **Figure 5.7B**) and was a good predictor of improvement in cognition (AUC=0.80, *p*=0.002) (right panel of **Figure 5.7B**).

5.4.3 Microstate Dynamics in Patients with Treatment-Resistant Depression vs Healthy Controls

Patients revealed an increased duration and decreased frequency of microstates compared to healthy subjects (left panels of **Figure 5.8A, C**). By comparing healthy subjects with patients before treatment, a significant main effect of *Group* (F=4.9; *df*=1,127; *p*=0.03; η_p^2=0.04), *Microstate Class* (F=12.2; *df*=3,381; *p*<0.0001; η_p^2=0.09) and an interaction effect of *Group x Microstate Class* approaching significance (F=2.1; *df*=3,381; *p*=0.09; η_p^2=0.02) was observed in the **duration** of microstates. A main effect of *Group* (F=4.4; *df*=1,127; *p*=0.03; η_p^2=0.03), *Microstate Class* (F=7.6; *df*=3,381; *p*=0.01; η_p^2=0.06) and an interaction effect of *Group x Microstate Class* (F=3.8; *df*=3,381; *p*=0.01; η_p^2=0.03) was observed in the **frequency** of microstates.

A significant main effect of *Microstate Class* (F=1.0; *df*=3,381; *p*<0.0001; η_p^2=0.08) and a significant interaction of *Group x Microstate Class* (F=3.5; *df*=3,381; *p*=0.01; η_p^2=0.03) was observed with the **coverage** feature between healthy subjects and patients at baseline (**Figure 5.8E**). The main effect of *Group* was not significant (F=2.1; *df*=1,127; *p*=0.2; η_p^2=0.02).

5.4.4 Microstate Dynamics in Patients with Treatment-Resistant Depression following Seizure Therapy vs Healthy Controls

The longer duration and lower frequency of microstates in patients persisted following treatment (**Figure 5.8B, D**). Comparing healthy subjects with patients after seizure therapy, there was a main effect of *Group* (F=19.4; *df*=1,98; *p*<0.0001; η_p^2=0.17) and *Microstate Class* (F=12.1; *df*=3,294; *p*<0.0001; η_p^2=0.11) in the **duration** feature. The interaction of *Group x Microstate Class* was not significant (F=1.1; *df*=3,294; *p*=0.36; η_p^2=0.01). In the **frequency** feature, there was a significant main effect of *Group* (F=16.3; *df*=1,98; *p*=0.0001; η_p^2=0.14) and *Microstate Class* (F=6.1; *df*=3,294; *p*=0.0005; η_p^2=0.06). The interaction of *Group x Microstate Class* was not significant (F=1.5; *df*=3,294; *p*=0.22; η_p^2=0.02).

The main effect of *Microstate Class* was significant in the **coverage** feature (F=8.8; *df*=3,294; *p*<0.0001; η_p^2=0.08). The main effect of Group (F=1.4; *df*=1,98; *p*=0.24; η_p^2=0.003) and the interaction effect of *Group x Microstate Class* (F=1.2; *df*=3,294; *p*=0.32; η_p^2=0.01) were not significant (**Figure 5.8E**).

5.4.5 Source Localization of the Global-Clustered Microstates

In **Figure 5.9**, the global-clustered microstates are shown alongside their corresponding eLORETA images. All states show a common neural generator in the posterior cingulate and cingulate gyrus. This has been shown in previous literature(Pascual-Marqui, Lehmann et al. 2014). State A was associated with the left superior and middle temporal gyrus (**Figure 5.9A**). State B with the cuneus and precuneus of the occipital lobe (**Figure 5.9B**). State C was best associated with the anterior cingulate, insula and cuneus and precuneus of the occipital lobe (**Figure 5.9C**). Finally, State D with the paracentral lobe of the frontal lobe, the precuneus of the parietal lobe, the parahippocampal gyrus of the limbic lobe, and the lingual gyrus of the occipital lobe (**Figure 5.9D**). See **Figure S 5.8**, **Figure S 5.9**, **Figure S 5.10**, and **Figure S 5.11** for more detailed images of source localization.

5.4.6 Power Spectral Density Analysis

5.4.6.1 Neuronal Oscillations in Patients vs Healthy Controls

No significant differences were observed in relative power between the healthy and patient groups after cluster-based permutation correction for multiple comparisons (**Figure S 5.12**).

5.4.6.2 Effect of Seizure Therapy on Neuronal Oscillations

Following cluster-based permutation correction, both ECT responders and non-responders revealed an increase in relative power of slow cortical oscillations (1-7Hz) and a decrease in relative power of oscillations above 10Hz (**Figure S 5.13A-B**). The increase in relative power of slow cortical oscillations was not observed in responders or non-responders of MST (**Figure S**

5.13C-D). Common to ECT and MST however, was a decrease in relative power above 17Hz. In responders of MST a decrease in relative power was observed above 17Hz. In non-responders of MST, a decrease in relative power was observed above 11Hz.

5.4.7 Microstate Dynamics and Neuronal Oscillations

In the ECT group, the increase in slow oscillations was not correlated with microstate characteristics (**Figure S 5.14**). Only changes in the low beta and high beta power bands correlated with change in duration and frequency of microstates. A significant change in duration of microstates following ECT was specific to State A. However, changes in low beta and high beta were correlated with changes in duration of States B and C. Significant change in frequency of microstates following ECT was specific to States B, C and D. Yet, change in low beta correlated with change in frequency of States A and D and change in high beta correlated with change in frequency of States A, B and D. In the MST group, none of the power bands significantly correlated with microstate characteristics (**Figure S 5.15**).

5.5 Discussion

This study presents the first evidence for the modulation of resting-state EEG microstate dynamics by seizure therapy in patients with treatment-resistant depression. First, several changes in microstate dynamics following seizure therapy suggested that ECT selectively modifies global brain-network dynamics. An increase in the duration of State A and a decrease in the frequency of States B, C and D were associated with response to seizure therapy (specifically ECT). Although there was a change in network dynamics following MST, it was not network-specific and it was not specific to responders. However, a decrease in the frequency of State B was associated with improvement in cognition following MST. In addition, baseline microstate dynamics were shown to predict suicidal ideation response to MST (shorter duration of States A, B, C and a higher frequency of State D). Finally, patients revealed increased duration and decreased frequency of microstates compared to healthy subjects. Following seizure therapy, this difference was greater between patients and healthy subjects. Collectively, these findings provide

insight into the role of global network dynamics in the potential mechanism of action of seizure therapy for treatment-resistant depression.

Although the most effective treatment for treatment-resistant depression is ECT, its underlying mechanism of action is not clearly understood. In this study, a significant increase in the duration of State A and decrease in the frequency of States B, C and D was observed in responders of seizure therapy (specifically ECT). We hypothesize that this might reflect a relative stabilization (or reduction) of microstate dynamics since it infers that the microstate occurs for a longer duration of time and is less variable (i.e., more stable). This finding may support one of the main theories on the efficacy of ECT, the anticonvulsant hypothesis (Coffey, Lucke et al. 1995, Sackeim 1999). The anticonvulsant hypothesis suggests that an activation of inhibitory mechanisms initially occurs to inhibit seizures caused by ECT but eventually leads to the inhibition of hyperactive networks in depression, which may lead to reduced global network dynamics. There is a large amount of accumulating evidence for the anticonvulsant hypothesis, such as increased cortical GABA (Sanacora, Mason et al. 2003), decreased regional brain metabolism (Hoy, Thomson et al. 2013), increased slow-wave EEG activity as well as decreased seizure duration and increased seizure threshold over the course of ECT (Sackeim 1999). The potential stabilization of microstate (i.e., global network) dynamics following seizure therapy further adds to this line of evidence.

In a recent comprehensive meta-analysis of resting-state fMRI studies (Kaiser, Andrews-Hanna et al. 2015), depression was associated with aberrant interactions between the salience, frontoparietal and default-mode networks, argued to be the facilitators of depressive symptoms. With recent progress in the integration of fMRI and EEG data, a few studies have explored the association between cortical microstate activity and resting-state fMRI networks (Jann, Kottlow et al. 2010, Musso, Brinkmeyer et al. 2010, Schwab, Koenig et al. 2015). In *Britz et al.* (Britz, Van De Ville et al. 2010), the salience and frontoparietal networks were associated with States C and D. As hypothesized in this study, the frequency of States C and D decreased following seizure therapy and these changes were associated with treatment response (specifically ECT). In addition, change in State D duration relative to the change in State C duration correlated with improvement in depressive symptoms following ECT. This suggests that the interaction between

the neural generators underlying these microstates may be impaired in patients with treatment-resistant depression, and may be linked to the role of State C as a dynamic "switching network". This role of State C has been widely postulated in schizophrenia-related microstate research (Rieger, Hernandez et al. 2016).

As an alternate to ECT, MST was proposed to minimize cognitive side effects while maintaining antidepressant efficacy. A few studies have even associated MST with improvement in cognition including visual-spatial learning, memory and phonological tasks (Lisanby, Luber et al. 2003, Kayser, Bewernick et al. 2011, Kayser, Bewernick et al. 2015). The association between change in State B frequency and improvement in cognition with MST may be in line with these findings since State B has been associated with the parietal and occipital-parietal areas of spatial-visualization and verbalization (Britz, Van De Ville et al. 2010, Milz, Pascual-Marqui et al. 2016). In addition to cognition, MST was previously associated with remission of suicidal ideation. Baseline markers of inhibitory neurotransmission were shown to predict therapeutic efficacy of MST in reducing suicidal ideation (Sun, Farzan et al. 2016). In our study, baseline microstate dynamics (of all four states) predicted the therapeutic efficacy of MST in reducing suicidal ideation. Source localization revealed that all four microstates had in common the posterior cingulate cortex and the precuneus, regions linked with the default-mode network. This suggests that MST may be able to target the impaired default-mode network in treatment-resistant depression (Kaiser, Andrews-Hanna et al. 2015). The coil position of MST in this study supports this hypothesis as the greatest induced electrical field was over one of the hubs of the default-mode network, the dorsomedial prefrontal cortex. We suggest that MST modulates neural networks impaired in treatment-resistant depression but may not be as robust as ECT due to the sub-optimal induced electric field potentials.

As mentioned, patients that have previously received two or more courses of antidepressants with no clinical outcome are treatment-resistant. Detailed fMRI studies between healthy subjects and patients with and without treatment-resistant depression have indicated that different functional connectivity patterns may be associated with treatment-resistance (Guo, Sun et al. 2011, Wu, Li et al. 2011, Yamamura, Okamoto et al. 2016), potentially due to the previous antidepressant exposure in treatment-resistant depression. Several neuroimaging studies have

shown that treatment can change resting-state brain dynamics regardless of clinical outcome (Mayberg, Brannan et al. 1997, Kennedy, Evans et al. 2001, Mayberg 2003, Fu, Steiner et al. 2013). In this study, a longer duration and lower frequency of microstates were observed in patients with treatment-resistant depression compared to healthy subjects. We link this effect to the neurotropic medications previously taken by these patients. Although the patients in this study did not respond to their previous treatments, we hypothesize that each treatment they received may have had an effect on global brain dynamics. It has been shown that benzodiazepines and antipsychotics can modulate microstate dynamics (Kinoshita, Strik et al. 1995). Following seizure therapy, a larger increase in duration and decrease in frequency of microstates was observed in our study, suggesting that antidepressants and seizure therapy may modulate global brain dynamics in a similar manner. Considering the association of these changes to therapeutic outcome in seizure therapy, we hypothesize that seizure therapy overcomes the inadequacy of medications in treatment-resistant depression through a stronger impact on network dynamics. However, in the absence of a control group (i.e., patients without treatment-resistant depression) and longitudinal assessments, it remains to be investigated whether this is an effect of treatment-resistance, medication, or both.

The higher efficacy of seizure therapy in treatment-resistant depression has been linked with the stimulation of thalamic oscillatory pacemakers and re-setting of neural dynamics (Farzan, Boutros et al. 2014). Studies have also highlighted the importance of temporal variability in affective and cognitive brain functions (Tononi, Sporns et al. 1994, Stam, Jones et al. 2006, Rubinov and Sporns 2010). A recent study showed a link between timescale-dependent and region-specific modulation of temporal complexity and the affective and cognitive impacts of seizure therapy (Farzan, Atluri et al. 2017). For example, association between change in complexity and improvement in depressive symptoms was localized to the fronto-central and parieto-occipital regions (Farzan, Atluri et al. 2017). In the current study, by observing the temporal stability of brain network dynamics rather than temporal complexity, we provide complimentary evidence suggesting that the therapeutic impact of seizure therapy is network-specific. States C and D were consistently associated with response to seizure therapy and these states were in part localized to the frontal, parietal and occipital regions.

There is strong evidence suggesting that seizure therapy impacts neuronal oscillations. Traditional power spectral density analysis has shown that the slowing of EEG oscillations following ECT is associated with improvement in depressive symptoms (Sackeim, Luber et al. 1996). Power spectral density analysis in our study also revealed increased slow wave activity following ECT. However, this effect was observed in both responders and non-responders. In comparison, microstate analysis illustrated that ECT responders show significant changes in certain microstates (A in duration and B, C, & D in frequency) while non-responders do not. In addition, changes in microstate dynamics were not associated with the increase in power of slow oscillations. Microstate analysis also demonstrated that patients with treatment-resistant depression reveal different global brain dynamics compared to healthy subjects and also that microstate D may play an important role in this difference. While we identified differences in microstate characteristics, no significant differences were observed in power between patients with treatment-resistant depression and healthy subjects. These findings suggest that microstate analysis provide additional and perhaps independent information compared to power spectral density analysis.

There are some limitations to this study. First, treatment-resistant depression may be too broad and heterogeneous to be treated in a homogenous manner. Although patients are grouped together under the definition of treatment-resistant depression (i.e., failed to respond to two or more antidepressants), as seen in this study, they still show heterogeneity in their response to treatments such as seizure therapy. This suggests that there is still considerable heterogeneity in the population of treatment-resistant depression which can translate to variability in the derived neurophysiological markers such as microstate characteristics. Results of this study will therefore need validation with a larger sample size. In addition, due to the small sample size, treatment parameters such as stimulation location (bilateral or unilateral for ECT) and frequency (for MST) could not be controlled. The potential effects these parameters may have on microstate characteristics need to be explored and validated by larger samples in future work. Preregistration of such future studies in advance of data collection and analysis is encouraged.

5.6 Conclusions

The present study provides insight into the mechanism of action of successful seizure therapy for treatment-resistant depression using resting-state EEG microstate analysis. First, an increased duration and decreased frequency of microstates was observed in responders of seizure therapy, specifically ECT. This provides complementary evidence to recent neuroimaging studies which suggest that seizure therapy may stabilize global network dynamics in treatment-resistant depression. In MST, this modulation was a trend-level effect, implying that ECT may have a stronger impact on global neural networks than MST. Second, baseline microstate dynamics were indicative of MST-related improvement in cognition and suicidal ideation. Third, contrary to our hypothesis, we showed reduced global network dynamics in treatment-resistant depression when compared with healthy subjects. We hypothesized that this may be caused by previous antidepressants taken by these patients. Seizure therapy was shown to further reduce these dynamics and this reduction was associated with clinical response. This suggests that antidepressant medications and seizure therapy may have an analogous effect on network dynamics. However, only the modulation of global network dynamics by seizure therapy may be associated with clinical response. Finally, state-specific changes in microstate dynamics were observed in responders of seizure therapy. Microstates previously linked to resting-state networks known to be disrupted in depression (C and D), were associated with seizure therapy response. Further work is required to evaluate microstates as a therapeutic target in developing novel antidepressant treatments.

5.7 Tables

Table 5.1 - Clinical Data Table

	TRD	*n*	ECT Pre & Post	*n*	MST Pre & Post	*n*
Demographic Characteristics						
Age, years: mean (std)	45.7 (14.4)	75	46.8 (15.8)	22	42.0 (13.4)	24
Sex, M/F	31/44	75	8/14	22	12/12	24
Clinical Characteristics						
Illness Duration: mean (std)	20.3 (13.2)	74	19.0 (12.0)	22	20.3 (13.7)	24
On Medications (Antidepressants or Benzodiazepines) (Yes/No)	60/10	70	20/2	22	20/4	24
Avg. No. of Treatment Sessions: mean (std)	-	-	14.2 (5.2)	22	20.2 (6.19)	24
Site of Treatment [a]	-	-	RUL UB RUL UB then BL BL	14 6 2	DMPFC	24
Stimulation Frequency: [a]	-	-	-	-	100 Hz 60 Hz 50 Hz 25 Hz	12 1 2 9
Clinical Assessments						
HRSD, % change following Seizure Therapy: mean (std)	36.7 (30.0)	46	44.8 (28.6)	22	29.3 (29.9)	24
Initial HRSD scores, mean(std)	26.6 (4.44)	74	24.5 (3.81)	22	28.1 (4.73)	24
Post HRSD scores, mean(std)	-	-	13.0 (6.19)	22	19.3 (7.96)	24
HRSD, Responders/Nonresponders [b]	20/26	46	13/9	22	7/17	24
MoCA, change following seizure therapy: mean (std)	0.56 (3.32)	46	-3 (2.97)	6	1.57 (2.69)	21
MoCA: # patients with improvement > 0	12	27	0	6	12	21
BDI, % change: mean (std)	-	-	49.2 (31.2)	17	-	-
BDI, Responders/Nonresponders [b]	-	-	9/8	17	-	-
MADRS, % change: mean (std)	-	-	50.0 (32.6)	18	-	-
MADRS, Responders, Nonresponders [b]	-	-	11/7	18	-	-
SSI, change: mean (std)	-	-	-	-	6.4 (6.4)	20
SSI, Responders/Nonresponders [b]	-	-	-	-	16/4	20

ECT: electroconvulsive therapy; **MST:** magnetic seizure therapy

[a] **RUL UB**: Right Unilateral Ultra-Brief Pulse Width; **BL:** Bitemporal (brief pulse width); **DMPFC:** Dorsomedial Prefrontal Cortex

[b] **Response for HRSD/BDI/MADRS/SSI** defined as >=50% improvement in score

5.8 Figures

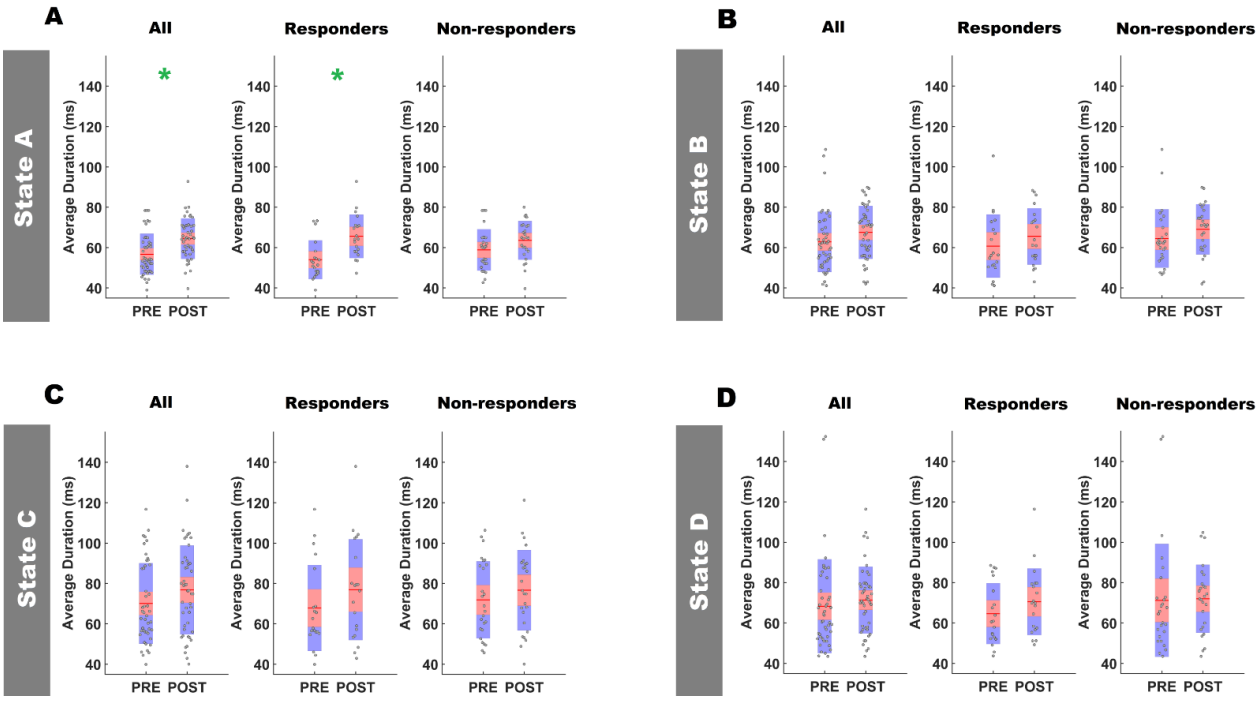


Figure 5.1 - Effect of seizure therapy (ECT and MST) on the average duration of all four microstates.
In each subplot, the raw data is plotted on top of a boxplot showing the mean (red line), 95% confidence interval (red area) and 1 standard deviation (blue area). Significant comparisons are marked with a green (*). **(A)** Left panel: Following seizure therapy (ECT and MST), there was a significant increase in the duration of State A (y-axis) (p<0.0001). Middle panel: This increase was specific to responders of seizure therapy (p<0.0001). **(B)** No significant changes were observed in the duration of State B. **(C)** No significant changes were observed in the duration of State C. **(D)** No significant changes were observed in the duration of State D.

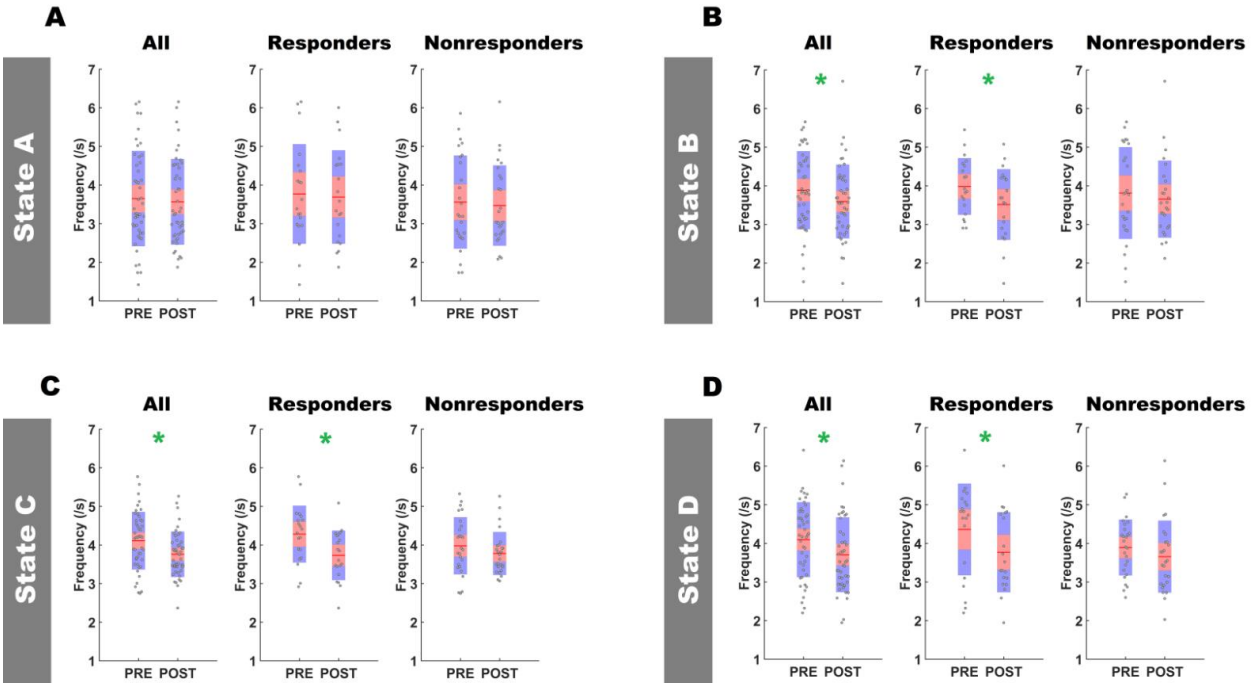


Figure 5.2 - Effect of seizure therapy (ECT and MST) on the frequency of all four microstates.
In each subplot, the raw data is plotted on top of a boxplot showing the mean (red line), 95% confidence interval (red area) and 1 standard deviation (blue area). Significant comparisons are marked with a green (*). **(A)** No significant changes were observed in the frequency of State A. **(B) Left panel:** A decrease in the frequency of State B (y-axis) was observed following seizure therapy (p=0.03). **Middle panel:** This decrease in frequency of State B was specific to responders of seizure therapy (p=0.01). **(C) Left panel**: A decrease in the frequency of State C (y-axis) was observed following seizure therapy (p=0.004). **Middle panel:** This decrease in frequency of State C was specific to responders of seizure therapy (p=0.0004). **(D) Left panel:** A decrease in the frequency of State D (y-axis) was observed following seizure therapy (p=0.0008). **Middle panel:** This decrease in frequency of State D was specific to responders of seizure therapy (p=0.04).

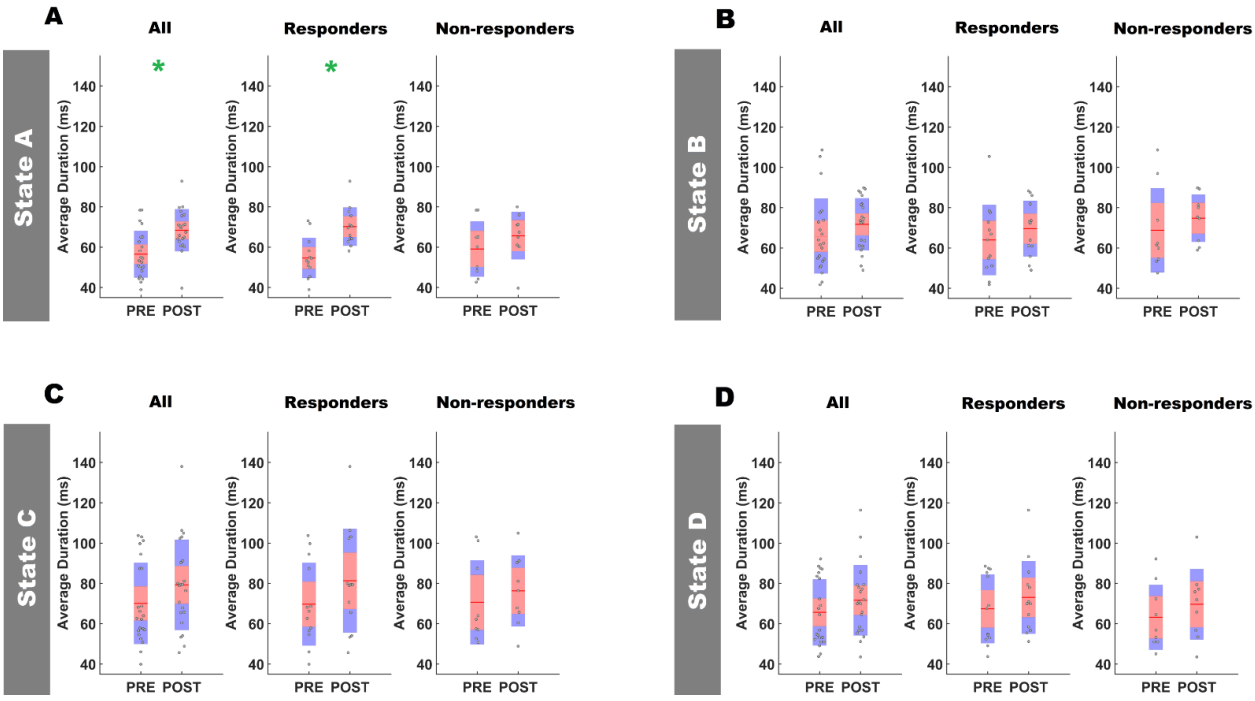


Figure 5.3 - Effect of electroconvulsive therapy (ECT) on the average duration of all four microstates.

In each subplot, the raw data is plotted on top of a boxplot showing the mean (red line), 95% confidence interval (red area) and 1 standard deviation (blue area). Significant comparisons are marked with a green (*). **(A) Left panel:** Following ECT, there was a significant increase in the duration of State A (y-axis) (p<0.0001). **Middle panel:** This increase was specific to responders of ECT (p<0.0001). **(B)** No significant changes were observed in the duration of State B. **(C)** No significant changes were observed in the duration of State C. **(D)** No significant changes were observed in the duration of State D.

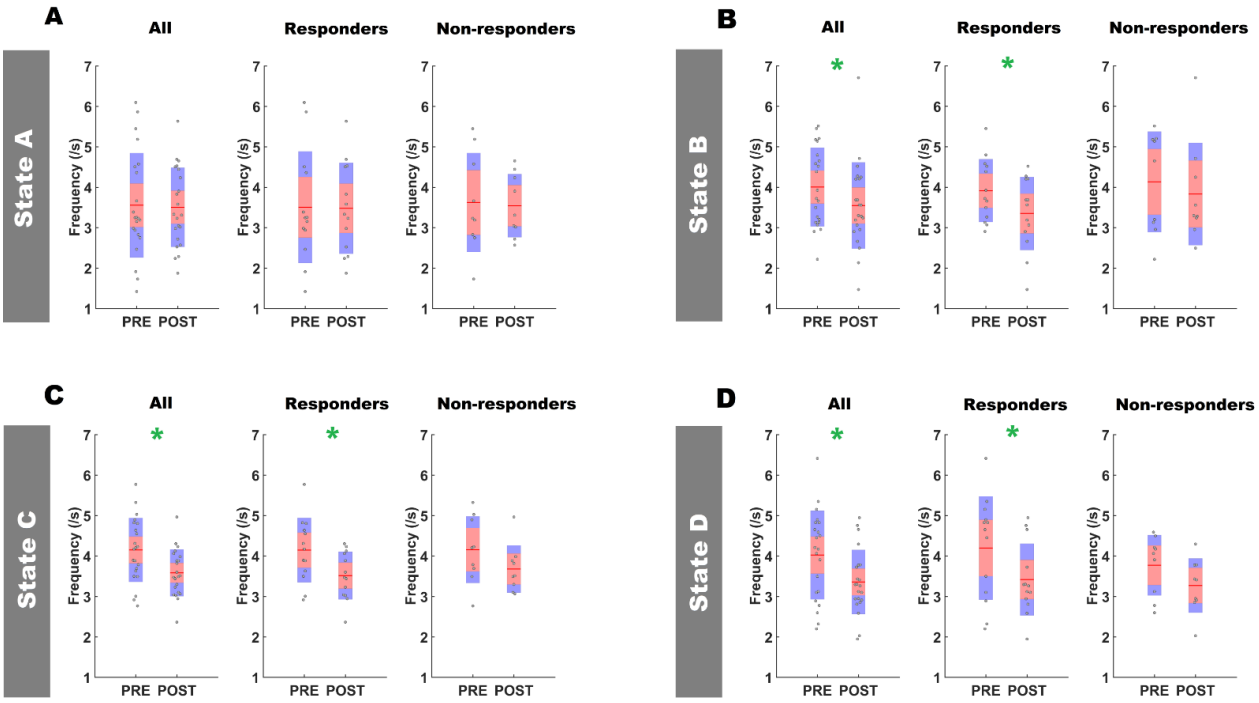


Figure 5.4 - Effect of electroconvulsive therapy (ECT) on the frequency of all four microstates.
In each subplot, the raw data is plotted on top of a boxplot showing the mean (red line), 95% confidence interval (red area) and 1 standard deviation (blue area). Significant comparisons are marked with a green (*). **(A)** No significant changes were observed in the frequency of State A. **(B) Left panel:** A decrease in the frequency of State B (y-axis) was observed following ECT (p=0.03). **Middle panel:** This decrease in frequency of State B was specific to responders of ECT (p=0.03). **(C) Left panel:** A decrease in the frequency of State C (y-axis) was observed following ECT (p=0.002). **Middle panel:** This decrease in frequency of State C was specific to responders of ECT (p=0.008). **(D) Left panel:** A decrease in the frequency of State D (y-axis) was observed following ECT (p=0.0003). **Middle panel:** This decrease in frequency of State D was specific to responders of ECT (p=0.04).

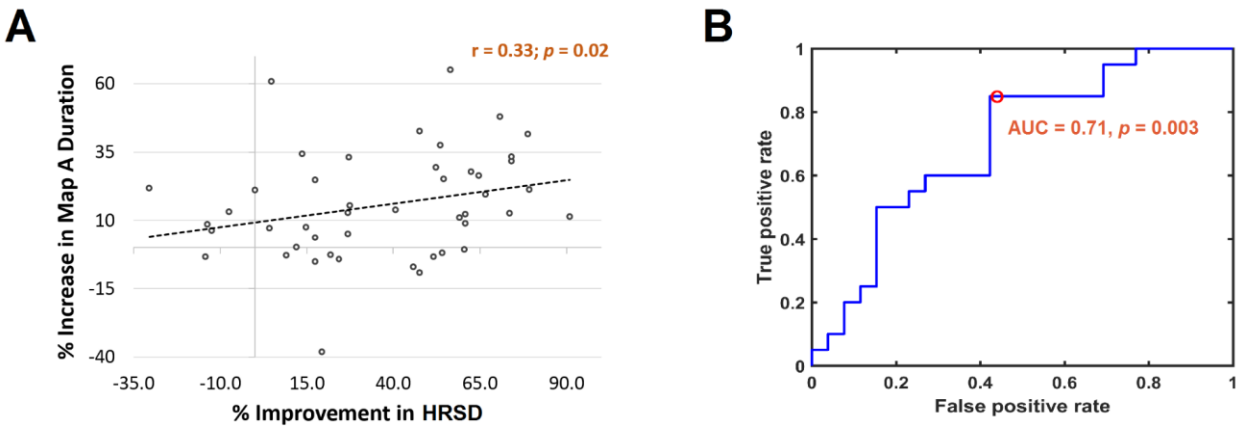


Figure 5.5 - Change in the duration of State A following seizure therapy (ECT+MST) correlated with improvement in depressive symptoms (HRSD).
For the receiver operating characteristic (ROC) curve (right panels), the x-axes represents the false positive rate (1-specificity) and the y-axes represents the true positive rate (sensitivity). The red circle depicts the optimum operating point of the ROC curve. The area under the curve (AUC) at this optimum point is specified on the graph. **(A)** An increase in the duration of State A significantly correlated with improvement in depressive symptoms (x-axis), Hamilton Rating Scale for Depression (HRSD) (r=0.33, *p*=0.02). X-axis represents change in HRSD (pre-post)/pre*100). Y-axis represents change in coverage of State D (post-pre)/pre*100). **(B)** Change in State A duration was also a fair predictor of response to seizure therapy (AUC=0.71, *p*=0.003).

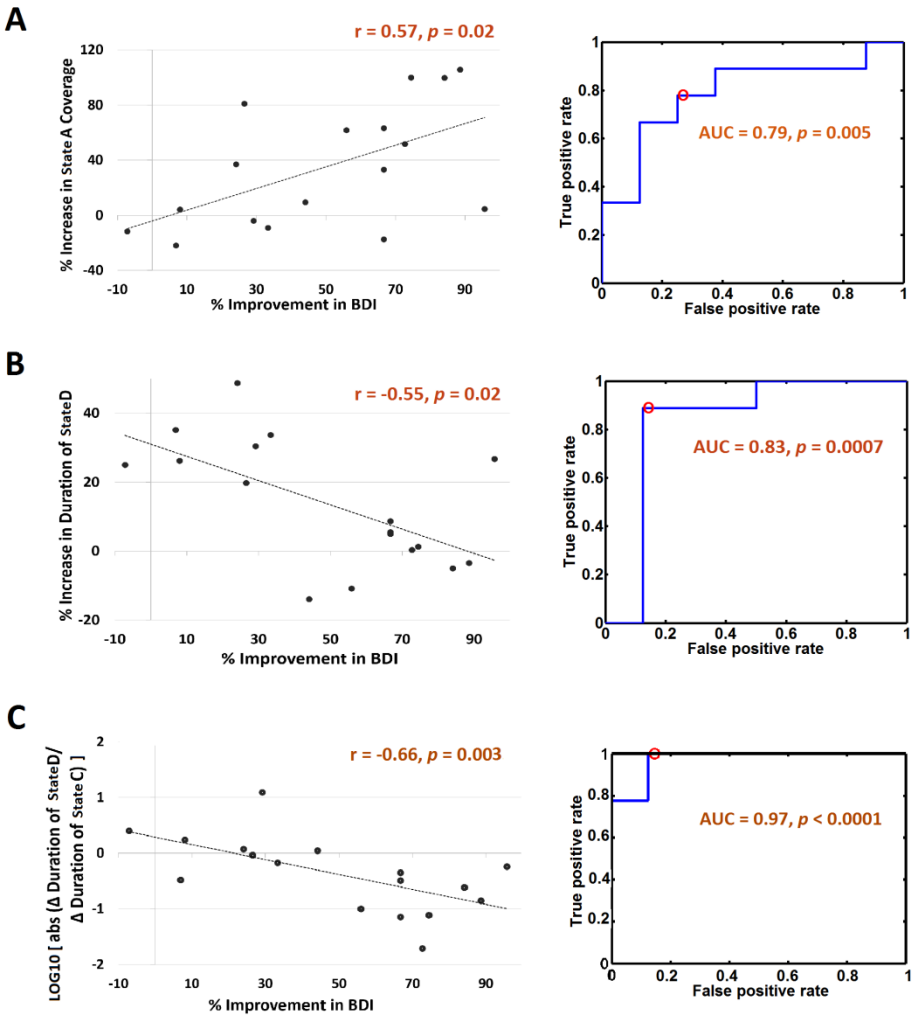


Figure 5.6 - Changes in microstate characteristics following electroconvulsive therapy (ECT) correlated with improvement in self-rated depressive symptoms.
For the receiver operating characteristic (ROC) curves (all right panels), the x-axes represents the false positive rate (1-specificity) and the y-axes represents the true positive rate (sensitivity). The red circle depicts the optimum operating point of the ROC curve. The area under the curve (AUC) at this optimum point is specified on the graph. **(A) Left panel:** An increase in the coverage of State A significantly correlated with improvement in self-rated depressive symptoms (x-axis), Beck's Depression Inventory scale (BDI) (r=0.57, *p*=0.02). X-axis represents change in BDI (pre-post)/pre*100). Y-axis represents change in coverage of State D (post-pre)/pre*100). **Right panel:** Change in State A coverage was also a strong predictor of response to ECT (BDI) (AUC=0.79, *p*=0.005). **(B) Left panel:** A decrease in the duration of State D was significantly correlated with improvement in BDI (r= −0.55, *p*=0.02). X-axis represents change in BDI (pre-post)/pre*100). Y-axis represents change in duration of State D (post-pre)/pre*100). **Right panel:** Change in State D duration was also a strong predictor of response to ECT (BDI) (AUC=0.83, *p*=0.0007). **(C) Left panel:** The correlation between State D duration and BDI remained significant when the change in duration of State D was presented relative to the change

in duration of State C (r= −0.66, *p*=0.003). X-axis represents change in BDI (pre-post)/pre*100). Y-axis represents the log of the absolute ratio between change in duration of State D (post-pre)/pre*100) over the change in duration of State C (post-pre)/pre*100). **Right panel:** Ratio of change in State D duration over the change in State C duration was an excellent predictor of self-rated response to ECT (BDI) (AUC=0.97, *p*<0.0001).

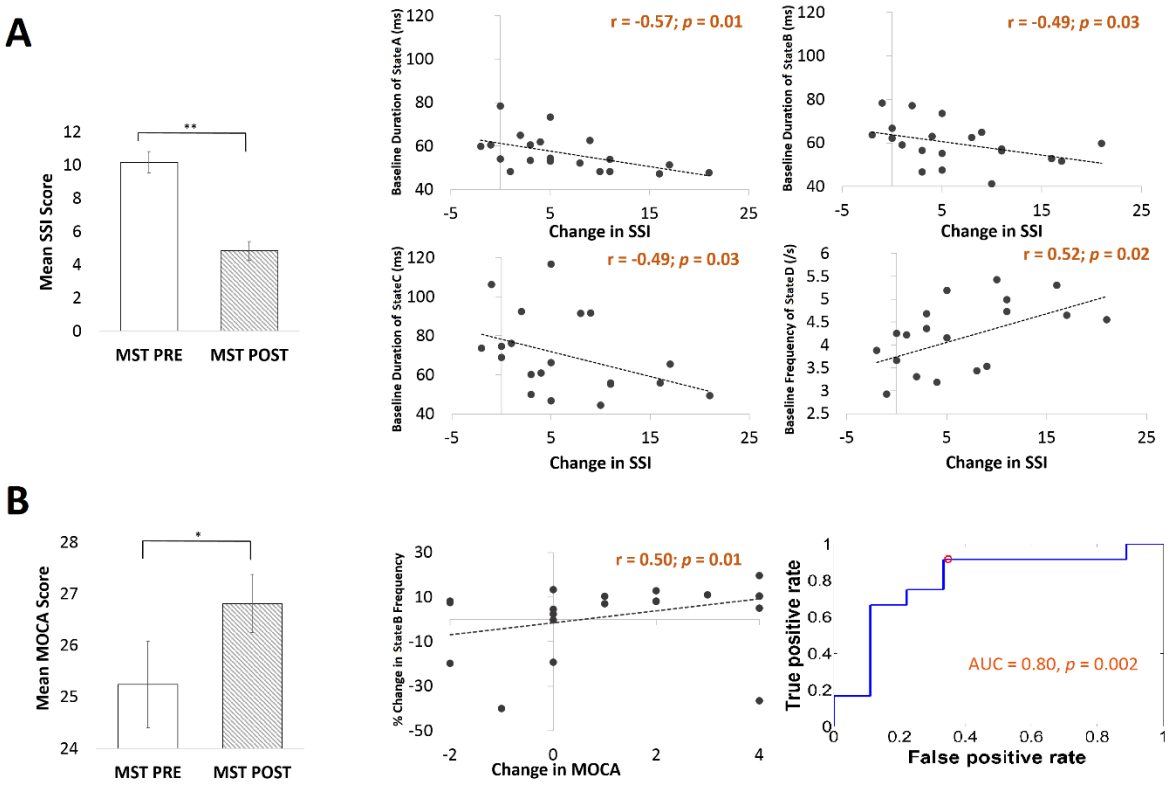


Figure 5.7 - Effect of magnetic seizure therapy (MST) on microstate characteristics.
(A) Left panel: Following MST, patients showed a significant improvement in scale for suicidal ideation (SSI). **Middle and right panels:** Reduction in SSI was significantly associated with baseline resting-state microstate characteristics of all microstate classes. X-axes represents change in SSI score (Post-Pre) and y-axes represents baseline characteristics of each microstate A, B, C and D**. (B) Left panel:** Following MST, patients showed a significant improvement in cognition scores (Montreal Cognitive Assessment (MoCA)). **Middle panel:** A decrease in the frequency of State B following MST correlated with improvement in cognition scores. X-axis represents change in MoCA (Post-Pre) and y-axis represent change in State B frequency ((post-pre)/pre*100). **Right panel:** This decrease in frequency was also a strong predictor of cognitive score outcome. The x-axis represents the false positive rate (1-specificity) and the y-axis represents the true positive rate (sensitivity). The red circle depicts the optimum operating point of the receiver operating characteristic curve. The area under curve (AUC) at this optimum point is specified on the graph (AUC=0.80, *p*=0.002).

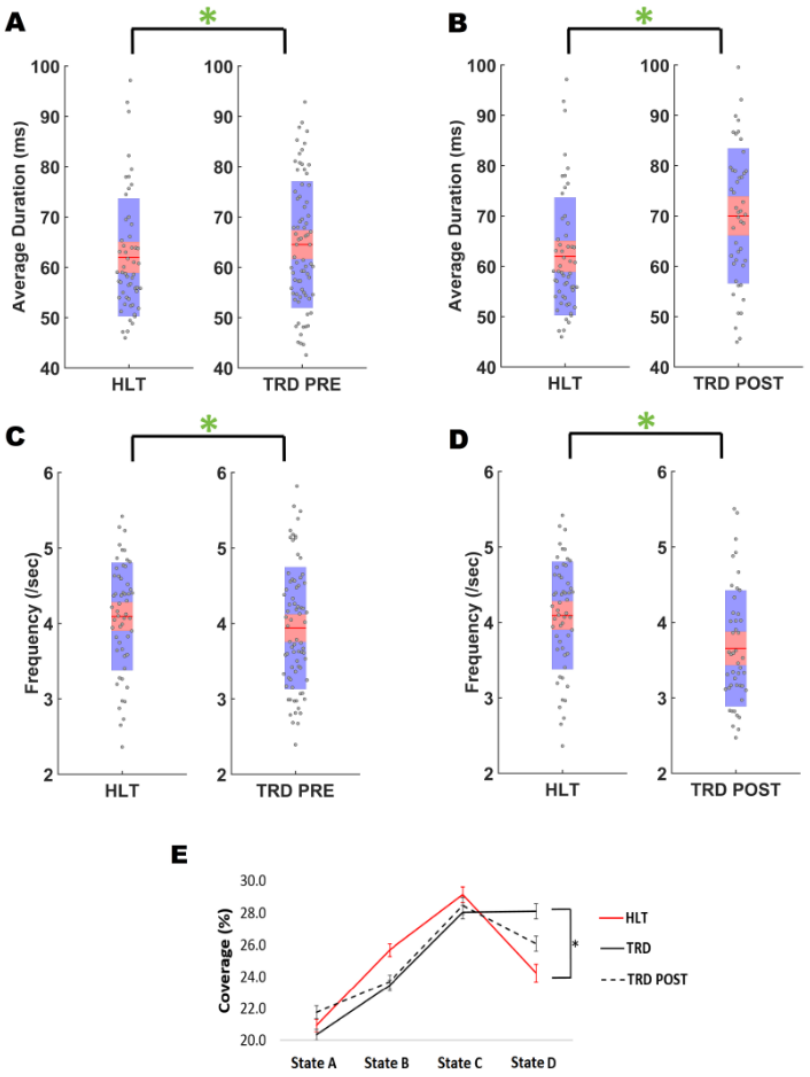


Figure 5.8 - Microstate characteristics of treatment-resistant depression compared to healthy (HLT) subjects before and after seizure therapy.
In each subplot, the raw data is plotted on top of a boxplot showing the mean (red line), 95% confidence interval (red area) and 1 standard deviation (blue area). All comparisons shown in **(A) to (D)** were significant. **(A) & (B)** Patients showed a longer duration (p=0.03) and lower frequency (p=0.03) of microstate dynamics than healthy subjects. **(C) & (D)** Following seizure therapy, patients showed a much longer duration ($p < 0.0001$) and lower frequency ($p = 0.0001$) of microstates than healthy subjects. In all plots of (**A-D**, x-axes represents the subject group. In **(A-B)**, y-axis represents the duration of all microstates in milliseconds (main effect of group in ANCOVA). In **(C-D)**, y-axis represents the frequency of all microstates per second (main effect of group in ANCOVA). **(E)** Seizure therapy (ECT and MST) was shown to normalize the high coverage of State D in patients compared to healthy subjects (p=0.01). X-axis represents each microstate class. Y-axis represents the percent coverage of all microstates, and each line in the graph represents a subject group (interaction effect of M*icrostate Class x Group* in ANCOVA).

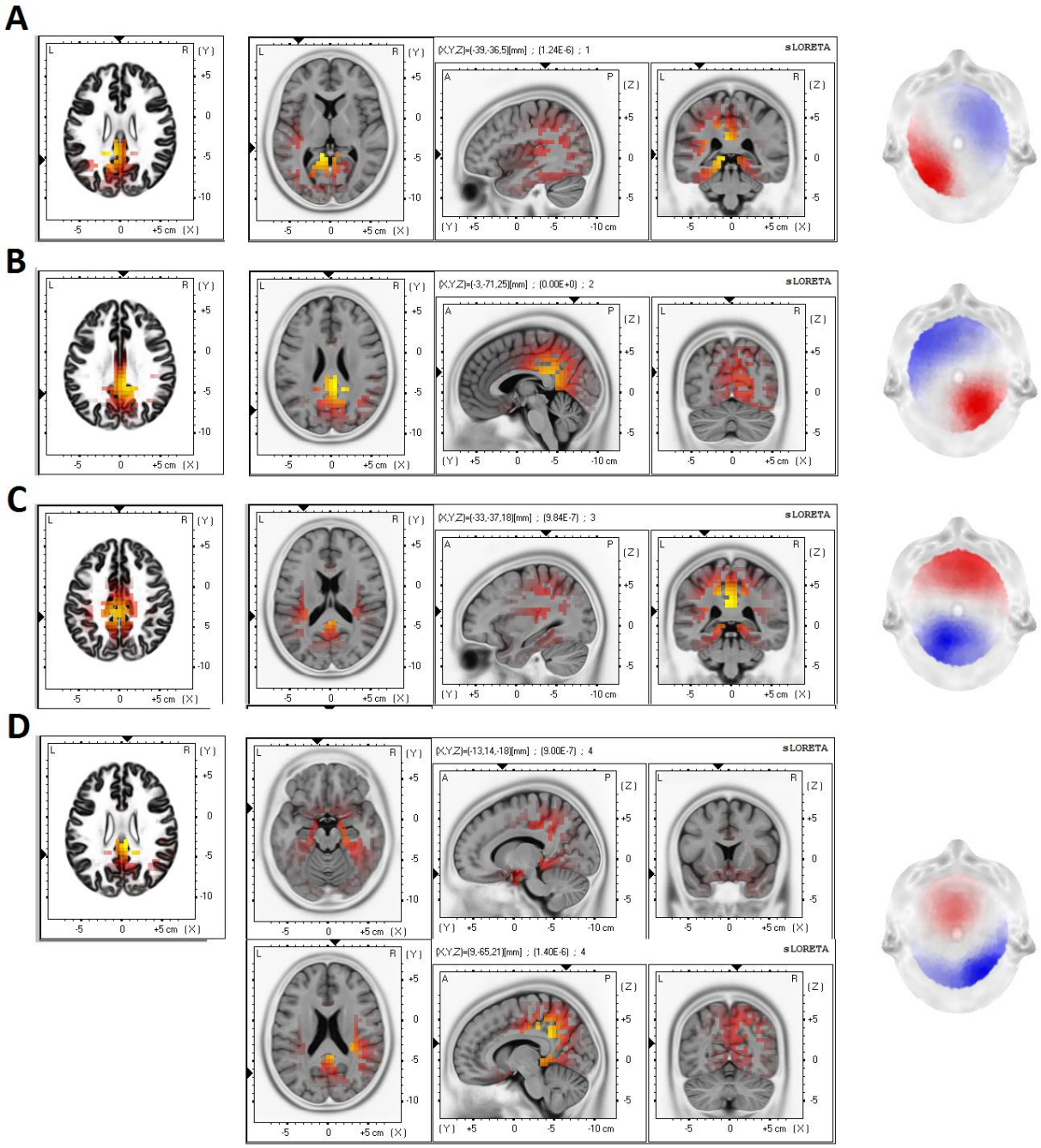


Figure 5.9 - Global microstate classes clustered over all groups and all subjects with their source (eLORETA) images.

All microstates show activation in the posterior cingulate gyrus **(A)** Microstate A was shown to be associated with the left superior and middle temporal gyrus. **(B)** Microstate B was associated with the cuneus and precuneus of the occipital lobe. **(C)** Microstate C was associated with the anterior cingulate, insula and cuneus and precuneus of the occipital lobe. **(D)** Microstate D was associated with the paracentral lobe of the frontal lobe, the precuneus of the parietal lobe, the parahippocampal gyrus, and the lingual gyrus of the occipital lobe.

5.9 Supplementary Material

Table 5.2 - Planned Comparison Results for MST Group

Average Duration				
	State A	**State B**	**State C**	**State D**
Responders	t=2.00; df=7; *p*=0.09; Cohen's *d*=0.51	t=1.14; df=7; *p*=0.30; Cohen's *d*=0.31	t=0.80; df=7; *p*=0.45; Cohen's *d*=0.19	t=1.64; df=7; *p*=0.15; Cohen's *d*=0.55
Non-responders	t=1.39; df=17; *p*=0.18; Cohen's *d*=0.45	t=1.44; df=17; *p*=0.17; Cohen's *d*=0.35	t=1.16; df=17; *p*=0.26; Cohen's *d*=0.22	t=0.48; df=17; *p*=0.64; Cohen's *d*=0.09
Frequency				
	State A	**State B**	**State C**	**State D**
Responders	t=-0.56; df=7; *p*=0.60; Cohen's *d*=0.15	t=-1.26; df=7; *p*=0.25; Cohen's *d*=0.35	t=-2.97; df=7; *p*=0.02; Cohen's *d*=0.70	t=-0.79; df=7; *p*=0.46; Cohen's *d*=0.24
Non-responders	t=-0.62; df=17; *p*=0.54; Cohen's *d*=0.08	t=-0.70; df=17; *p*=0.49; Cohen's *d*=0.08	t=-0.29; df=17; *p*=0.78; Cohen's *d*=0.08	t=-0.52; df=17; *p*=0.61; Cohen's *d*=0.11
Coverage				
	State A	**State B**	**State C**	**State D**
Responders	t=0.32; df=7; *p*=0.76; Cohen's *d*=0.11	t=-0.35; df=7; *p*=0.74; Cohen's *d*=0.13	t=-0.25; df=7; *p*=0.81; Cohen's *d*=0.07	t=0.19; df=7; *p*=0.86; Cohen's *d*=0.07
Non-responders	t=0.24; df=17; *p*=0.81; Cohen's *d*=0.06	t=0.54; df=17; *p*=0.60; Cohen's *d*=0.12	t=0.45; df=17; *p*=0.66; Cohen's *d*=0.14	t=-0.77; df=17; *p*=0.45; Cohen's *d*=0.21

p-values are un-corrected; p-values are not significant after correction for 4 comparisons (4 states)

Table 5.3 - Correlation Results for MST Group

	State A	**State B**	**State C**	**State D**
Duration	r = -0.20; *p*=0.35	r = -0.02; *p*=0.94	r = -0.16; *p*=0.46	r = 0.09; *p*=0.66
Frequency	r = -0.13; *p*=0.56	r = 0.23; *p*=0.28	r = -0.12; *p*=0.59	r = 0.02; *p*=0.93
Coverage	r = -0.11; *p*=0.60	r = 0.28; *p*=0.19	r = -0.16; *p*=0.45	r = 0.03; *p*=0.87

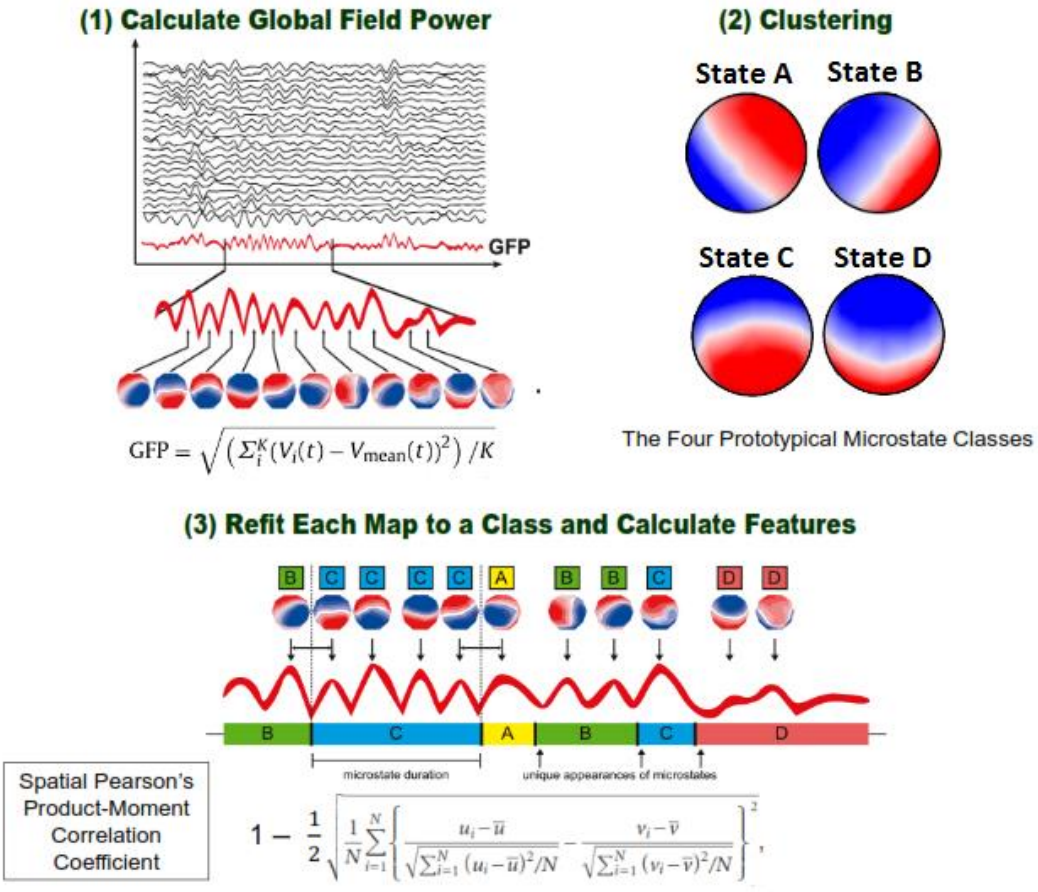


Figure S 5.1 - Overview of Microstate Analysis Procedure (Adapted from *Khanna et al., PloS one, 2014*).

The above figure highlights the three main steps of microstate analysis. **(1)** The global field power over the entire scalp at each time point is calculated. The topographical maps at the peaks of this global field power signal are passed on to the clustering algorithm. **(2)** Four prototypical microstate classes (i.e., states) are defined through the clustering algorithm. **(3)** Each of the topographical maps at the peaks of the GFP signal are correlated with each of the four prototypical microstate classes. The class with the highest correlation is assigned to each peak. Finally, characteristics such as duration, frequency and coverage are defined for each microstate class.

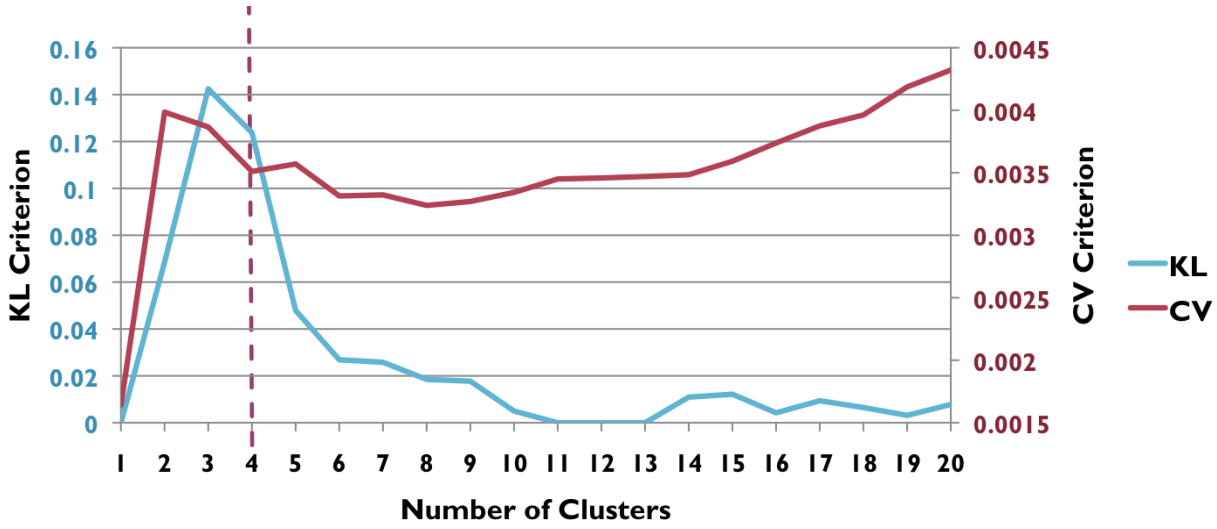


Figure S 5.2 - A data-driven approach to defining the optimum number of clusters for microstate analysis.

The x-axis represents the number of microstate clusters derived from the data (1 to 20). The left y-axis represents the Krzanowski-Lai (KL) criterion, which is based on global explained variance and is a measure of the quality of clustering. KL must be maximized for optimal clustering. The right y-axis represents the cross-validation (CV) criterion, which is based on the predictive residual variance (i.e., mean-squared error) and must be minimized for optimal clustering. The number of clusters that minimized CV and maximized KL, as shown on the figure, was 4.

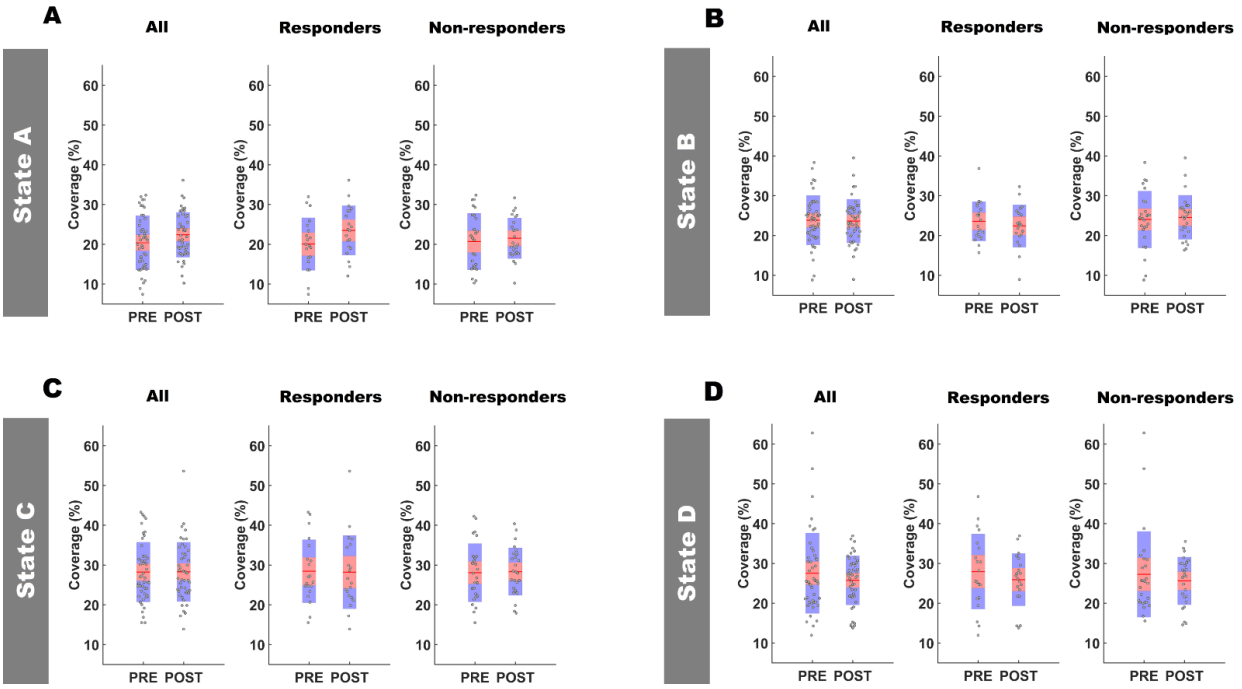


Figure S 5.3 - Effect of seizure therapy (ECT and MST) on the coverage of all four microstates.
In each subplot, the raw data is plotted on top of a boxplot showing the mean (red line), 95% confidence interval (red area) and 1 standard deviation (blue area). Significant comparisons are marked with a green (*). **(A-D)** No significant changes were observed in the duration of State A-D.

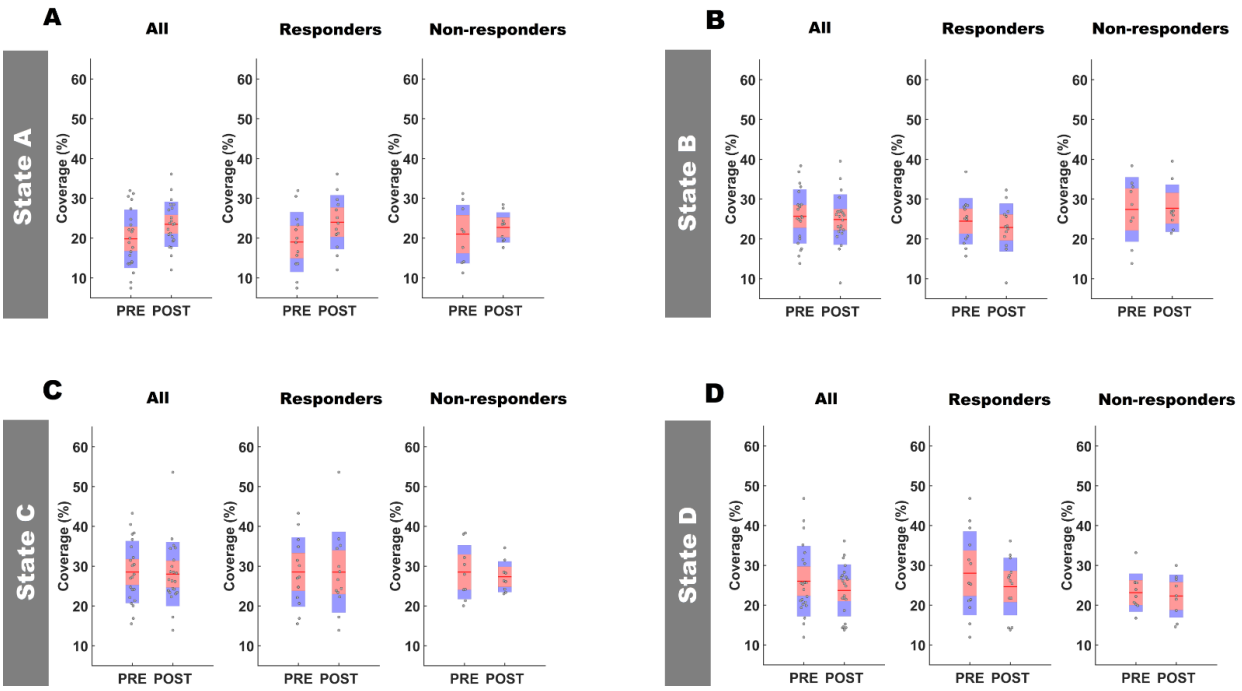


Figure S 5.4 - Effect of electroconvulsive therapy (ECT) on the coverage of all four microstates.
In each subplot, the raw data is plotted on top of a boxplot showing the mean (red line), 95% confidence interval (red area) and 1 standard deviation (blue area). Significant comparisons are marked with a green (*). **(A-D)** No significant changes were observed in the duration of State A-D.

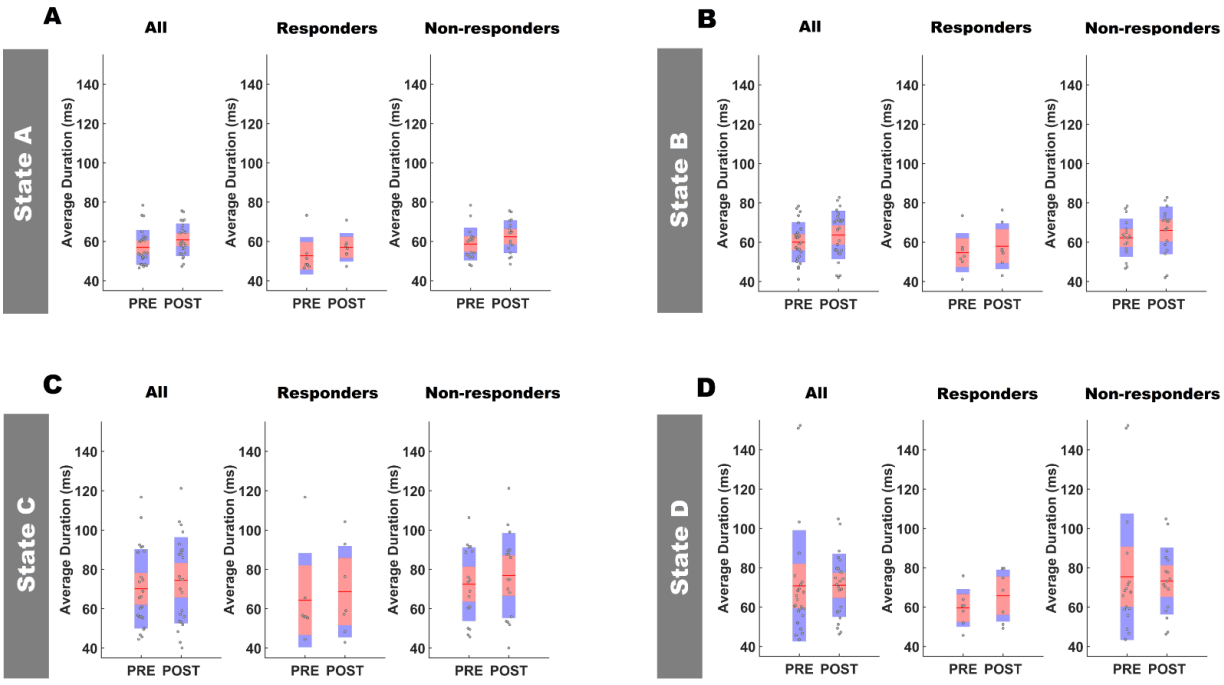


Figure S 5.5 - Effect of magnetic seizure therapy (MST) on the average duration of all four microstates.

In each subplot, the raw data is plotted on top of a boxplot showing the mean (red line), 95% confidence interval (red area) and 1 standard deviation (blue area). Significant comparisons are marked with a green (*). **(A-D)** No significant changes were observed in the duration of State A-D.

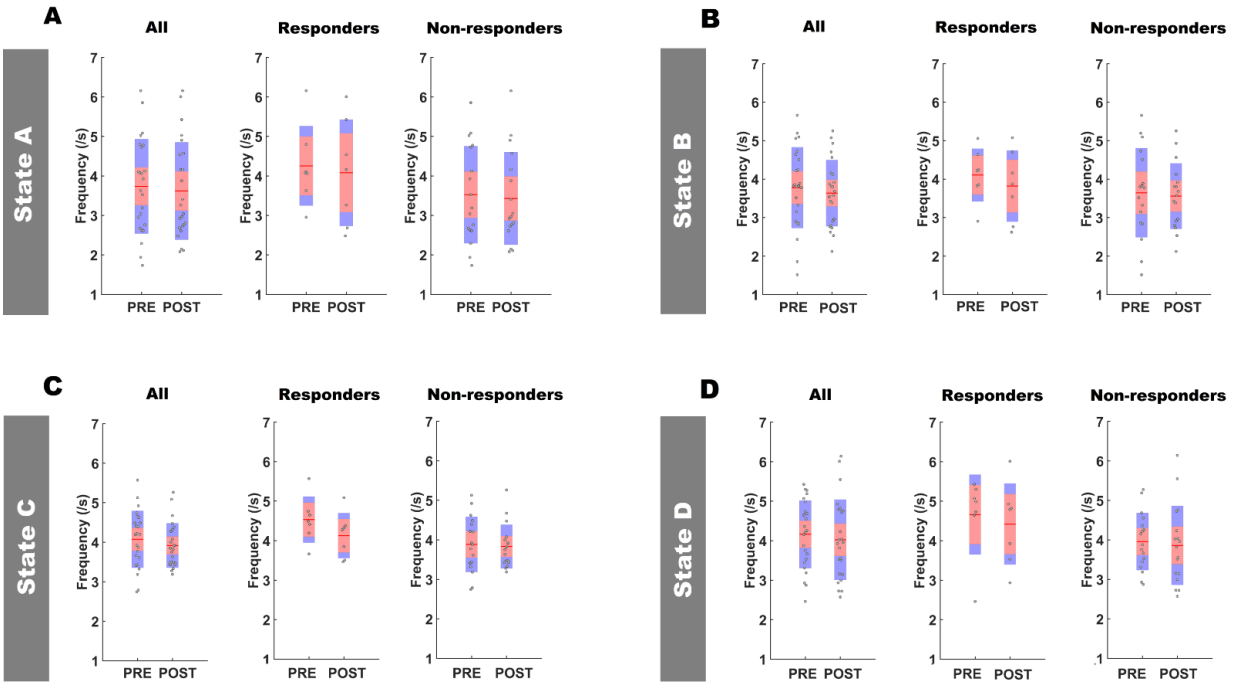


Figure S 5.6 - Effect of magnetic seizure therapy (MST) on the frequency of all four microstates.
In each subplot, the raw data is plotted on top of a boxplot showing the mean (red line), 95% confidence interval (red area) and 1 standard deviation (blue area). Significant comparisons are marked with a green (*). **(A-D)** No significant changes were observed in the duration of State A-D.

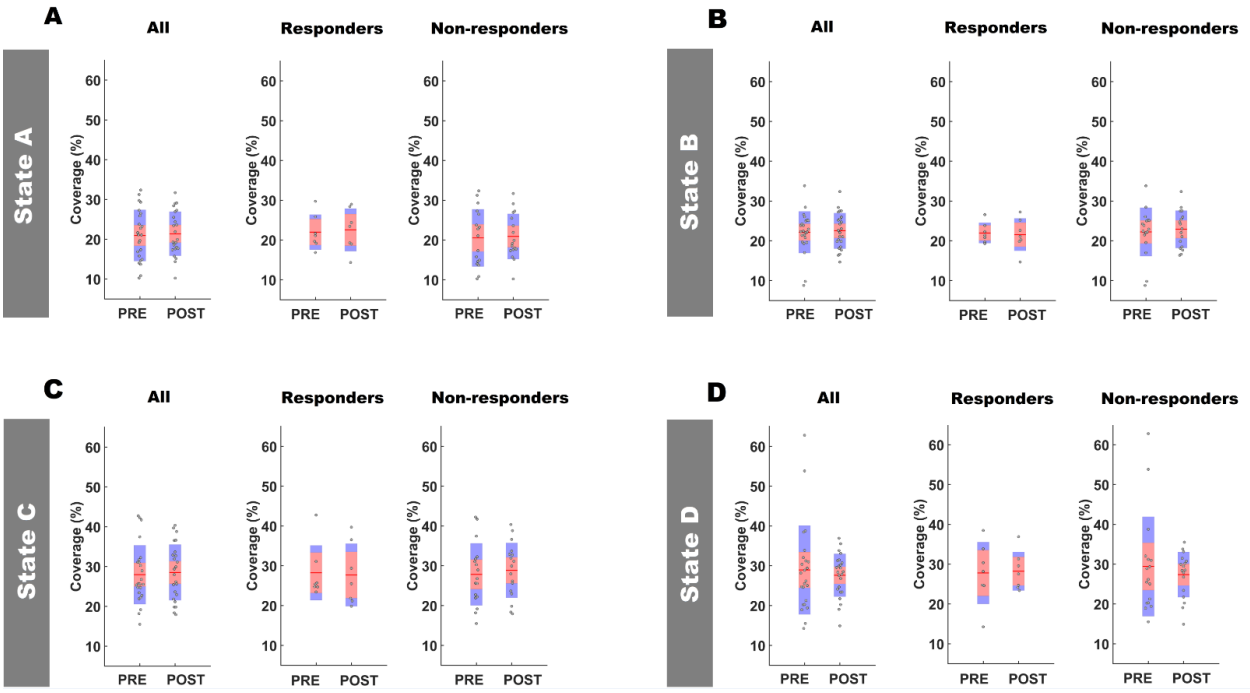


Figure S 5.7 - Effect of magnetic seizure therapy (MST) on the coverage of all four microstates.
In each subplot, the raw data is plotted on top of a boxplot showing the mean (red line), 95% confidence interval (red area) and 1 standard deviation (blue area). Significant comparisons are marked with a green (*). **(A-D)** No significant changes were observed in the duration of State A-D.

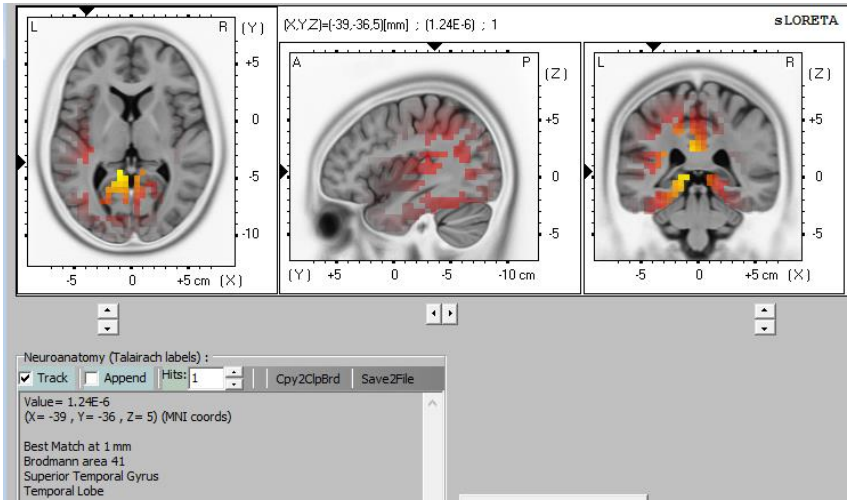


Figure S 5.8 - Global-Clustered "Microstate A" was localized to regions reported in previous work (i.e., superior temporal gyrus).

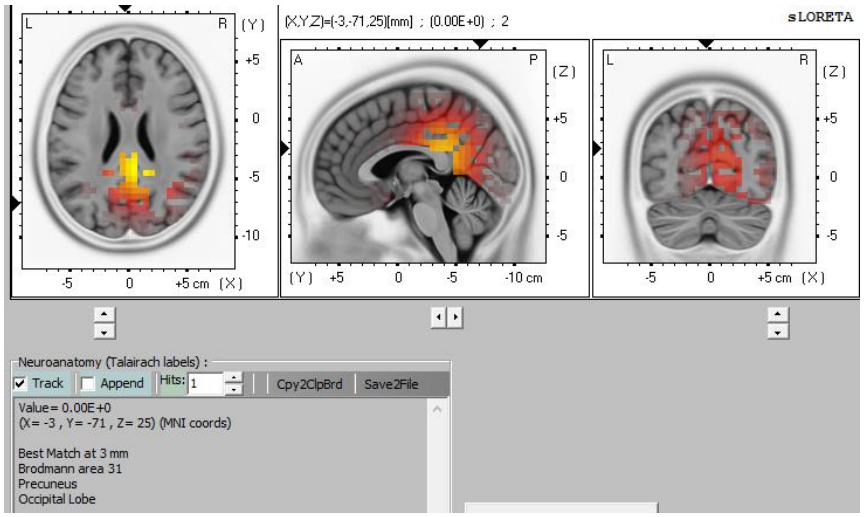


Figure S 5.9 - Global-Clustered "Microstate B" was localized to regions reported in previous work (i.e., occipital lobe).

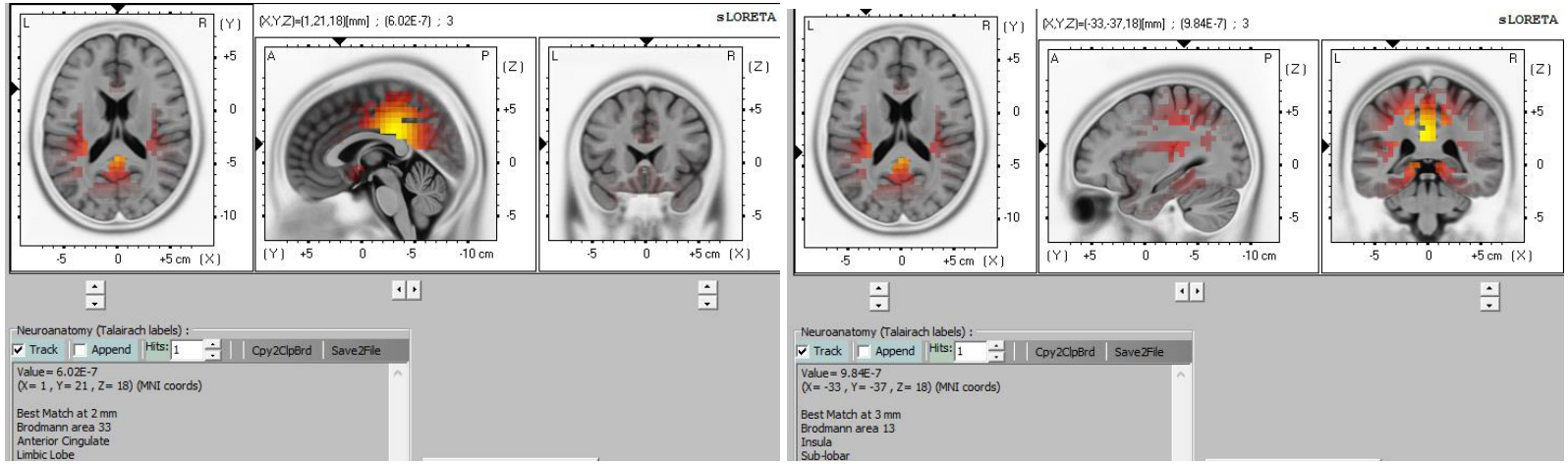


Figure S 5.10 - Global-Clustered "Microstate C" was localized to regions reported in previous work (i.e., anterior cingulate cortex and insula).

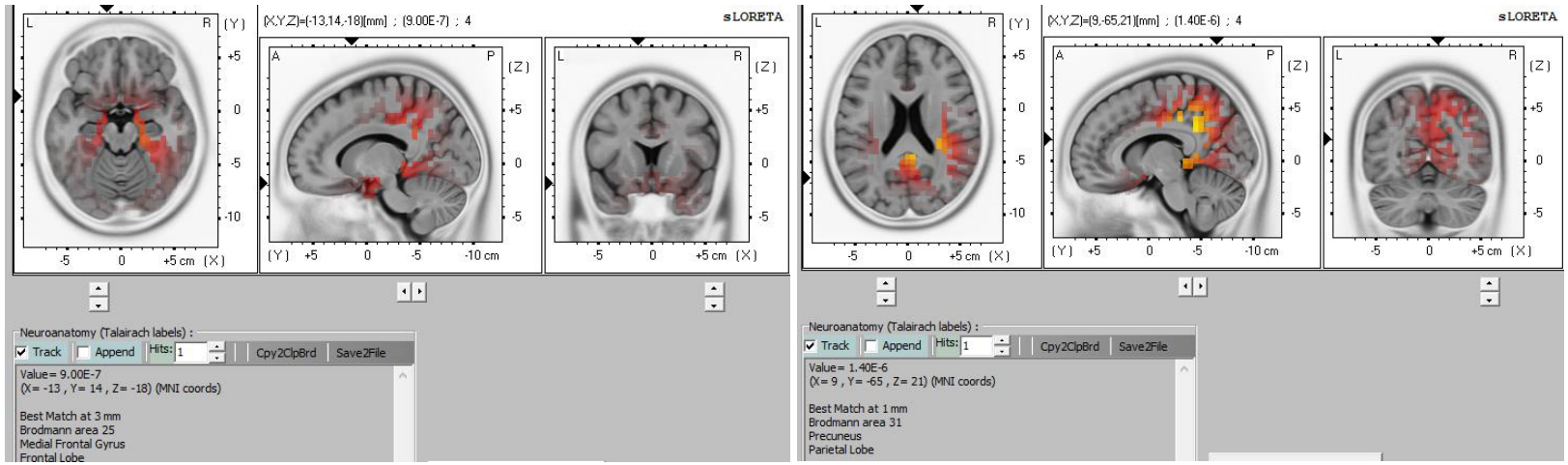


Figure S 5.11 - Global-Clustered "Microstate D" was localized to regions reported in previous work (i.e., frontal and parietal regions).

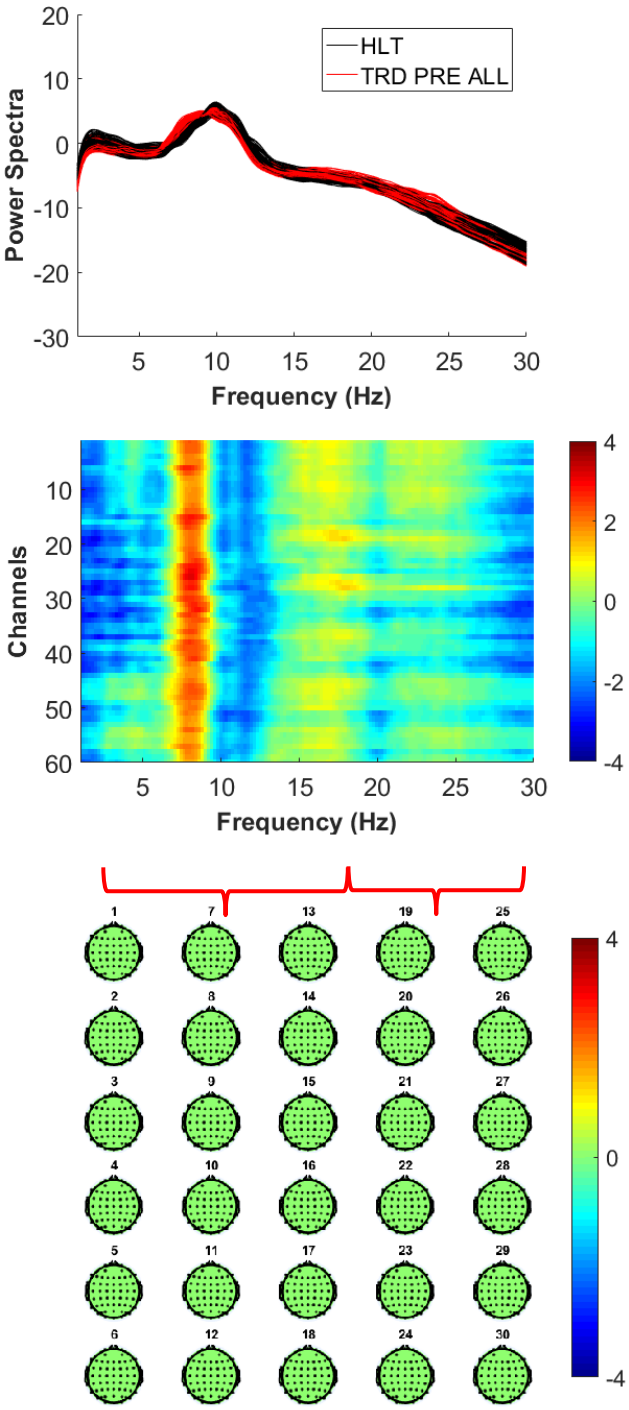


Figure S 5.12 - Power analysis was performed between the healthy (HLT) subjects (n=55) and treatment-resistant depression (TRD) patients (n=75). Top. Relative power spectrum is shown across all 60 electrodes for the healthy and TRD groups (black and red lines, respectively). **Middle.** Parametric unpaired t-test map illustrates the t-statistic comparing relative power between the healthy and TRD groups (blue: reduced power in TRD group compared to healthy subjects; red: increased power). **Bottom.** Topographic plots illustrate significant t-map clusters ($P < 0.05$) following correction for multiple comparisons, using a cluster-based nonparametric permutation test.

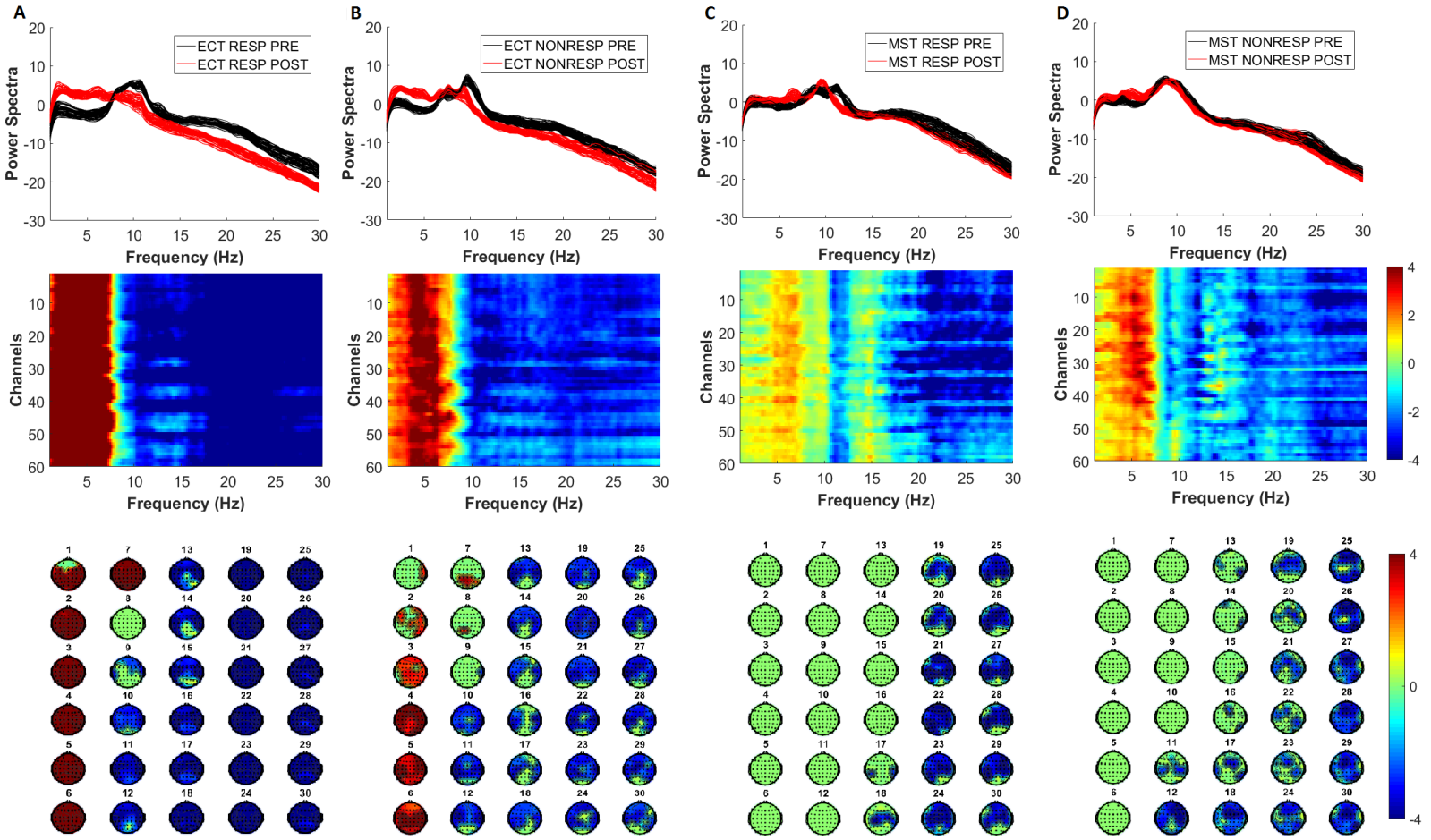


Figure S 5.13 - Power analysis was performed on resting-state EEG collected before and after treatment for responders (n=13) and non-responders (n=9) of electroconvulsive therapy (ECT) (A-B) and responders (n=7) and non-responders (n=17) of magnetic seizure therapy (MST) (C-D). ***Top.*** Relative power spectrum is shown across all 60 electrodes before and after a course of treatment (black and red lines, respectively). ***Middle.*** Parametric paired t-test maps illustrate the t-statistic comparing relative power between the healthy and TRD groups (blue: reduced power in TRD group compared to healthy subjects; red: increased power). ***Bottom:*** Topographic plots illustrate significant t-map clusters ($P < 0.05$) following correction for multiple comparisons using a cluster-based nonparametric permutation test.

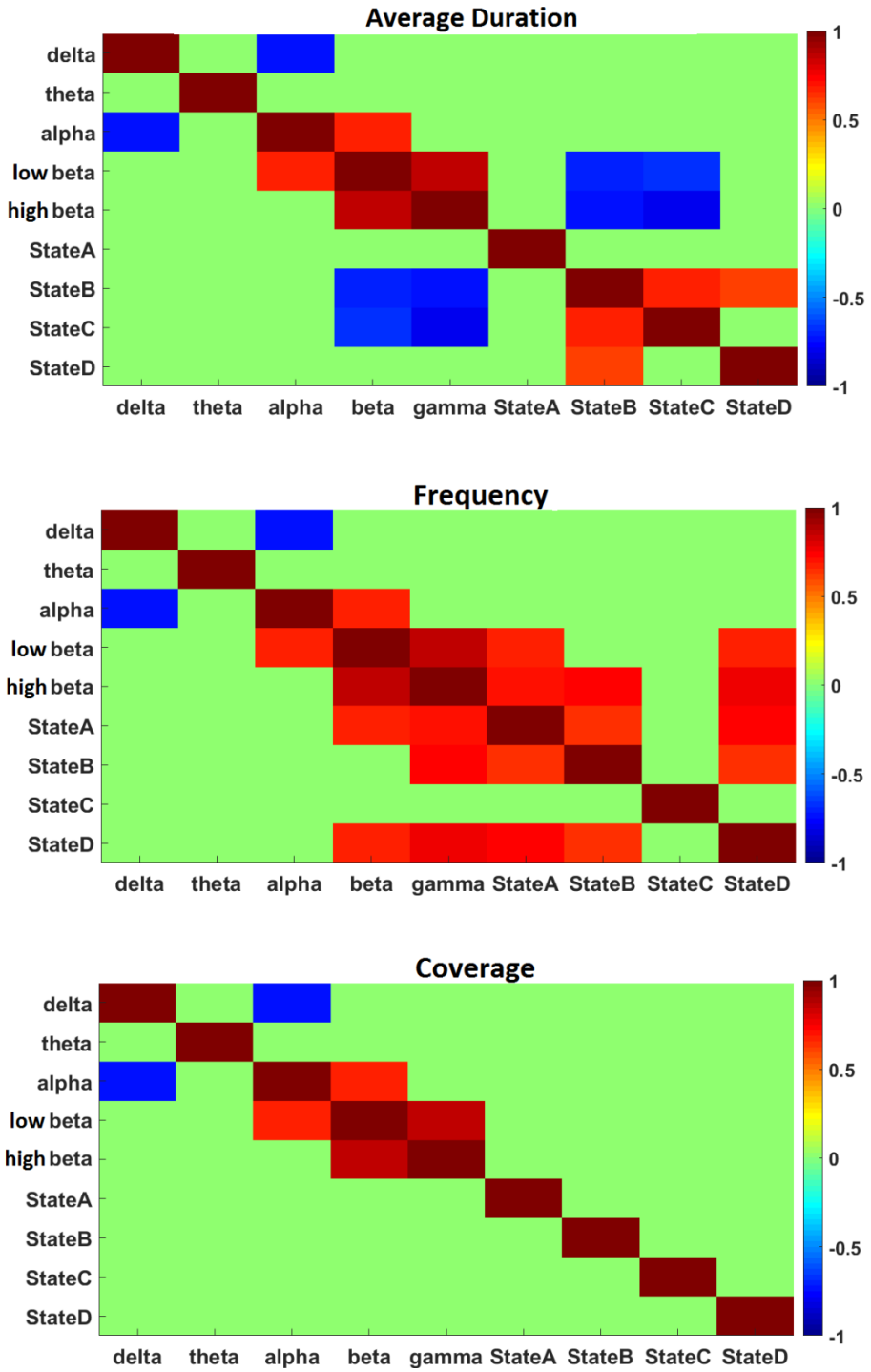


Figure S 5.14 – Correlation between Change in Microstate Characteristics and Change in EEG Power following ECT.

Bonferroni-corrected correlation matrix (Spearman's correlation) (4 maps x3 features = 12 comparisons) between the change in microstate map temporal parameters following Electroconvulsive therapy only (***Top:*** average duration; ***Middle:*** frequency; ***Bottom*****:** coverage) and change in EEG band power following treatment. Power bands were defined as follows: delta – 1:4Hz, theta – 4:8Hz; alpha – 8:12Hz; low beta – 12:15Hz; and high beta 15:30Hz.

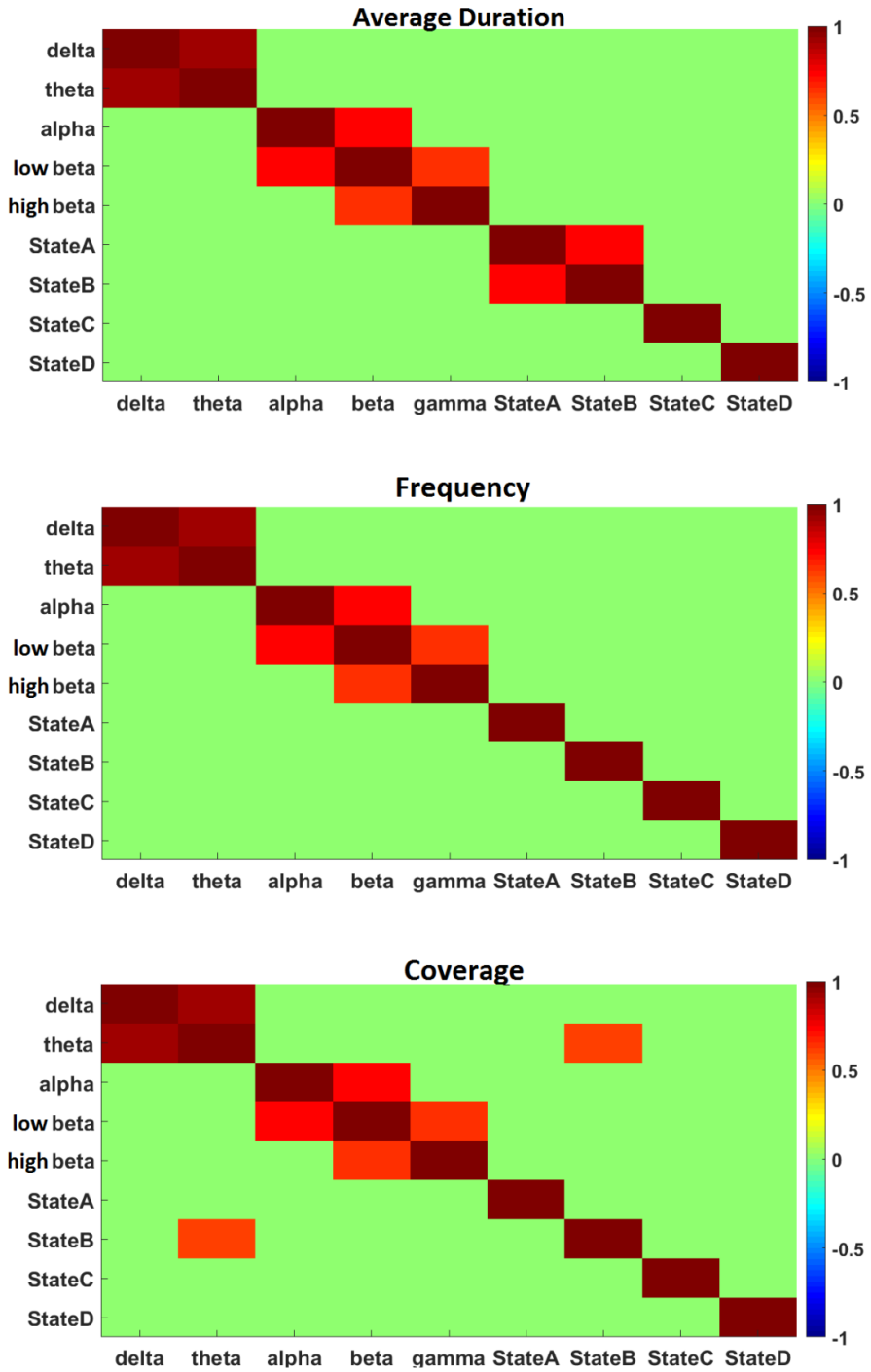


Figure S 5.15 – Correlation between Change in Microstate Characteristics and Change in EEG Power following MST.

Bonferroni-corrected correlation matrix (Spearman's correlation) (4 maps x3 features = 12 comparisons) between the change in microstate map temporal parameters following Magnetic Seizure Therapy only (***Top:*** average duration; ***Middle:*** frequency; ***Bottom*****:** coverage) and change in EEG band power following treatment. Power bands were defined as follows: delta – 1:4Hz, theta – 4:8Hz; alpha – 8:12Hz; low beta – 12:15Hz; and high beta 15:30Hz.

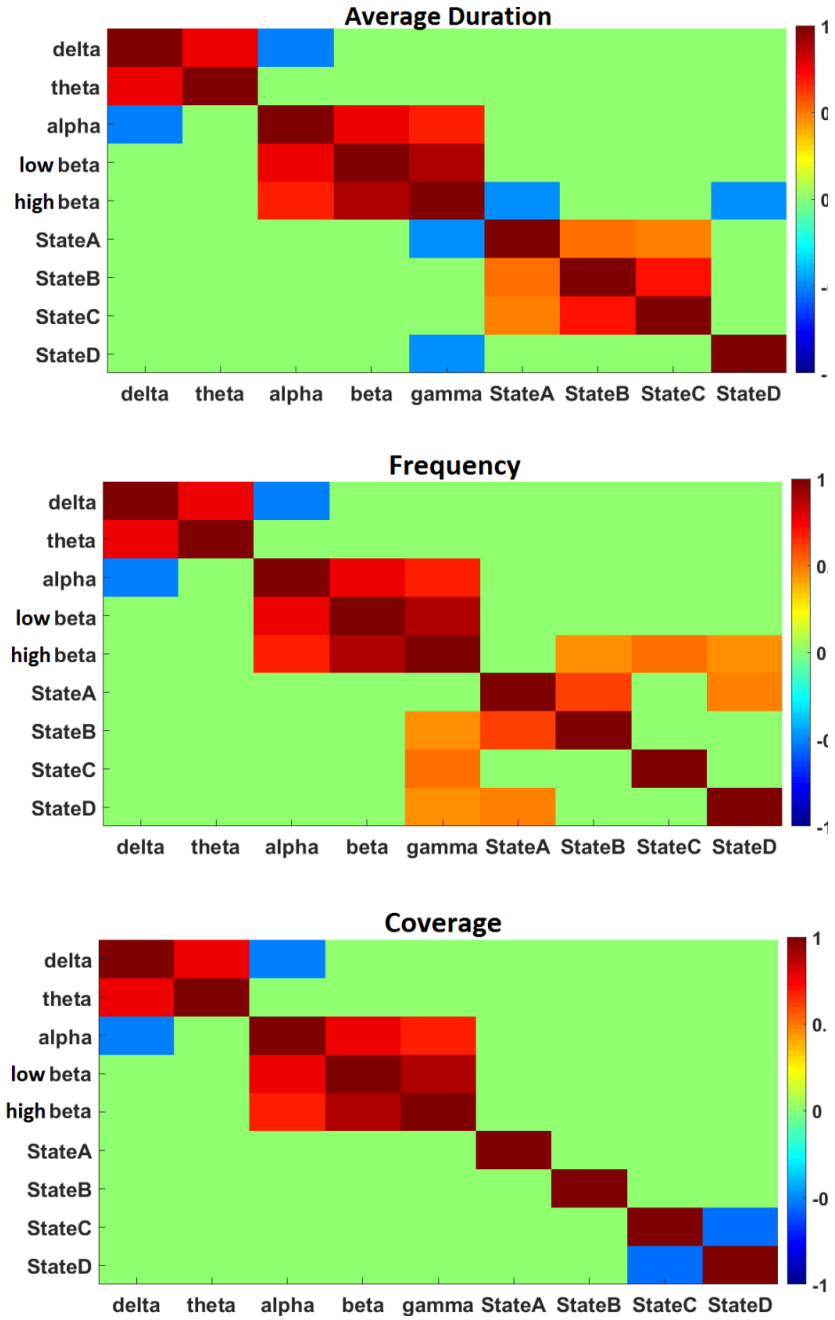


Figure S 5.16 – Correlation between Change in Microstate Characteristics and Change in EEG Power (ECT + MST).

Combining both the Electroconvulsive Therapy (ECT) and the Magnetic Seizure Therapy (MST) groups together (i.e., seizure therapy), we show the Bonferroni-corrected correlation matrix (Spearman's correlation) (4 maps x3 features = 12 comparisons) between the change in microstate map temporal parameters following seizure therapy (T***op:*** average duration; ***Middle:*** frequency; ***Bottom*****:** coverage) and change in EEG band power following treatment. Power bands were defined as follows: delta – 1:4Hz, theta – 4:8Hz; alpha – 8:12Hz; low beta – 12:15Hz; and high beta 15:30Hz.

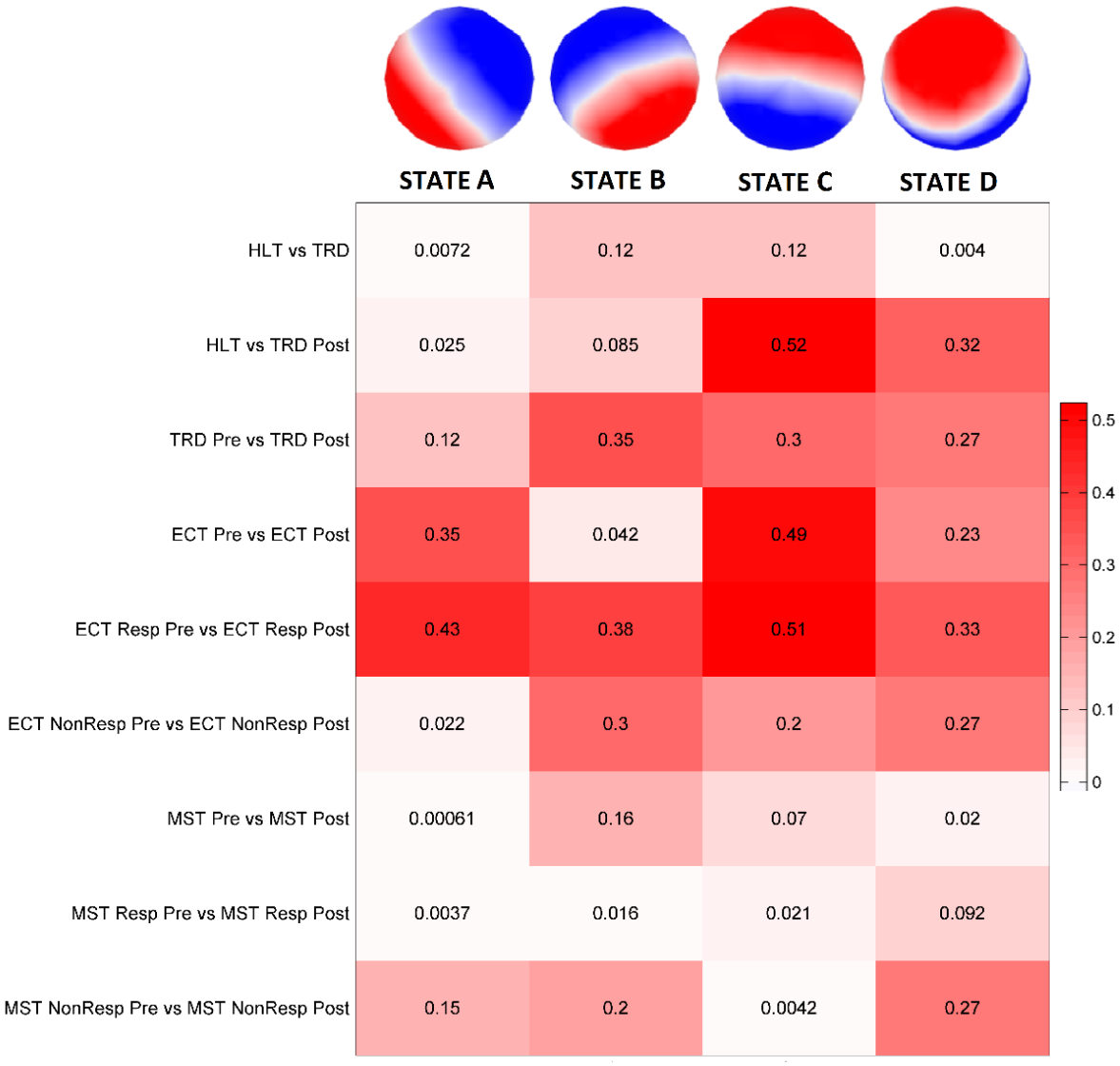


Figure S 5.17 - Global map dissimilarity (GMD) values.
Shown for each comparison between subject groups (i.e., healthy (HLT), treatment-resistant depression (TRD), electroconvulsive therapy (ECT) and magnetic seizure therapy (MST)) or between conditions within a subject group (pre, post, responders, non-responders). Global microstate classes clustered over all subjects are shown at top. Global map dissimilarity quantifies the variation in scalp electrical potential configuration of two microstate topographies from different groups or conditions. It can provide insight on whether different (or similar) neuronal generators are implicated within the two compared groups. GMD values (shown inside the matrix) are modified to range from 0 to 1, where 0 represents identical microstates (regardless of polarity) and 1 represents highly dissimilar microstates. In general, high global map dissimilarity values were observed between the groups/conditions that displayed the most significant differences in microstate characteristics.

Section III: Investigating the Targets of Pharmacotherapy using EEG Measures of Neural Dynamics

Chapter 6 – Characterizing the Modulation of Neural Dynamics during Escitalopram Treatment in Major Depressive Disorder: A CAN-BIND Study

In this chapter, we apply EEG frequency analysis and multiscale entropy analysis to extract power and complexity measures of neural oscillations. We aim to identify whether these measures can provide insight into targets of escitalopram for successful treatment outcome.

Contents of this chapter were prepared for publication. Please note that during the publication process, errors may be discovered which could affect the content.

Sravya Atluri, Willy Wong, Yasaman Vaghei, Andrey Zhdanov, Zafiris Daskalakis, Daniel Blumberger, Benicio N. Frey, Peter Giacobbe, Raymond Lam, Glenda M. MacQueen, Roumen Milev, Daniel J Müller, Sagar V. Parikh, Susan Rotzinger, Claudio N. Soares, Colleen A. Brenner, Fidel Vila-Rodriguez, Steven C. Strother, Mary Pat McAndrews, Killian Kleffner, Esther Alonso, Stephen R. Arnott, Rudolf Uher, Jane A. Foster, Sidney H. Kennedy and Faranak Farzan. "Characterizing the Modulation of neural dynamics during escitalopram treatment in major depressive disorder: A CAN-BIND Study."

6.1 Abstract

A wide range of antidepressant mediations are currently available for the treatment of major depressive disorder. However, they seem to show a delayed onset of action and the average efficacy of antidepressants is low. It is suggested that antidepressants have a delayed onset of action because adaptive changes in brain regions linked to mood and emotion need to occur for therapeutic effect. We hypothesized that such adaptive changes can be characterized by non-invasive measures of neural dynamics. To test this hypothesis, we used resting-state, eyes-closed EEG data collected from 107 patients (age = 36.3+/-12.5; 70 females) through the Canadian Biomarker Integration Network in Depression (CANBIND-1) study. Patients received 8 weeks of escitalopram treatment and EEG assessments were performed at baseline, week 2 and week 8. Power spectral and multiscale entropy analyses were performed. Results from this study suggest that escitalopram potentially modulates neural activity in a frequency- and spatio-temporal specific manner for therapeutic effect. A decrease in alpha and beta oscillations (baseline to week 8) and an increase in complexity in mid-coarse timescales (week 2 to week 8) were observed in responders to escitalopram. These changes were source-localized to regions associated with the default-mode network and the cingulate cortex, both known to be impaired in depression. Non-response to escitalopram was predicted by early changes in neural dynamics (baseline to week 2) prior to any clinical record of improvement in symptoms. Specifically, non-response was linked to a widespread decrease in delta, theta and beta oscillations and a decrease in complexity in mid-coarse timescales in the left fronto-central regions (approaching significance) from baseline to week 2, potentially suggesting an early medication-induced perturbance of neural dynamics in non-responders. From week 2 to week 8 of therapy no changes were observed in non-responders. With further evaluation, these characteristics of neural dynamics may have the potential to be reliable treatment targets and can potentially guide the development of faster-acting, more effective drugs for major depression.

6.2 Introduction

Major depressive disorder is one of the most prevalent health disorders (Andrade, Caraveo-Anduaga et al. 2003, Kessler, Berglund et al. 2003, Hardeveld, Spijker et al. 2010) yet treatment selection approaches are still rudimentary. Due to a poor understanding on the pathophysiology of depression and the wide variety of antidepressants available for treatment, patients with major depression undergo a trial-and-error process to identify the best medication. As a result, symptoms might last for months or even years (Solomon, Keller et al. 1997). Furthermore, over a third of all patients with major depression do not respond to two or more of the standard antidepressant medications and are termed treatment-resistant (Fava 2003, Berlim and Turecki 2007). Treatment efficacy may be improved with better insight into the long-lasting effects associated with therapeutic efficacy of antidepressants.

The therapeutic efficacy of antidepressant medications is well-recognized and their acute mechanisms are relatively well-understood (Stahl 1998). However, it is still unclear how the short-term neuropharmacological actions of antidepressant medications might translate to a long-term improvement in mood. With recent technological advancements in non-invasive methods for investigating neurophysiological changes, it may be possible to provide insight into the targets of successful treatments for depression. We hypothesized that measures of these neurophysiological changes may be sensitive enough to monitor the effects of treatment and treatment response over the course of an antidepressant. Given that several weeks of treatment are needed before patients exhibit clinical benefits (Quitkin, Rabkin et al. 1984, Gelenberg and Chesen 2000, Frazer and Benmansour 2002), antidepressants are suggested to induce adaptive changes in brain regions linked to mood and emotion (Harmer, Goodwin et al. 2009, Davidson, Irwin et al. 2003).

Several studies have investigated neurophysiological changes in the brain following a course of antidepressant treatment using electroencephalography (EEG) measures. Most of these studies focused on monitoring changes in EEG oscillatory activity between responders and non-responders of antidepressants. Response to antidepressants was associated with baseline EEG oscillations including high alpha (Ulrich, Renfordt et al. 1986, Ulrich, Haug et al. 1988, Ulrich,

Haug et al. 1994), occipital alpha asymmetry (Bruder, Stewart et al. 2001, Bruder, Sedoruk et al. 2008) and high prefrontal theta (Knott, Telner et al. 1996, Pizzagalli, Pascual-Marqui et al. 2001, Mulert, Juckel et al. 2007). Few studies studied the long-term effects of antidepressants on EEG oscillations but provided conflicting or insignificant results (Tarn, Edwards et al. 1993, Kwon, Youn et al. 1996, Knott, Mahoney et al. 2002, Bruder, Sedoruk et al. 2008). However, frequency measures do not encapsulate all the information contained in the EEG signal and may overlook other important properties. For example, it may be important to consider other aspects of the EEG signal such as the spatial and temporal dynamics of neural activity, and especially the non-linear properties of these dynamics such as complexity, to provide a broader view into the neurological mechanisms of successful treatments for depression.

Complexity analysis is one approach for studying the complex, non-linear dynamics of neural signals. The theoretical background of complexity analysis and its application to physiological systems has been previously discussed (Tononi, Sporns et al. 1994, Tononi and Edelman 1998, Sporns, Tononi et al. 2000, Costa, Goldberger et al. 2005). In addition, abnormalities in the complexity of neural signals were associated with developmental changes and mood disorders (McIntosh, Vakorin et al. 2013, Yang and Tsai 2013). A few studies have also shown that complexity analysis of neural signals offers a unique approach to understanding the pathophysiology of depression (Nandrino, Pezard et al. 1994, Li, Tong et al. 2008, Méndez, Zuluaga et al. 2012) as well as predicting response to treatment in depression (Thomasson, Pezard et al. 2000, Méndez, Zuluaga et al. 2012, Okazaki, Takahashi et al. 2013, Farzan, Atluri et al. 2017, Jaworska, Wang et al. 2018). Of the various methods available for complexity analysis, multiscale entropy investigates complexity at small and large temporal scales to reveal information on both local and global neuronal processing (Vakorin, Lippé et al. 2011, McIntosh, Vakorin et al. 2013, McDonough and Nashiro 2014).

In this study, we investigated changes in spectral content as well as in complexity of neural signals using power spectral density and multiscale entropy analysis of resting-state, eyes-closed EEG data. Data was collected over an eight-week course of escitalopram treatment at three time points (baseline, end of week 2, and end of week 8). We hypothesized that EEG measures of frequency and complexity will provide distinct information on the effects of antidepressant

treatment. In addition, we hypothesized that antidepressants will modulate both EEG neural oscillations and complexity of neural dynamics in a frequency and timescale dependent manner and that these changes will be specific to regions previously shown to be impaired in depression. Finally, we hypothesized that these frequency and spatio-temporal dynamics at baseline and the early change in these dynamics (2 weeks into treatment) may predict response to escitalopram.

6.3 Methods

Patients and data. In the CAN-BIND-1 study, 211 participants aged 18-60 years who met the DSM-IV requirements for major depressive disorder were recruited and completed the baseline visit. Of this group, 180 patients received 8 weeks of standardized escitalopram treatment (10-20mg). At the 8 week visit, responder or non-responder status was determined as $\geq$ 50% decrease in Montgomery–Åsberg Depression Rating Scale (MADRS) score from baseline. There were 85 responders and 95 non-responders at week 8 (see (Kennedy, Lam et al. *In Press*, 2018) for further details). Detailed descriptions on the clinical data, research protocol and data acquisition have been published previously (Lam, Milev et al. 2016, Baskaran, Farzan et al. 2017, Farzan, Atluri et al. 2017, Kennedy, Lam et al. *In Press*, 2018). In a subset of patients that were recruited at four sites participating in EEG acquisition, eight minutes of resting-state, eyes-closed EEG data was collected. The participating sites included: University of British Columbia (UBC), Toronto General Hospital (TGH), Queens University (QNS), and the Centre for Addiction and Mental Health (CAMH). The demographic and clinical characteristics are presented in **Table 6.1**.

EEG data was collected at baseline (within 3 days before the start of the treatment trial), at the end of week 2 (i.e., two weeks after the beginning of the trial) and at the end of week 8 (i.e., eight weeks after the beginning of the trial). Of the 180 patients, 124 patients were recruited to participate in EEG assessments. Fifteen patients were excluded because data from one or more of the visits was missing. Two patients were excluded because data from at least one or more visits were noisy. Therefore, this study included resting-state, eyes-closed EEG data collected from 107 patients at baseline, at the end of week 2, and at the end of week 8.

Treatment. Escitalopram was administered in an open-label manner, starting at 10 mg daily, which was increased to 20 mg daily at week 2 or later if clinically necessary. For patients who were unable to tolerate the 20 mg dose, the dose could be reduced to 10 mg at the discretion of the treating psychiatrist. Participants were clinically assessed every 2 weeks throughout the study period (8 weeks) including baseline (before administration of study medication). As mentioned, primary outcome measure was ≥ 50% decrease in MADRS from baseline to week 8. Of the 107 patients included in this study, 50 were responders and 57 were non-responders.

Inter-site Data Harmonization. All EEG datasets were standardized to the following parameters: 58 EEG electrodes common to all sites (excludes eye electrodes), 0.05-100Hz bandpass filter, Cz reference, 512Hz sampling rate. The EEG files were then exported as an EEGLAB (Delorme and Makeig 2004) dataset. Data was standardized using MATLAB R2012b-R2016a with the EEGLAB toolbox (v12.0.2.6b). Complete descriptions on the standardization of EEG data across sites in the CAN-BIND-1 study were recently published (Farzan, Atluri et al. 2017).

Data Preprocessing. During pre-processing, the 8-minutes of EEG data was divided into 2-s continuous epochs, bandpass-filtered between 1-80Hz, and notch-filtered at 60Hz. Using EEGLAB, independent component analysis was used to extract eye, muscle and electrode artifacts. Deleted EEG channels were interpolated using spherical spline interpolation (Perrin, Pernier et al. 1989) and data was re-referenced to an average reference. This preprocessing pipeline is currently made available as ERPEEG (http://www.tmseeg.com/multisiteprojects/).

EEG Power Spectral Density Analysis. The EEGLAB function spectopo was used to obtain the power spectrum for each electrode from 1 to 50Hz.

Multiscale entropy analysis was performed using the methods outlined in (Costa, Goldberger et al. 2005, Farzan, Atluri et al. 2017). Using the sample entropy equation, multiscale entropy was examined across all 58 electrodes with the coarse-graining process (for 70 scales). Sample entropy quantifies the variability of time series by estimating the predictability of amplitude patters across a time series. In our analysis, two consecutive data points were used for data

matching (m=2) and data points were considered to match if their absolute amplitude difference was less than 15% (i.e., r =0.15) of the standard deviation of the time series. Multiscale entropy was calculated for a 30-s continuous epoch.

EEG source localization was performed using an open-source application, Brainstorm (Tadel, Baillet et al. 2011). First, the locations of our 58 EEG electrode sites were co-registered to the ICBM152 MRI template in Brainstorm. The forward solution was then calculated using the OpenMEEG BEM head model (Gramfort, Papadopoulo et al. 2010) and the inverse solution was derived using sLORETA (Pascual-Marqui 2002), with the solution space constrained to the cortex surface. To localize the neural activity, we used the Destrieux Atlas, which provides 148 reconstructed sources in the MNI co-ordinate space (Destrieux, Fischl et al. 2010). After the 58-channel EEG data was mapped to the 148 reconstructed sources, multiscale entropy and power spectrum measures were calculated for all subjects at these sources.

Statistics. Subjects were grouped into two groups of antidepressant responders and non-responders: subjects were grouped as responders if there was a 50% or higher change in MADRS relative to baseline, and non-responders otherwise. Analysis of variance was used to (i) examine the effect of antidepressant response on multiscale entropy (1–70 timescales), absolute power (1–50Hz frequencies) for the main effect of Response (Responder, Non-Responder) and Time (Baseline; Week 2 and Week 8), and the interaction effect of Response x Time; and (ii) to examine the effect of site on multiscale entropy (1–70 timescales), absolute power (1– 50Hz) in responders and non-responders for the main effect of Site (UBC; TGH; QNS; CAM) and Time (Baseline; Week 2 and Week 8) across 58 electrodes in sensor space and 148 reconstructed sources in source space. Bootstrapping was used to correct for multiple comparisons in the analysis of variance. For post-hoc t-test comparisons, cluster-based non-parametric permutation tests were used to correct for multiple comparisons (Maris and Oostenveld 2007). In this multi-dimensional data [58 channels (or 148 reconstructed sources); 70 timescales; 50 frequencies], significance is assigned to the probability of clusters formed by pooling significant t-test results (p<0.05) adjacent along all dimensions of the data. The significance of each cluster is evaluated against the probability distribution of all clusters obtained over 1000 permutations. Identical parameters were used across all cluster-based permutation tests: threshold statistic of P<0.05,

identical channel (or reconstructed source) neighborhood matrices, and 1000 permutations using the monte-carlo approach where cluster statistics where computed as the maximum sum of cluster values.

Analysis of variance, post-hoc paired t-test and independent sample t-test analyses were used to calculate the original test statistics. Spearman correlation coefficient was used to examine the association between change in complexity and symptom severity or cognitive score. Similarly, cluster-based non-parametric permutation test was applied to correct for the multiple comparisons in the correlation analyses.

Throughout the paper, except otherwise noted, reported statistics are corrected p-values, and descriptive values indicate mean and standard deviation unless otherwise stated. Percent change (i.e. %) in multiscale entropy is calculated as: [(post treatment value - baseline value)/baseline value] *100; change in power is calculated as: [post treatment value - baseline value]; and percent change in MADRS was calculated as [(baseline score - post treatment)/baseline score] *100.

6.4 Results

6.4.1 Changes in neural oscillations following escitalopram treatment

There was a significant main effect of Time [mean F=11.5 (5.75 to 36.8)], Response [mean F=3.77 (3.59 to 3.94)] and a significant interaction effect of Time x Response [mean F=4.98 (3.51 to 9.13)] across several frequencies and channels. In addition, there was a significant main effect of Time [mean F=11.3 (5.33 to 36.8)], Site [mean F=4.62 (2.49 to 11.2)], and a significant interaction effect of Site x Time [mean F = 3.52 (2.42 to 6.77)]. In source space, there was a significant main effect of Time [mean F = 10.4 (5.62 to 34.6)], Response [mean F = 4.64 (3.30 to 6.09)] and a significant interaction effect of Time x Response [mean F = 5.08 (3.44 to 9.95)] and across several frequencies and ROIs. In addition, there was a significant main effect of Time [mean F = 10.3 (5.46 to 34.6)], Site [mean F = 5.40 (2.39 to 19.0)], and a significant interaction effect of Site x Time [mean F = 3.84 (2.30 to 11.8)].

All post-hoc tests are summarized below and were controlled for site. Post-hoc tests, comparing responders and non-responders at each time point (baseline, week 2 and week 8), were not significant.

6.4.1.1 Changes in neural oscillations over the course of treatment (baseline to week 8)

Post-hoc analyses for responders revealed a significant decrease in the high alpha, low beta and gamma bands (p=0.003, controlled for site: p=0.009) (left panel of **Figure S 6.1**). The significant decrease in high-alpha oscillations (10-12Hz) was shown to be widespread. The significant decrease in beta oscillations (12-30Hz) was observed in frontal (FP1, FP2, FPZ, AF3, AF4, F5, F3, FZ, F2, F8, FT7, FT8), fronto-central (FC5, FCZ, FC4, FC6), central (C1, C2, C5, C6), centro-parietal (CPZ, CP2, CP3, CP4, CP5, CP6), temporal (T7, T8, TP8), all parietal, all parieto-occipital and all occipital electrodes. Finally, the significant decrease in gamma oscillations (30-50Hz) was observed in the parietal (P1, P2, P3, P7), parieto-occipital (POZ, PO4), and occipital (O1, O2) electrodes. In source space (left panel of **Figure S 6.2**), a significant decrease in alpha, beta and gamma oscillations was also observed (p=0.008, controlled for site: p=0.02). The decrease in high-alpha oscillations (10-12Hz) was observed in the superior frontal gyrus and sulcus, left ACC, aMCC, pMCC, dPCC, vPCC, insular and insula regions, subcentral gyrus and sulci, subparietal sulcus, parieto-occipital sulcus, occipital gyrus and sulcus, occi-temporal gyrus, temporal gyrus and sulcus regions and the right occipital pole. The decrease in beta (12-30Hz) oscillations was observed in the aMCC, vPCC, insular regions, post-central gyrus, subparietal sulcus, inferior and lat-occipital gyrus and sulcus, temporal gyrus and sulcus, right occipital pole and the orbital gyri regions. Finally, the decrease in gamma (30-50Hz) oscillations was associated with the aMCC, paracentral lobule and sulcus, inferior occipital gyrus and the right temporal pole.

Post-hoc analysis for non-responders revealed a significant decrease in delta, theta, alpha, beta and gamma bands (p=0.006, controlled for site: p=0.01) (right panel of **Figure S 6.1**). A decrease in delta oscillations (1-4Hz) was observed in the central (C4) and centro-parietal (CP2)

electrodes. A decrease in the theta oscillations (4-8Hz) was observed in the frontal (FP1, FP2, AF3, AF4, F1, F2, F3, F5, F6, F8), fronto-central (FCZ, FC1, FC3, FC6), central (CZ, C2, C4, C5), centro-parietal (CPZ, CP1, CP2, CP3, CP5, CP6), parietal (P1, P3, P4, P5, P6, P7), parieto-occipital (PO3, PO7, PO8), and the occipital (OZ, O1, O2) electrodes. At the 10Hz alpha frequency, significant electrodes included frontal (FP1, AF3, F7, F3, F1), fronto-central (FC3, FC5), central (CZ, C3, C5), and parietal (P4, P6) regions. Across the alpha band (8-12Hz), significant electrodes included centro-parietal (CP2, CP5), parieto-occipital (PO3, PO7) and occipital (O1, OZ) regions. A decrease in beta oscillations (12-30Hz) was observed in the frontal (FP1, FP2, AF3, AF4, F1, F2, F3, F4, F5, F6, F7, FT7, FT8), fronto-central (FCZ, FC1, FC3, FC4, FC5, FC6), all central, centro-parietal (CP1, CP2, CP3, CP5, CP6), all parietal except P2, parieto-occipital (PO3, PO4, PO7, PO8), and all occipital electrodes. Finally, the decrease in gamma oscillations (30-50Hz) was observed in centro-parietal (CP5, CP6), parietal (P4, P6, P8), parieto-occipital (PO8) and occipital (OZ, O2) electrodes. In source space (right panel of **Figure S 6.2**), the decrease in delta, theta, alpha and beta oscillations was also significant (p=0.03, controlled for site: p=0.04). In delta-theta oscillations (1-8Hz), significant regions included the aMCC, insular regions, central sulcus occipital gyrus and sulcus, lat occi-temporal gyrus, and the right pre-occipital notch regions. Across the alpha band (8-12Hz), significant regions included the sup occipital gyrus (left) and the lat occi-temporal gyrus. In beta (12-30Hz) oscillations, significant regions were the transverse frontopolar gyri, inferior frontal sulcus (left), aMCC, insular regions, some occipital gyrus and sulcus, some temporal gyrus and sulcus, the middle and superior frontal sulcus (right), occi-temporal, and the right occipital pole.

Correlation analysis in sensor space (left panels of **Figure S 6.3**) revealed that a significant increase in theta oscillations (4-8Hz) was correlated with improvement in depressive symptoms (p=0.001) in frontal (F6, F7, FT8), fronto-central (FC4, FC6), central (C4), centro-parietal (CP2, CP5, CP6), tempro-parietal (TP8), and parietal (P6) electrodes. In addition, a significant decrease in the alpha, beta and gamma oscillations correlated with improvement in depressive symptoms (p=0.001). In the high-alpha band (10-12Hz) the effect was observed in frontal (FCZ, FC1, FC3), central (CZ, C1, C2, C6), centro-parietal (CPZ, CP2, CP4), parietal (P2, P4) and parieto-occipital (POZ) electrodes. In the beta band (12-30Hz), the effect was observed in centro-parietal (CP4), parietal (PZ, P2, P3) and parieto-occipital (POZ, PO4) electrodes and lastly, in the gamma band

(30-50Hz) the effect was observed in P2 and POZ. In source space (right panel of **Figure S 6.3**), significant correlations were observed in the delta, theta, alpha, beta and gamma bands (p=0.001). Significant regions in the delta band (1-4Hz) included the marginal cingulate cortex, dPCC, precuneus, cuneus, paracentral lobule and sulcus, central sulcus, parieto-occipital sulcus, pre-occipital notch, middle occipital sulcus and the superior temporal sulcus. In alpha oscillations, significant regions were global. In theta oscillations (4-8Hz), significant regions included the marginal cingulate sulcus, dPCC, precuneus, cuneus, paracentral lobule and sulcus, postcentral gyrus & sulcus, precentral gyrus, central sulcus, superior parietal lobule, preoccipital notch, subparietal sulcus, parieto-occipital sulcus, the left vPCC, left lat occ-temporal gyrus & sulcus and the right occipital pole. In beta oscillations (12-30Hz), significant regions included the dPCC, vPCC, precuneus, cuneus, short insular gyri, insula sup-circular sulcus, angular gyrus, intra&trans-parietal sulcus, subparietal sulcus, parieto-occipital sulcus, med occi-temp parahip gyrus, med occi-temp lingual sulcus, inferior temporal sulcus and the right occipital pole and preoccipital notch. Finally, regions identified in the **gamma** band (30-50Hz) were the left fronto-marginal gyrus, left transverse frontopolar gyri, left superior frontal sulcus, and the right straight gyrus.

6.4.1.2 Early changes in neural oscillations (baseline to week 2)

Post-hoc analyses for responders showed a decrease in mid-high beta and low gamma oscillations (18-42Hz) (p=0.03, controlled for site: p=0.07) in central (C1, C4), centro-parietal (CP1, CP2, CP3, CPZ), parietal (P7, P5, P3, P1, PZ, P2, P8) and parieto-occipital (PO7, PO3, POZ, PO4, PO8) and occipital (O1, OZ, O2) electrodes (left panel of **Figure S 6.4**). In source space (left panel of **Figure S 6.5**), this translated to the temporal pole, occi-temporal sulcus, occi-temporal gyrus, inferior occipital gyrus and sulcus, superior occipital gyrus, precentral gyrus, postcentral gyrus, superior precentral gyrus, paracentral lobule and sulcus, temporal pole, aMCC, vPCC, pMCC, short insular gyrus and other regions (p=0.04, controlled for site: p=0.08). In addition to the central-parieto-occipital cluster, another cluster was found only in the source space (p=0.04, controlled for site: p=0.08) showing an increase in high beta and gamma oscillations (25-50Hz). The regions in this cluster include fronto-inferior orbital, middle frontal gyrus, precentral gyrus, postcentral gyrus, superior precentral sulcus, occipital pole, orbital sulci,

lateral orbital sulcus, medial orbital sulcus, suborbital sulcus, short insular gyri, long insular gyrus, central sulcus of insula, postcentral sulcus, superior precentral sulcus, middle occipital sulcus, lateral-superior temporal gyrus, inferior temporal sulcus, superior temporal sulcus, straight gyrus and subcallosal gyrus.

Post-hoc analyses for non-responders revealed a significant global decrease in delta and theta oscillations (1-8Hz) (p=0.03, controlled for site: p=0.05). In addition, there was a significant decrease in beta (12-30Hz) and low-gamma (30-40Hz) oscillations (p=0.01, controlled for site: p=0.02) in all frontal electrodes, fronto-central (FC1, FCZ, FC2, FC4, FC6), central (CZ, C4, C6, C1, C3, C5), all centro-parietal, all parietal, parieto-occipital (PO7, PO3, PO4, PO8) and occipital (O1, OZ, O2) electrodes. Significant regions in low-gamma band (30-40Hz) include parietal (P3, P1, PZ, P2) and centro-parietal (CP5, CP6) electrodes (right panel of **Figure S 6.4**). In source domain (right panel of **Figure S 6.5**), the two clusters approached significance: delta-theta (1-8Hz) (p=0.09, controlled for site: p=0.10) and beta-gamma: 12-50Hz (p=0.08, controlled for site: p=0.09). Regions in the delta-theta cluster (1-8Hz) included the inferior frontal sulcus, superior frontal sulcus, ACC, aMCC, pMCC, marginal cingulate sulcus, insula regions, postcentral gyrus and sulcus, inferior precentral sulcus, superior precentral sulcus, intra and trans-parietal sulcus, occipital regions including the occipital pole, lateral occi-temporal gyrus and sulcus, temporal regions, lateral, medial and suborbital sulcus. Regions in the beta cluster (12-30Hz) include the inferior frontal sulcus, middle frontal sulcus, parieto-occipital sulcus, lateral occi-temporal gyrus and sulcus, plan-tempo superior temporal gyrus, and orbital gyri. In gamma (30-50Hz), regions included the right middle frontal sulcus, right superior frontal sulcus and the right pMCC.

Correlation analysis revealed that an increase in delta-theta (1-8Hz) and beta (18-30Hz) oscillations significantly correlated with improvement in depressive symptoms (left panels of **Figure S 6.6**). For the delta-theta (1-8Hz) cluster (p=0.001), this effect was observed in the frontal (AF3, F7, F5, FZ, F4, F6), fronto-central (FC1, FC5, FCZ, FC2, FC4, FC6), central (C3, C4, C5, C6), centro-parietal (CP2, CP3, CP4, CP5, CP6), tempro-parietal (TP7, TP8), parietal (P1, P3, PZ, P2, P4), parieto-occipital (PO3, PO7) and occipital (O1, OZ, O2) electrodes. In beta (18-30Hz) oscillations (p=0.001), this effect was observed in the frontal (AF3, F7, F5, FZ, F4,

F6, F8), fronto-central (FC1, FC4, FC5, FC6), central (C5, C6), centro-parietal (CP3, CP5, CP6), parietal (P2, P4, P6) and parieto-occipital (PO3, PO7) electrodes. In source space (right panel of **Figure S 6.6**) significant regions in the delta-theta cluster (p=0.001) included the middle frontal gyrus, inferior and superior frontal sulcus, ACC, central sulcus, pre and post-central gyrus, postcentral sulcus, paracentral lobule and sulcus, superior parietal lobule, superior parietal lobule, superior occipital gyrus, middle occipital sulcus, lateral orbital sulcus, middle occipital sulcus, lat-sup temporal gyrus, subparietal sulcus, preoccipital notch and the lateral orbital sulcus regions. In source space for beta oscillations, significant regions (p=0.001) in the left hemisphere were the front-inf orbital & triangular, pMCC, marg-cingulate sulcus, paracentral lobule & sulcus and the subparietal sulcus. Significant regions in the right hemisphere were the pMCC, vPCC, precentral gyrus, central sulcus and the preoccipital notch. In gamma oscillations, significant regions (p=0.001) included the pMCC and the dPCC. In the right hemisphere, significant regions included the vPCC, postcentral gyrus, precentral gyrus, central sulcus, postcentral sulcus, and the inferior & superior precentral sulcus.

6.4.1.3 Changes in neural oscillations over latter time course of treatment (week 2 to week 8)

Post-hoc analyses for responders showed a decrease (p=0.006, controlled for site: p=0.06) in the high-alpha and beta bands (left panel of **Figure S 6.7**). In the high-alpha band (10-12Hz), this effect was observed in the fronto-central (FC1, FC3), temporal (T7), central (C3, C5), centro-parietal (CP3), parietal (P1), parieto-occipital (PO3, PO7) and occipital (O1) electrodes. In the beta band (12-30Hz), this effect was observed in the frontal (FP1, AF3, AF4, F5, F3, F1, FZ, F2, F4, F6, F8, FT7, FT8), fronto-central (FCZ, FC1, FC3, FC4, FC5, FC6), temporal (T7), central (C5, C6), centro-parietal (CP3, CP4, CP5, CP6), parietal (P1, P2, P3, P6) and parieto-occipital (PO3, PO7) electrodes. Results were not significant in the source space. Post-hoc analysis for non-responders did not show any significant results (right panel of **Figure S 6.7**).

Correlation analysis in sensor space revealed that a decrease in neural oscillations (delta, theta, alpha, beta and gamma) was significantly correlated with improvement in depressive symptoms (p=0.001) (left panels of **Figure S 6.8**). In the delta band (1-4Hz), the effect was observed in

frontal (AF3, FZ, F5), fronto-central (FC5), centro-parietal (CP4, CP6), parietal (P1, P2) and parieto-occipital (PO3, PO7) electrodes. In the theta band (4-8Hz), the effect was observed in frontal (FP1, AF3, F5, F7, FT7), fronto-central (FC1, FC5), parietal (P1, P3) and parieto-occipital (PO3, PO7) electrodes. In the high-alpha band (10-12Hz), the effect was observed in frontal (F5), central (C3, C5) and centro-parietal (CP3, CP4, CP5) electrodes. In the beta band (12-30Hz), the effect was observed in frontal (mostly in AF3, FZ, F4, F5), fronto-central (mostly FC1, FC4, FC5, FC6), central (C5, C6), centro-parietal (CP3, CP4, CP6), parietal (PZ, P1, P2, P3, P7,) and parieto-occipital (POZ, PO3, PO7) electrodes. Finally, in the gamma band (30-50Hz), the effect was observed in frontocentral (FC6), central (C6), parietal (P2) and parieto-occipital (POZ) electrodes. In the source space (right panel of **Figure S 6.8**), the correlation effect was observed across all bands. An increase in delta, theta, high beta, gamma and a decrease in high-alpha & mid-beta oscillations were significantly correlated with improvement in depressive symptoms (p=0.001 for both). For the increase in delta power (1-4Hz), significant regions included the left preoccipital notch, left & right superior temporal sulcus, right insula inf-circular sulcus, right intra- & trans-parietal sulcus and the right temporal gyrus. For the increase in theta band power (4-8Hz), significant regions in the right hemisphere included the dPCC, precuneus, insula inf- & sup-circular sulcus, long insular gyrus & central sulcus of insula, postcentral sulcus, angular gyrus, superior parietal lobule, intra- & trans-parietal sulcus, med occi-temp parahip gyrus and the temporal gyrus. For the decrease in high-alpha band power (10-12Hz), significant regions included the insula circular sulcus, central & post-central sulcus, temporal gyrus, med occi-temp lingual sulcus, left dPCC, left precuneus, left pre- & postcentral gyrus, right superior frontal sulcus and the right long insular gyrus & central sulcus of insula. For the decrease in mid-beta band power (18-22Hz), significant regions included the left parieto-occipital sulcus, left straight gyrus, right front-inf-opercular, right cuneus, right subcentral gyrus & sulci, right temporal gyrus & sulcus and the right orbital sulcus. For the increase in high-beta band power (22-30Hz), significant regions included the left dPCC, left precuneus, left precentral gyrus, left postcentral sulcus, left subparietal sulcus, right mid-occipital gyrus, right med occi-temp lingual & parahip gyrus and the right lat occi-temporal sulcus. Finally, the increase in gamma band power (30-50Hz) was observed in the dPCC, paracentral lobule & sulcus, precentral gyrus, central sulcus, left superior frontal gyrus & sulcus, left pMCC, left marg-

cingulate sulcus, left precuneus, left long insular gyrus & central sulcus of insula and the left subparietal sulcus.

6.4.2 Changes in temporal complexity of neural dynamics following escitalopram treatment

There was a significant main effect of Time [mean F = 8.27 (5.67 to 19.8)]], Response [mean F = 4.20 (3.67 to 4.63)] and the interaction effect of Time x Response [mean F = 6.13 (3.30 to 22.1)] and across several frequencies and channels. In addition, there was a significant main effect of Time [mean F = 8.22 (5.75 to 19.8)], Site [mean F = 3.89 (2.57 to 8.64)], and Site x Time [mean F = 3.71 (2.40 to 8.38)] interaction effect. In source space, there was a significant main effect of Time [mean F = 8.24 (5.52 to 19.1)], Response [mean F = 4.75 (3.54 to 9.25)] and the interaction effect of Time x Response [mean F = 5.76 (3.52 to 19.3)] and across several frequencies and ROIs. In addition, there was a significant main effect of Time [mean F = 8.39 (5.60 to 19.1)], Site [mean F = 5.74 (2.45 to 19.9)], and Site x Time [mean F = 3.78 (2.40 to 8.93)] interaction effect.

Post-hoc test and correlation results are presented below. All post-hoc tests were controlled for site. Post-hoc tests, comparing responders and non-responders at each time point (baseline, week 2 and week 8), were not significant.

6.4.2.1 Changes in temporal complexity over the course of treatment (baseline to week 8)

Post-hoc analysis did not show any significant results for responders and non-responders. Correlation analysis revealed that an increase in complexity in mid-high timescales (15-55) correlated with improvement in depressive symptoms (p=0.001). This effect was observed in the parietal (PZ, P2, P3, P5, P7) and parieto-occipital (PO3, PO7) electrodes (left panel of **Figure 6.1**). Highest correlations were seen in channel P7/scale 34 (r = 0.358, p = 0.0002) and in channel P7/scale 47 (r = 0.384, p < 0.0001) (right panels of **Figure 6.1**). In source space, this effect (p=0.001) was observed in the left: precuneus, vPCC, dPCC, superior parietal lobule and subparietal sulcus; and the right: precuneus, vPCC, insula inf-circular sulcus, superior parietal

lobule, parieto-occipital sulcus, preoccipital notch, med occi-temp lingual gyrus and sulcus, plan-tempo sup temporal gyrus and the inferior temporal sulcus (left panel of **Figure 6.2**). Highest correlation was seen in the dPCC at scale 22 (r = 0.270, *p* = 0.0049) (right panel of **Figure 6.2**).

6.4.2.2 Early changes in temporal complexity (baseline to week 2)

Post-hoc analysis for responders did not show any significant results (left panels of **Figure 6.3**). Post-hoc analysis for non-responders revealed a decrease in complexity in mid-coarse scales (20-65) approaching significance (*p*=0.08, controlled for site: *p*=0.09). This effect was observed in the frontal (AF3, FZ, F1, F3, F5, F7, FT7), fronto-central (FC1, FC3, FC5), temporal (T7), and the central (C1, C3) electrodes. Source analysis did not reveal any significant effects (right panels of **Figure 6.3**).

Correlation analysis in sensor space did not reveal any significant effects. However, in source space there was significant correlation between decrease in complexity in low timescales (5-20) and improvement in depressive symptoms (*p*=0.001). This effect was seen in left inferior frontal sulcus, left ACC, left insula ant-circular sulcus, left orbital gyri and sulci, left suborbital sulci and the right straight gyrus (left panel of **Figure 6.4**). Highest correlations were seen in the ACC at scale 11 (r = -0.250, *p* = 0.0093) (right panels of **Figure 6.4**).

6.4.2.3 Changes in temporal complexity over the latter time course of treatment (week 2 to week 8)

Post-hoc analysis for responders revealed a significant increase in complexity in mid-coarse scales (20-70) (*p*=0.03, controlled for site: *p*=0.04). This effect was observed in the frontal (FP2, AF3, AF4, FZ, F1, F2, F3, F4, F5, F6, F7, F8 FT7, FT8), all fronto-central except FC1, temporal (T7, T8, TP7, TP8), all central except C5, all centro-parietal except CPZ and CP4, all parietal except PZ, all parieto-occipital except PO4 and occipital (O1, OZ) electrodes (left panels of **Figure 6.5**). Source analysis did not reveal any significant effects. Post-hoc analysis for non-responders also did not show any significant results (right panels of **Figure 6.5**).

Correlation analysis revealed that an increase in complexity in fine (1-10), mid-high (15-50) and coarse (60-70) timescales was significantly correlated with improvement in depressive symptoms (p=0.001). In fine timescales (1-10), this effect was observed in the frontal (AF4, F4, F6), fronto-central (FC3, FCZ), temporal (T7, TP7), central (CZ, C2, C5), centro-parietal (CP3, CP4, CP6), parietal (P5, P6, P7, P8) and parieto-occipital (PO7) electrodes. In mid-high timescales (15-50), this effect was observed in the centro-parietal (CP5, CP6), parietal (P5, P6, P7, P8), and parieto-occipital (PO7) electrodes. In the coarser timescales (60-70), the effect was observed in the parietal (P5, P7) and parieto-occipital (PO7) electrodes (left panels of **Figure 6.6**). Highest correlations were seen in channel P7/scale 18 (r = 0.353, p = 0.0002) and channel P7/scale 41 (r = 0.343, p = 0.0003) (right panels of **Figure 6.6**). The effect was also significant in source space (p=0.001) (left panel of **Figure 6.7**). In fine timescales (1-10), significant regions included the dPCC, vPCC, precuneus, central sulcus, subparietal sulcus and the mid occipital gyrus. In the left hemisphere, additional regions included the pMCC, cuneus, superior frontal gyrus and sulcus, paracentral lobule and sulcus, angular gyrus, superior parietal lobule, mid-occipital gyrus, inf occipital gyrus & sulcus, preoccipital notch and the lat occi-temporal gyrus & sulcus. In the right hemisphere, regions included the middle frontal gyrus & sulcus, precentral gyrus, superior precentral sulcus, and the med occi-temp parahip gyrus. In mid-timescales (15-50), significant regions included the vPCC, left preoccipital notch, left lat occi-temporal sulcus and the left inferior temporal gyrus. Finally, in coarse timescales (60-70), significant regions were the left preoccipital notch and the left lat occi-temporal sulcus. Highest correlations were seen in the subparietal sulcus at scale 3 (r = 0.357, p = 0.0002) and in the lat occi-temporal sulcus at scale 23 (r = 0.339, p = 0.0003) (right panels of **Figure 6.7**).

6.4.3 Association between neural oscillations and mood

6.4.3.1 Association between baseline neural oscillations and mood

Correlation analysis in sensor space revealed that high power in alpha, mid-beta and gamma oscillations correlates with improvement in symptoms (p=0.001). In the alpha band (8-12Hz), the effect was observed in frontal (AF3, F3, F5), fronto-central (FCZ, FC3), central (CZ, C2, C3), centro-parietal (CP1, CP2, CP4), and parietal (P1, P2) electrodes. In the mid-beta band (18-22Hz), the effect was observed in the frontal (F5, F7), central (C1, C3), centro-parietal (CP1),

parietal (P1, P2, P3, P4) and parieto-occipital (POZ, PO4) electrodes. In the gamma band (30-50Hz), the effect was observed in parietal electrodes (P1, P3) (left panels of **Figure S 6.9**). This effect was also significant in the source space (p=0.001). Results in the alpha band corresponded to the superior frontal sulcus, subcentral gyrus and sulci, parieto-occipital sulcus and the orbital gyri. Results in the mid-beta band corresponded to several regions including the precentral gyrus, superior parietal lobule, parieto-occipital sulcus and the right occipital pole. No significant regions were observed in the gamma band (right panel of **Figure S 6.9**).

6.4.3.2 Association between neural oscillations 2 weeks after treatment and mood

Correlation analysis in sensor space revealed that high power in the alpha, beta and gamma oscillations correlates with improvement in symptoms (p=0.001) In the alpha band (8-12Hz), the effect was seen in frontal (AF3, FZ, F5, F6, F7), fronto-central (FC3), central (C3, C4, C5), temporal (T7, TP8), centro-parietal (CP2, CP3, CP4, CP5, CP6) and parietal (P1, P2) electrodes. In the low-beta band (12-18Hz), the effect was seen in frontal (F5, F6, F7) and parietal (P1) electrodes. In the mid-beta band (18-22Hz), the effect was seen in frontal (AF3, FZ, F5, F6, F7, F8), fronto-central (FC1, FC4, FC6), central (C3, C4, C5, C6), centro-parietal (CP3, CP5, CP6), parietal (P1, P2, P3) and parieto-occipital (POZ, PO4) electrodes. In the high-beta band (22-30Hz), the effect was seen in frontal (F5, F6, F7), fronto-central (FC1), central (C6), centro-parietal (CP6) and parietal (P1) electrodes. In the gamma band (30-50Hz), the effect was seen in frontal (F3, F5, F6, F7), fronto-central (FC1, FC4, FC6), central (C3, C6), centro-parietal (CP5), parietal (P1, P3) and parieto-occipital (PO4) electrodes (left panels of **Figure S 6.10**).

This cluster effect was also significant in source space in the alpha, beta and gamma oscillations (p=0.001). In the alpha band (8-12Hz), significant regions included the inferior frontal sulcus, aMCC, central sulcus, temporal pole, orbital gyri, medial orbital sulcus, and the right ACC and right occipital pole. In the low and high-beta bands (12-18Hz, 22-30Hz), significant regions included the inferior frontal sulcus, aMCC, and the orbital sulci. In the mid-beta band (18-22Hz), significant regions included the inferior frontal sulcus, aMCC, temporal pole, orbital gyri and sulci, the left parieto-occipital sulcus, right ACC, and the right occipital pole. Finally, in the

gamma band (30-50Hz), significant regions included the inferior frontal sulcus, aMCC, orbital gyri and sulci, pericallosal sulcus, and the right ACC (right panel of **Figure S 6.10**).

6.4.4 Association between temporal complexity of neural dynamics and mood

6.4.4.1 Association between baseline temporal complexity and mood

Correlation analysis revealed that a low complexity value at baseline correlates with improvement in depressive symptoms (p=0.001) across coarser timescales (30-70) (left panels of **Figure 6.8**). This effect was observed in the frontal (mostly FPZ, F5, F6, F7), fronto-central (FC1, FC3, FC4, FC5, FC6), temporal (T8), central (C1, C2, C3, C4, C6), centro-parietal (CPZ, CP1, CP2), parietal (PZ, P1, P2), parieto-occipital (POZ, PO4) and occipital (OZ, O2) electrodes. Highest correlations were seen in channel F5/scale 40 (r = -0.364, p = 0.0001), channel F6/scale 66 (r = -0.309, p = 0.0012), channel C1/scale 54 (r = -0.335, p = 0.0004) and channel O2/scale 58 (r = -0.302, p = 0.0016) (right panels of **Figure 6.8**). In source space (left panels of **Figure 6.9**), this effect was significant (p=0.001) in coarser timescales in brain regions including the aMCC, pMCC, middle frontal sulcus, insular regions, temporal pole, occipital and temporal regions as well as the left central regions and the right occipital pole. A positive cluster was also found to be significant (p=0.004) in fine timescales (12-15) in the postcentral gyrus and sulcus, intra- & trans-parietal sulcus, plan-tempo sup temporal gyrus, transverse temporal sulcus, orbital gyri, suborbital sulcus and the lat-fis posterior regions (all in left hemisphere). Highest correlations were seen in the superior precentral sulcus at scale 56 (r = -0.381, p = 0.0001) and in the orbital gyri at scale 14 (r = 0.253, p = 0.0086) (right panels of **Figure 6.9**).

6.4.4.2 Association between temporal complexity 2 weeks after treatment and mood

Correlation analysis revealed that a low complexity value at week 2 in coarser timescales (30-70) (p=0.001) and a high complexity value at week 2 in finer timescales (12-17) (p=0.001) correlates with improvement in depressive symptoms (left panels of **Figure 6.10**). The cluster in the coarser timescales was observed in frontal (mostly in FP1, FP2, FPZ, AF3, AF4, FZ, F4, F5, F6), fronto-central (FCZ, FC4, FC6), central (CZ, C1, C2), centro-parietal (CPZ, CP5, CP6), parietal

(PZ, P2), parieto-occipital (PO4, PO8), and occipital (OZ, O1) electrodes. Highest correlations were seen in channel FC2/scale 18 (r = 0.246, p = 0.01), channel CP6/scale 54 (r = -0.288, p = 0.0026), channel F6/scale 68 (r = -0.294, p = 0.0021) and channel FC6/scale 66 (r = -0.294, p = 0.0021) (right panels of **Figure 6.10**). The cluster in the finer timescales was observed in frontal (F2, F4), fronto-central (FC2, FC4) and central (CZ) electrodes. In source space (left panels of **Figure 6.11**), the cluster in the coarser timescales was significant (p=0.001) in several brain regions including the inferior frontal sulcus, superior frontal gyrus, ACC, aMCC, pMCC, dPCC, vPCC, short insular gyri, long insular gyrus and the central sulcus of insula, central sulcus, superior and inferior precentral sulcus, parieto-occipital sulcus, occipital pole, lat-sup temporal gyrus, and the temporal pole. Highest correlations were seen in the plan-polar sup temporal gyrus at scale 61 (r = -0.333, p = 0.0005), temporal pole at scale 65 (r = -0.301, p = 0.0016), ACC at scale 61 (r = -0.306, p = 0.0013) and med occi-temp lingual sulcus at scale 66 (r = -0.320, p = 0.0008) (right panels of **Figure 6.11**).

6.5 Discussion

Results from frequency and complexity analysis revealed a distinct pattern of early (baseline to week 2) and long-term (week 2 to week 8, or baseline to week 8) changes in patterns of neural activity with escitalopram treatment. In responders, early changes were local and frequency-specific. Subsequent to these early changes, over the next 6 weeks of treatment (week 2 to week 8), there were significant changes in frequency and complexity scores that correlated with improvement in depressive symptoms. In non-responders, early significant changes (baseline to week 2) were seen in several regions, frequencies and timescales and these early significant changes negatively correlated with improvement of depressive symptoms following the eight-week course of escitalopram. Following this early change, we observed no further significant changes in neural activity in non-responders from week 2 to week 8. These findings potentially reveal a differential effect of antidepressant medications on neural dynamics in responders and non-responders. In addition, these findings also provide evidence in support of using baseline markers of neural dynamics and early changes in neural dynamics with treatment to predict treatment outcome, and potentially inform individualized treatment.

Changes in neural oscillations following a course of antidepressant treatment have been reported in previous studies. Through this study, we replicated several of these results. For example, higher alpha at baseline was previously shown to be correlated with antidepressant response (Ulrich, Renfordt et al. 1986, Ulrich, Haug et al. 1988, Ulrich, Haug et al. 1994). Through our correlation analysis, we also showed that higher values of alpha at baseline and two weeks after treatment were correlated with antidepressant response (see **Figure S 6.9** and **Figure S 6.10**). Alpha asymmetry at baseline was also shown to differentiate responders and non-responders in previous studies (Bruder, Stewart et al. 2001, Bruder, Sedoruk et al. 2008) with largest differences in occipital sites in the right hemisphere. In this study, alpha power in the right hemisphere (frontal and occipital) at baseline and at week 2 was associated with treatment response. This was not observed in the left hemisphere (right panels of **Figure S 6.9**, **Figure S 6.10**).

In addition to the replication of some previous findings, this study provides additional insights on the longitudinal effect of antidepressants (over 8 weeks) on neural dynamics. Reports on longitudinal effects of antidepressants on neural dynamics are sparse and inconsistent for EEG frequency analysis (Tarn, Edwards et al. 1993, Kwon, Youn et al. 1996, Knott, Mahoney et al. 2002, Bruder, Sedoruk et al. 2008) and sparse for complexity analysis (Méndez, Zuluaga et al. 2012). In this study, we observed a decrease in alpha oscillations specific to responders and a decrease in theta oscillations specific to non-responders of escitalopram. We also demonstrated significant changes in beta oscillations associated with antidepressant response. Previous studies have shown that severity of depressive symptoms may be associated with increased alpha (Pollock and Schneider 1990, Bruder, Tenke et al. 2005, Korb, Cook et al. 2008), increased theta (Kwon, Youn et al. 1996) and dominant beta oscillations (Fingelkurts, Fingelkurts et al. 2006). Using source localization analysis in this study, responders also revealed significant changes in alpha oscillations (left panel of **Figure S 6.2**) in regions previously associated with mood regulation and the pathophysiology of depression such as the ACC (Mayberg, Brannan et al. 1997, Wu, Buchsbaum et al. 1999), cingulate cortex (Mayberg, Brannan et al. 1997, Greicius, Flores et al. 2007), orbitofrontal cortex (Greicius, Flores et al. 2007, Frodl, Bokde et al. 2010), occipital pole (Koch and Schultz 2014, Maller, Thomson et al. 2014), precuneus (Greicius,

Flores et al. 2007, Peng, Liddle et al. 2015) and the superior temporal cortex (Steele, Currie et al. 2007). Given these results, we suggest that antidepressants (such as escitalopram) do target regions associated with abnormal neural activity in depression for therapeutic response.

Measures from frequency analysis provide interesting insight into the mechanism of action of antidepressants and may also predict therapeutic outcome. However, as mentioned before, another aim of our study was to investigate whether other properties of the EEG signal might provide unique insight not provided by frequency analysis, such as complexity. To our knowledge, only two studies investigated whether pre-treatment neural complexity can predict response to antidepressant medication in depression (Méndez, Zuluaga et al. 2012, Jaworska, Wang et al. 2018). In *Mendez et al.,* a reduction in Lempel-Zev complexity of neural signals as measured by MEG was shown to be associated with response to mirtazapine. Lower complexity at baseline was also associated with antidepressant response in the current study. However, results cannot be easily compared to the current study because complexity was not calculated at multiple timescales in *Mendez et al*. In comparison, similar to our study, *Jaworski et al.* used multiscale entropy on resting-state EEG data to predict response to 12-weeks of escitalopram or bupropion plus placebo, or escitalopram and bupropion treatment in 36 patients with major depressive disorder. Of the 36 patients, 20 were responders. They showed that increased baseline complexity in mid-coarse timescales (frontal, central, parietal regions) was positively associated with treatment response and decreased complexity in fine timescales (fronto-central regions) was negatively associated with treatment response. In our study, decreased baseline complexity was associated with improvement in depressive symptoms in the same regions as *Jaworska et al.*, (frontal, central, parietal) (**Figure 6.8**). However, timescales and parameters for multiscale entropy analysis vary between our study and *Jaworska et al.*, and therefore, direct comparisons of results are not possible.

The trajectory of neural dynamics over the course of escitalopram treatment was shown to vary between responders and non-responders in the current study. While responders revealed transient changes in neural dynamics (from baseline to week 2 and then from week 2 to week 8), non-responders did not reveal any significant changes after 2 weeks of treatment (in frequency and complexity). We hypothesize that this differential effect may, in part, be attributed to variations

in: (i) neural transmission in the serotoninergic system, and/or (ii) neural plasticity in molecular systems between responders and non-responders known to be associated with the mechanism of action of selective serotonin re-uptake inhibitors (SSRIs) such as escitalopram (Zhong, Haddjeri et al. 2012). Previous studies have provided evidence that serotonin-receptor mediated shifts can be inferred from the EEG frequency spectrum (Saletu, Grünberger et al. 1986, McAllister-Williams and Massey 2003). A molecular marker of neural plasticity (polymorphism of the brain-derived neurotrophic factor (BDNF) Val66Met gene) was also previously associated with EEG alpha frequency oscillations (Gatt, Kuan et al. 2008, Zoon, Veth et al. 2013). In addition, at the level of large-scale neural circuits, plasticity was suggested to be reflected in the complexity of cortical neural signals. More specifically, transient increases and decreases in correlated activity at various timescales across brain regions are thought to reflect the rate of information generation and processing in the brain (McIntosh, Vakorin et al. 2013). The potential relationship between the modulation of these receptor-level, gene-level and circuit-level markers by antidepressants and changes in neural dynamics following antidepressant treatment may provide insight into the targets of successful antidepressants.

The mechanism of action of SSRIs (such as escitalopram) is generally assumed to be mediated by the binding of the drug to serotonin transporters resulting in increased extracellular concentrations of serotonin across the serotonergic pathways. However, a recent PET study in humans (Nord, Finnema et al. 2013) reported that a single dose of escitalopram may lead to an initial decrease in serotonin concentrations in occipital and temporal regions. Over the course of the medication, there may be a desensitization of inhibitory serotonin receptors and serotonin concentrations increase as expected. We hypothesize that the neurobiological effects of this initial decrease in serotonin concentrations may differ between responders and non-responders. Initial changes in neural oscillations (i.e., baseline to week 2) in responders were mainly seen in the occipital and also the temporal regions of the brain (left panel of **Figure S 6.5**). Non-responders revealed more widespread changes in neural oscillations leading to the hypothesis that perhaps the initial effect of escitalopram on serotonin receptors is local (occipital and temporal) in responders but may be more global in non-responders.

At the gene level, BDNF is known to be involved in neural cell proliferation and synaptic plasticity (Katz and Shatz 1996, Duman and Monteggia 2006). Reduced secretion of BDNF by the genetic polymorphism BDNF Val66Met, was previously reported to be associated with severity of depression (Jiang, Xu et al. 2005, Verhagen, Van Der Meij et al. 2010, Molendijk, Bus et al. 2011, Czira, Wersching et al. 2012) and normalization of BDNF levels was associated with antidepressant response (Sen, Duman et al. 2008, Zou, Ye et al. 2010, Molendijk, Bus et al. 2011). In addition, reduced secretion of BDNF was hypothesized to be associated with reduced functional connectivity and oscillatory activity in neuronal systems impaired in depression (Thomason, Yoo et al. 2009). Specifically, polymorphism of BDNF Val66Met was shown to be mediated by parieto-occipital alpha (Gatt, Kuan et al. 2008, Zoon, Veth et al. 2013) and increased parieto-occipital alpha was previously shown to be higher in depression (Pollock and Schneider 1990, Bruder, Tenke et al. 2005). In the current study, responders of escitalopram revealed a reduction in alpha oscillations after eight weeks of escitalopram treatment (left panel of **Figure S 6.1**) while non-responders revealed no changes in alpha power. Based on these results and the work from the previous studies, we hypothesize that non-responders may show impairments in the biological pathway including BDNF *Val66Met*, the neural networks generating EEG alpha oscillations and the resulting depressive symptoms underlying severity of depression.

Functional connectivity of resting-state neural networks was previously associated with complexity of (fMRI) neural signals (McDonough and Nashiro 2014). Although it has not been studied as extensively as receptor-level effects of antidepressants, studies have investigated changes in large-scale neural circuits following antidepressant treatment. In a recent meta-analysis (Kaiser, Andrews-Hanna et al. 2015), depression was associated with functional brain network dysfunction, specifically, hypoconnectivity in the frontoparietal network and hyperconnectivity in the default-mode network. Longitudinal studies monitoring the effect of antidepressants reported that antidepressants may normalize connectivity in the frontoparietal network (Alexopoulos, Hoptman et al. 2012) and also in the default-mode network (Andreescu, Butters et al. 2009, Wu, Andreescu et al. 2011, Li, Liu et al. 2013). In the current study, reduced complexity and neural oscillations were observed in regions associated with the DMN (precuneus, PCC) as well as the cingulate cortex (ACC, PCC). In addition, baseline complexity

and neural oscillations in these regions were associated with antidepressant response. Together, these results further add to the current line of neuroimaging evidence for the disruption and modulation of large-scale neural networks by antidepressants. The identification and future assessment of such markers are important for clinical translation.

6.6 Conclusions

In conclusion, several frequency and spatiotemporal-specific changes in neural dynamics were associated with response to escitalopram in patients with major depressive disorder. Results from prior studies associating baseline neural dynamics with antidepressant response were replicated and new insights into the longitudinal effects of antidepressants on neural dynamics were revealed. Responders to escitalopram were associated with reduced alpha and beta oscillations in the fronto-central and parieto-occipital regions (baseline to week 8) and increased complexity in mid-coarse timescales in fronto-central and centro-parietal regions (week 2 to week 8). These were source-localized to regions associated with the DMN (precuneus, PCC) and the cingulate cortex (ACC, PCC), known to be impaired in depression. Non-responders were associated with early changes in neural dynamics (baseline to week 2). Specifically, widespread effects were observed in high-delta, theta and beta oscillations and decreased complexity approached significance in mid-coarse timescales in the left fronto-central regions. No changes were observed in non-responders following 2 weeks of escitalopram treatment suggesting an early, potentially medication-induced perturbance of neural dynamics by medications. This perturbance was hypothesized to rise from potential abnormalities in the serotonergic and/or neuroplasticity mechanism of action of escitalopram specific to non-responders. Future work should investigate extended longitudinal effects of antidepressants on neural dynamics in patients (i.e., following the completion of a treatment and over several treatment trials) to better understand the transition from non-response to treatment-resistance. The combination of genetics and neuroimaging in the future is expected provide better insight into the specific targets of antidepressants. Uncovering reliable targets of treatment can improve the efficacy of current treatments for depression and potentially contribute towards to the development of novel treatments with improved targets for high efficacy.

6.7 Tables

Table 6.1 - Demographic and Clinical Characteristics

	UBC	TGH	QNS	CAM	ALL	Responders	Non-responders
N	49	36	15	7	107	50	57
Age (mean, std)	35.5 +/- 11.4	34.2 +/- 12.2	46.9 +/- 11.7	30.4 +/- 13.0	36.3 +/- 12.5	36.4 +/- 13.0	36.2 +/- 12.0
Gender (M/F)	17/32	12/24	8/7	0/7	37/70	16/34	21/36
MADRS Baseline (mean, std)	28.3 +/- 5.93	32.6 +/- 5.71	29.3 +/- 4.38	28.0 +/- 5.16	29.9 +/- 5.90	29.5 +/- 6.00	30.2 +/- 5.85
MADRS week2 (mean, std)	21.5 +/- 7.43	24.8 +/- 11.2	22.9 +/- 5.13	21.1 +/- 3.98	22.8 +/- 8.53	19.8 +/- 8.38	25.4 +/- 7.83
MADRS week8 (mean, std)	13.8 +/- 8.70	18.1 +/- 11.7	17.3 +/- 9.56	15.7 +/- 6.01	15.8 +/- 9.90	7.7 +/- 5.06	22.9 +/- 7.22
Change in MADRS (baseline to week8) (mean, std) (%)	51.4 +/- 31.1	44.8 +/- 31.9	40.0 +/- 35.9	44.5 +/- 16.9	47.1 +/- 31.3	73.8 +/- 16.3	23.0 +/- 20.8
Responders / Nonresponders	24 / 25	17 / 19	6 / 9	3 / 4	50 / 57	-	-

6.8 Figures

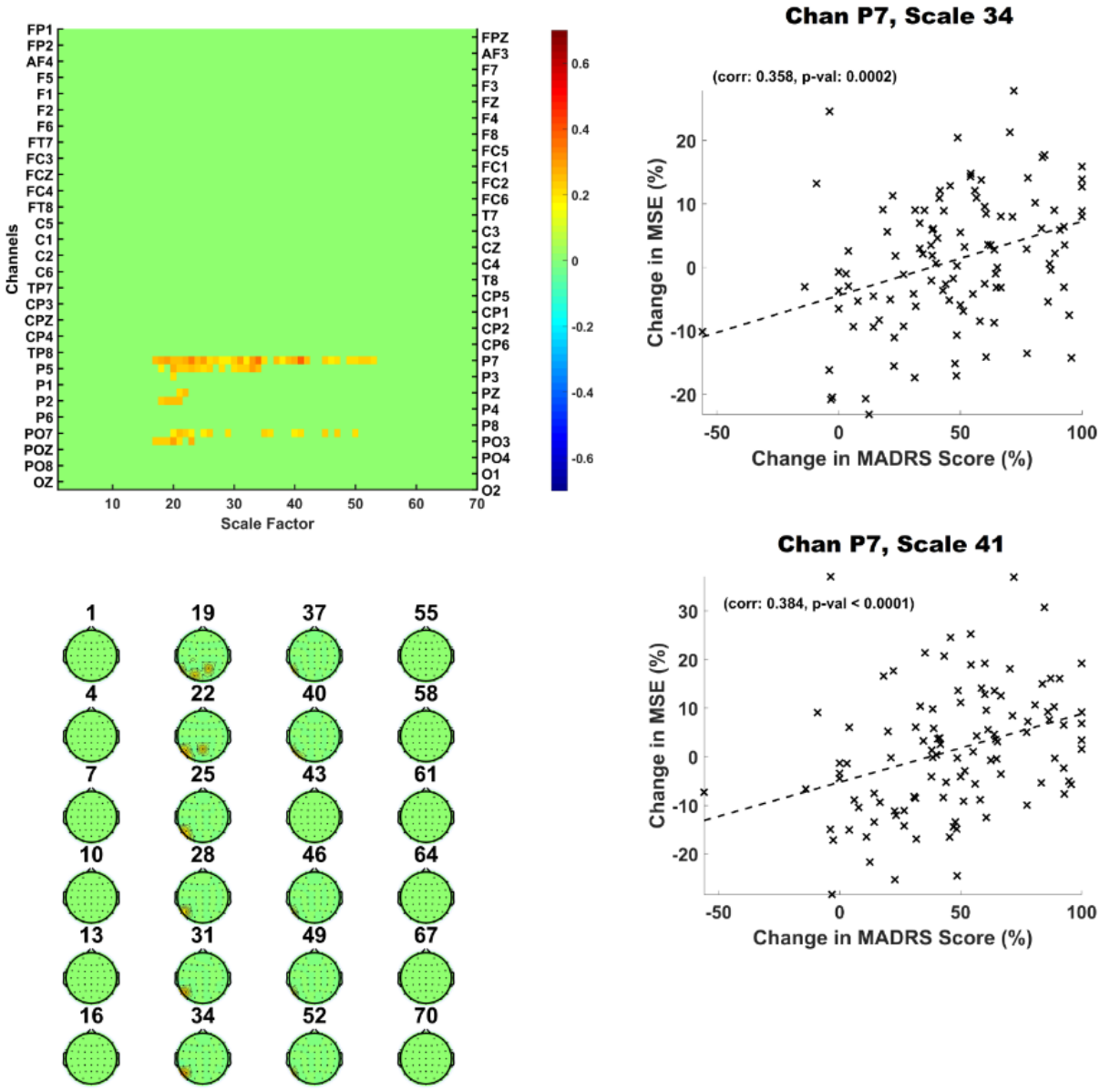


Figure 6.1 – Association between Modulation of Complexity (Baseline to Week 8) and Improvement in Mood

Left Top. Image shows significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and MSE in 107 patients receiving escitalopram (blue: negative correlation; red: positive correlation). Cluster-based correction for multiple comparisons was performed. **Left Bottom.** Topographies illustrate all the significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and MSE. **Right.** Scatter plots highlight the time-scale and region-specific association between percent change in MSE (y-axes) and percent change in MADRS (x-axis) with highest correlations.

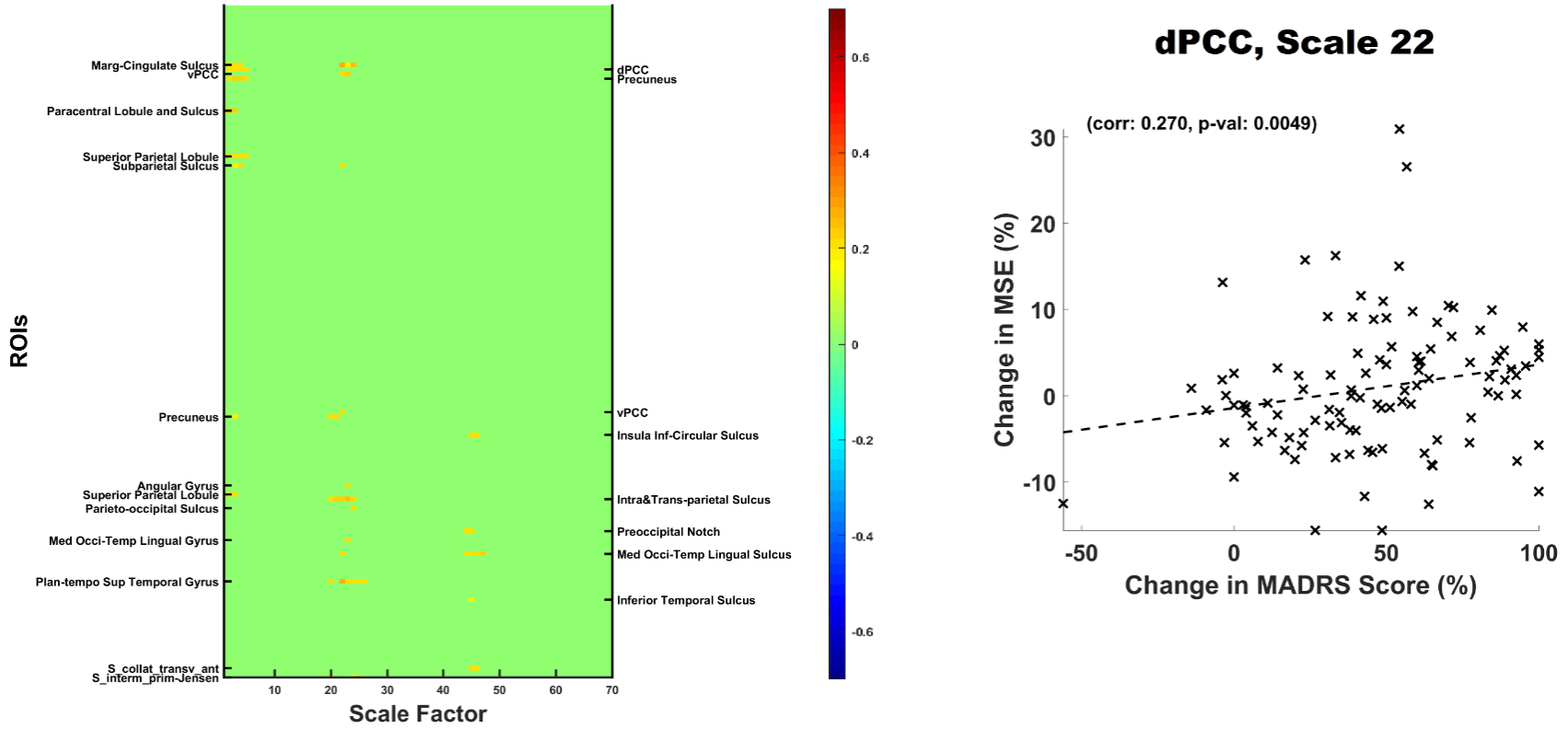


Figure 6.2 – Association between Modulation of Complexity (Baseline to Week 8) and Improvement in Mood in Source Space. **Left.** Image shows significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and MSE in 107 patients receiving escitalopram (blue: negative correlation; red: positive correlation). Cluster-based correction for multiple comparisons was performed. **Right.** Scatter plot highlights the time-scale and region-specific association between percent change in MSE (y-axes) and percent change in MADRS (x-axis) with highest correlation.

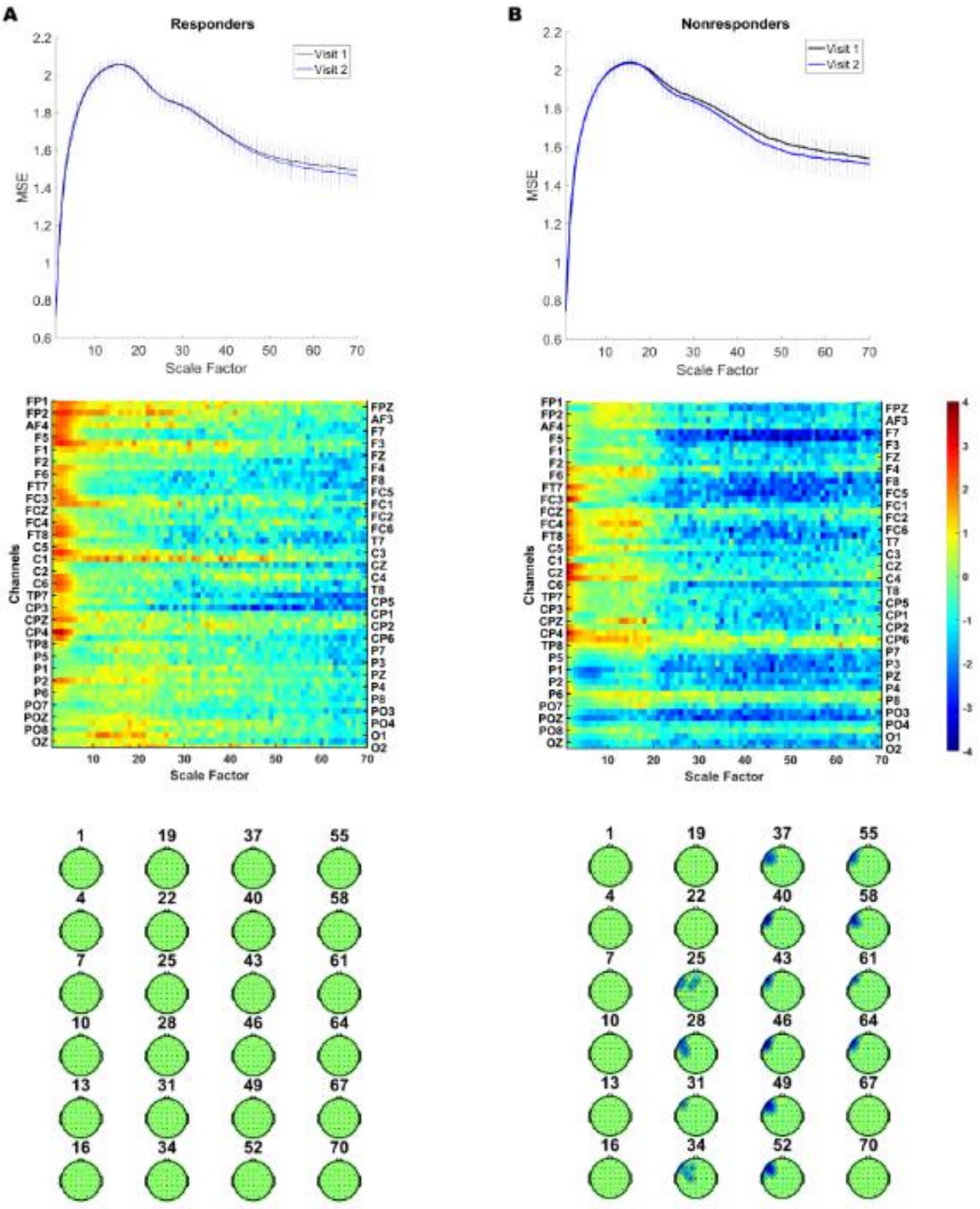

A
Responders
Visit 1
Visit 2
MSE
Scale Factor
B
Nonresponders
Visit 1
Visit 2
MSE
Scale Factor
Channels
Scale Factor
Channels
Scale Factor

Figure 6.3 - Effect of Escitalopram on Complexity of Neural Dynamics (Baseline to Week 2).
Top. Waveforms depict multiscale entropy (MSE) at baseline (black line) and 2 weeks (blue line) into escitalopram treatment in responders **(A)** and non-responders **(B)**. The lines represent the average MSE (y-axes) across electrodes (dots) for time-scales 1 to 70 (x-axes). **Middle.** Images show the original post-hoc test statistics comparing MSE post to pre-treatment across all electrodes (1 to 58) and all time-scales (1 to 70) (blue: decreases; red: increases following treatment) for responders **(A)** and non-responders **(B)** to escitalopram. **Bottom.** Each topography reflects the significant t-maps following correction for multiple comparison, using cluster-based non-parametric permutation test, depicting only the significant clusters $p<0.05$ and setting to 0 non-significant pixels.

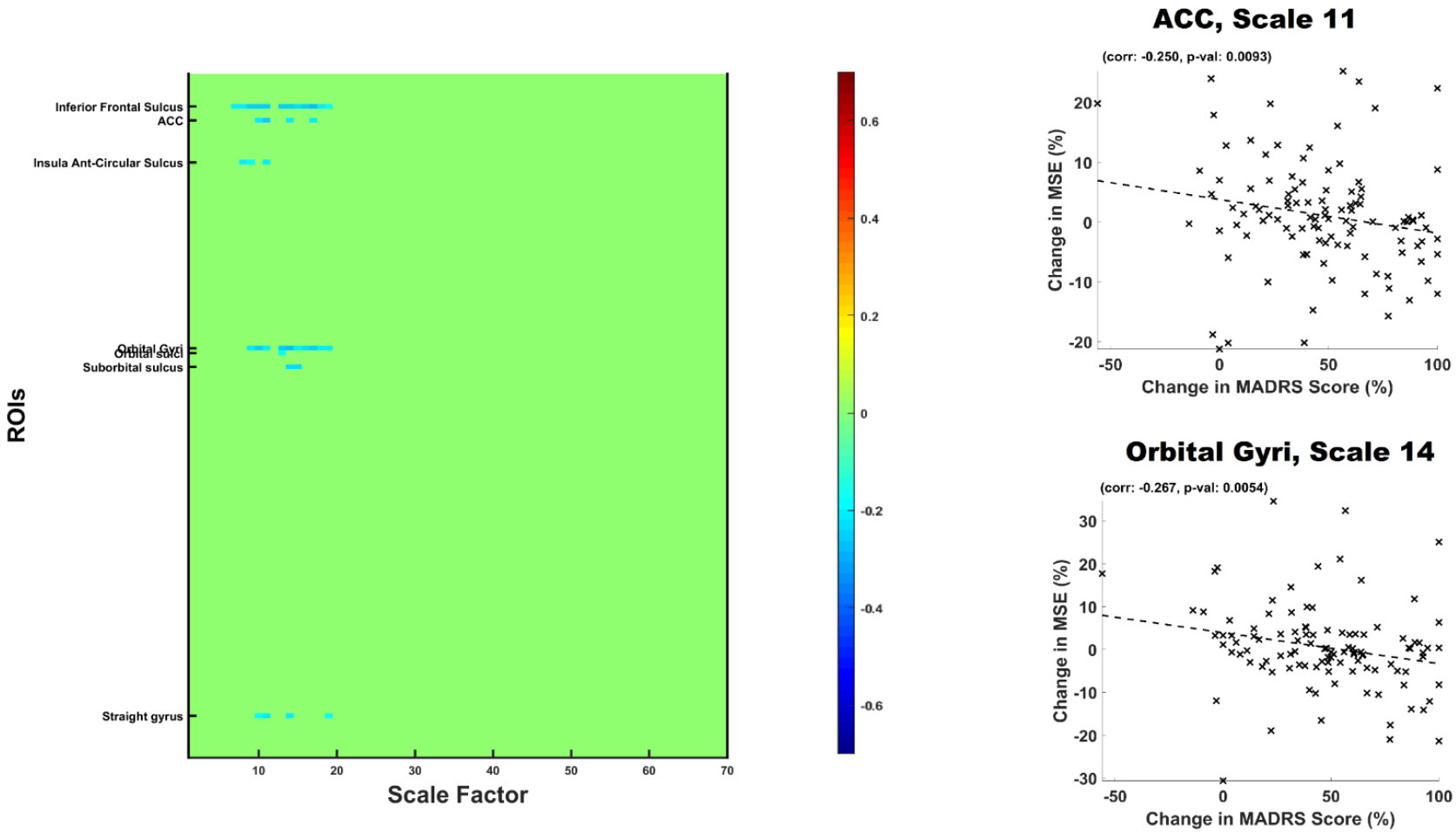


Figure 6.4 – Association between Modulation of Complexity (Baseline to Week 2) and Improvement in Mood in Source Space. **Left.** Image shows significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and MSE in 107 patients receiving escitalopram (blue: negative correlation; red: positive correlation). Cluster-based correction for multiple comparisons was performed. **Right.** Scatter plot highlights the time-scale and region-specific association between percent change in MSE (y-axes) and percent change in MADRS (x-axis) with highest correlations.

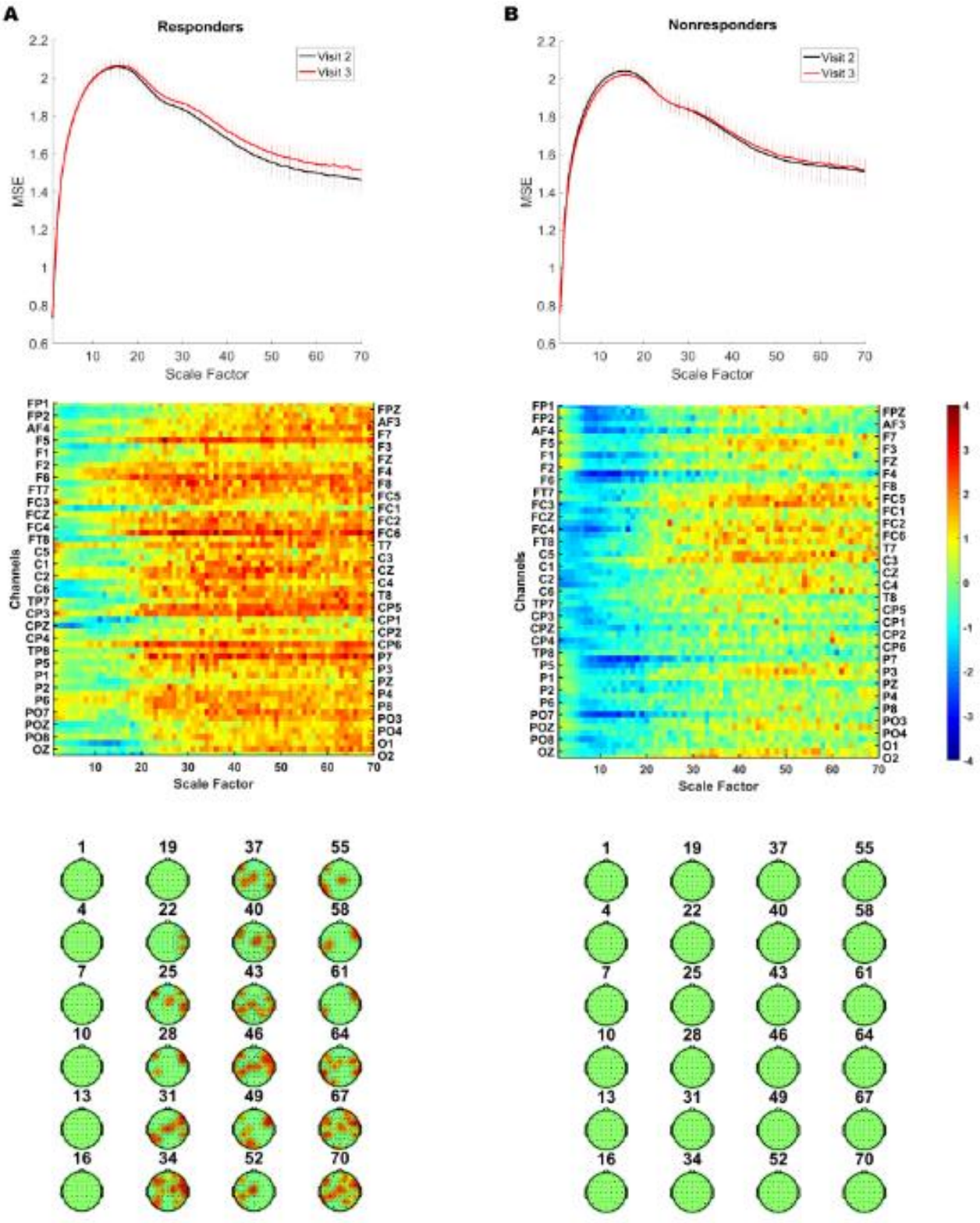

A
Responders
Visit 2
Visit 3
MSE
Scale Factor
B
Nonresponders
Visit 2
Visit 3
MSE
Scale Factor
Channels
Scale Factor
Channels
Scale Factor

Figure 6.5 – Effect of Escitalopram on Complexity of Neural Dynamics (Week 2 to Week 8).
Top. Waveforms depict multiscale entropy (MSE) at week 2 (blue lines) and 8 weeks (red lines) into escitalopram treatment in responders **(A)** and non-responders **(B)**. The lines represent the average MSE (y-axes) across electrodes (dots) for time-scales 1 to 70 (x-axes). **Middle.** Images show the original post-hoc test statistics comparing MSE post to pre-treatment across all electrodes (1 to 58) and all time-scales (1 to 70) (blue: decreases; red: increases following treatment) for responders **(A)** and non-responders **(B)** to escitalopram. **Bottom.** Each topography reflects the significant t-maps following correction for multiple comparison, using cluster-based non-parametric permutation test, depicting only the significant clusters $p<0.05$ and setting to 0 non-significant pixels.

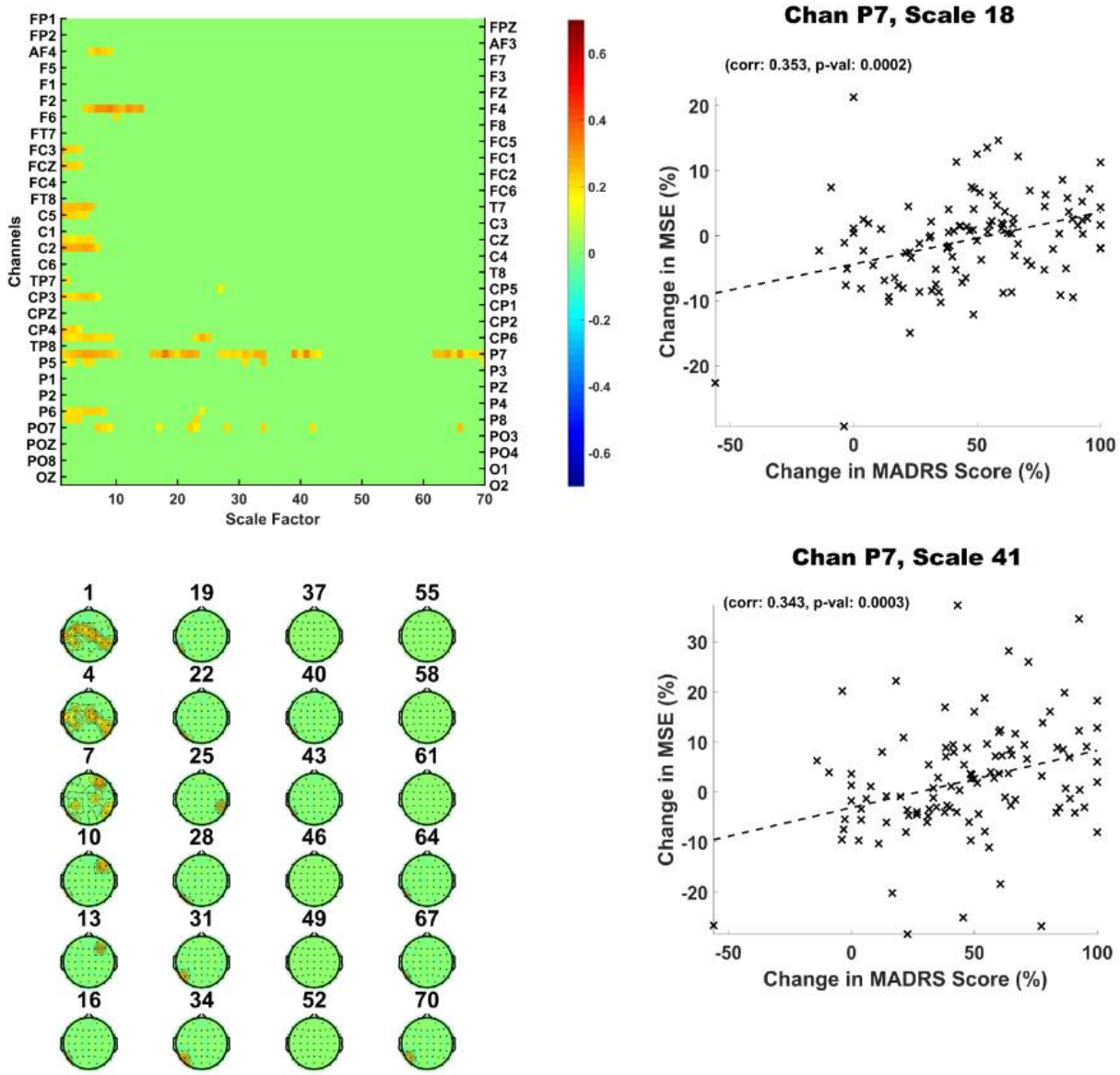


Figure 6.6 – Association between Modulation of Complexity (Week 2 to Week 8) and Improvement in Mood.

Left Top. Image shows significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and MSE in 107 patients receiving escitalopram (blue: negative correlation; red: positive correlation). Cluster-based correction for multiple comparisons was performed. **Left Bottom.** Topographies illustrate all the significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and MSE. **Right.** Scatter plots highlight the time-scale and region-specific association between percent change in MSE (y-axes) and percent change in MADRS (x-axis) with highest correlations.

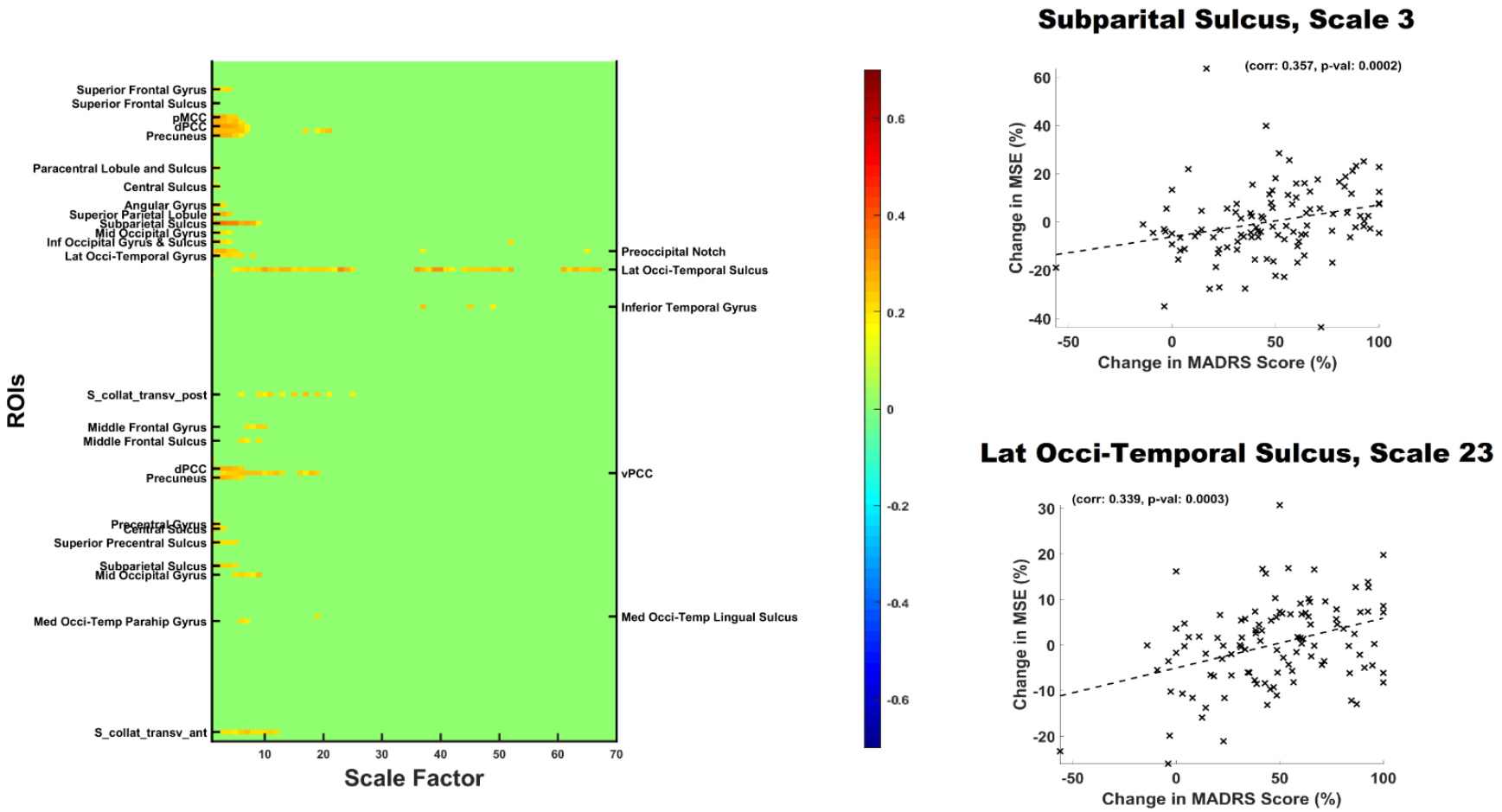


Figure 6.7 – Association between Modulation of Complexity (Week 2 to Week 8) and Improvement in Mood in Source Space. **Left.** Image shows significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and MSE in 107 patients receiving escitalopram (blue: negative correlation; red: positive correlation). Cluster-based correction for multiple comparisons was performed. **Right.** Scatter plot highlights the time-scale and region-specific association between percent change in MSE (y-axes) and percent change in MADRS (x-axis) with highest correlations.

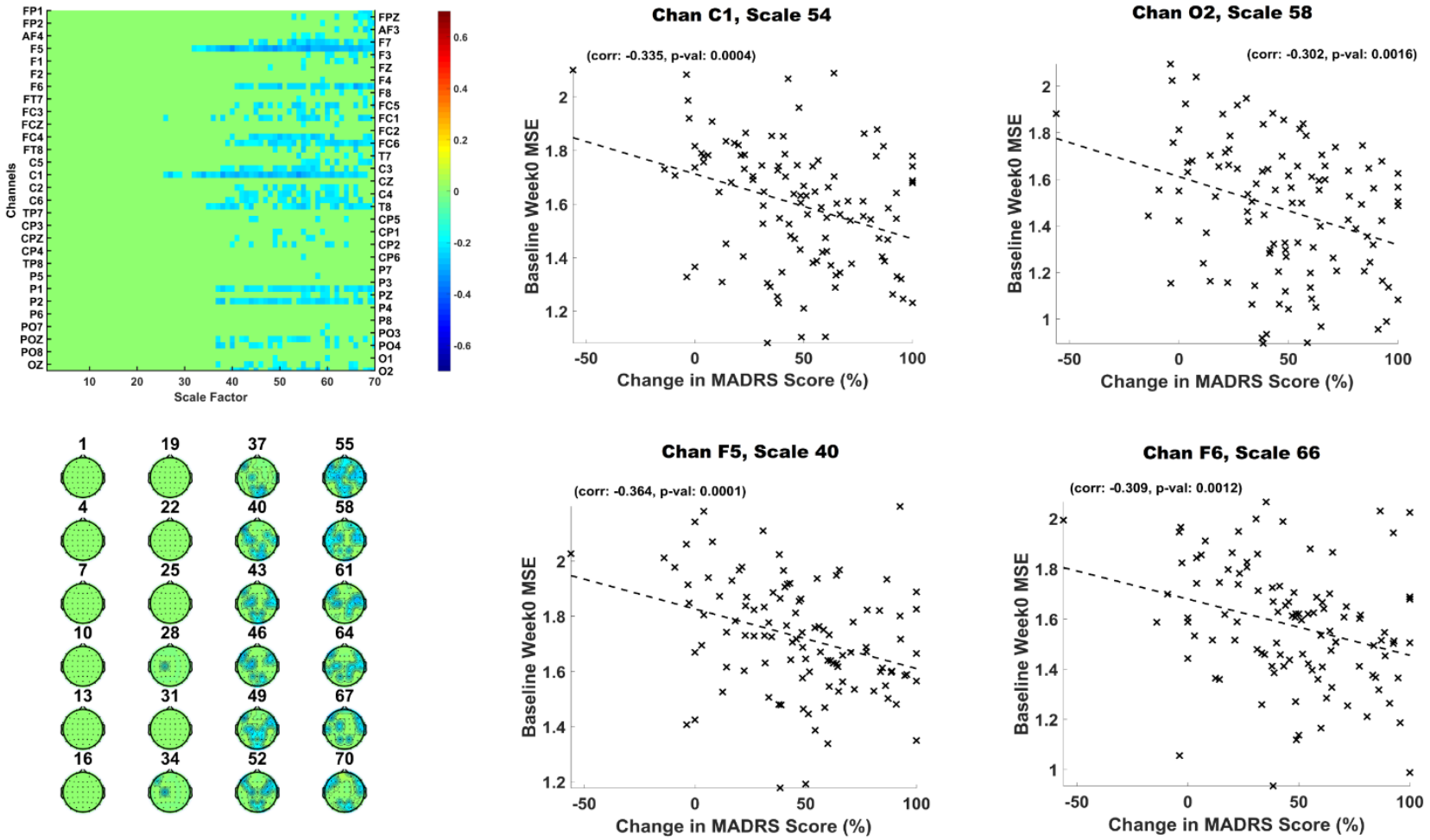


Figure 6.8 – Association between Baseline Complexity and Improvement in Mood.
Left. Image shows significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and baseline MSE in 107 patients receiving escitalopram (blue: negative correlation; red: positive correlation). Cluster-based correction for multiple comparisons was performed. **Right.** Scatter plots highlight the time-scale and region-specific association between percent baseline MSE (y-axes) and percent change in MADRS (x-axis) with highest correlations.

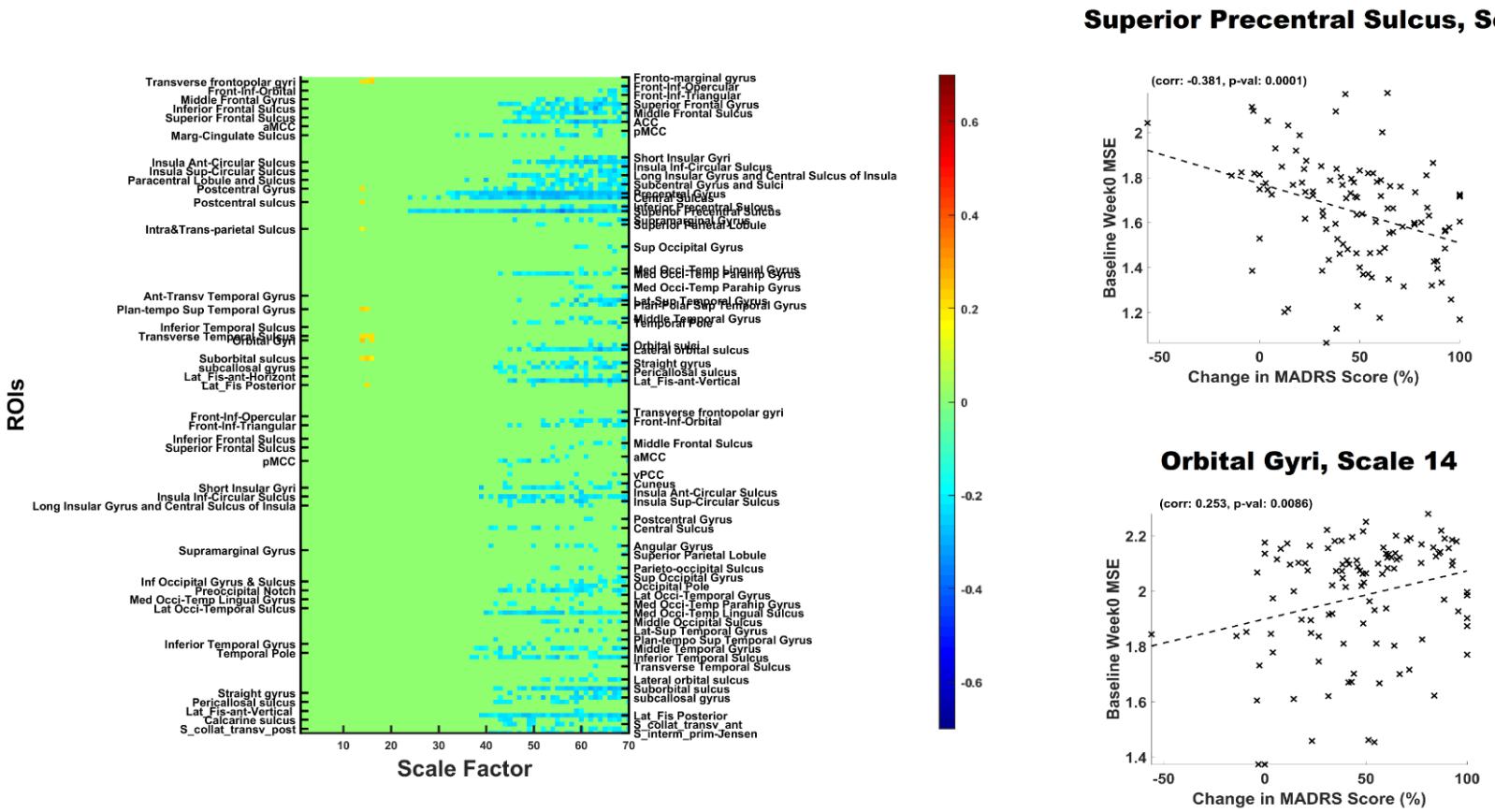


Figure 6.9 – Association between Baseline Complexity and Improvement in Mood in Source Space.
Left. Image shows significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and baseline MSE in 107 patients receiving escitalopram (blue: negative correlation; red: positive correlation). Cluster-based correction for multiple comparisons was performed. **Right.** Scatter plots highlight the time-scale and region-specific association between baseline MSE (y-axes) and percent change in MADRS (x-axis) with highest correlations.

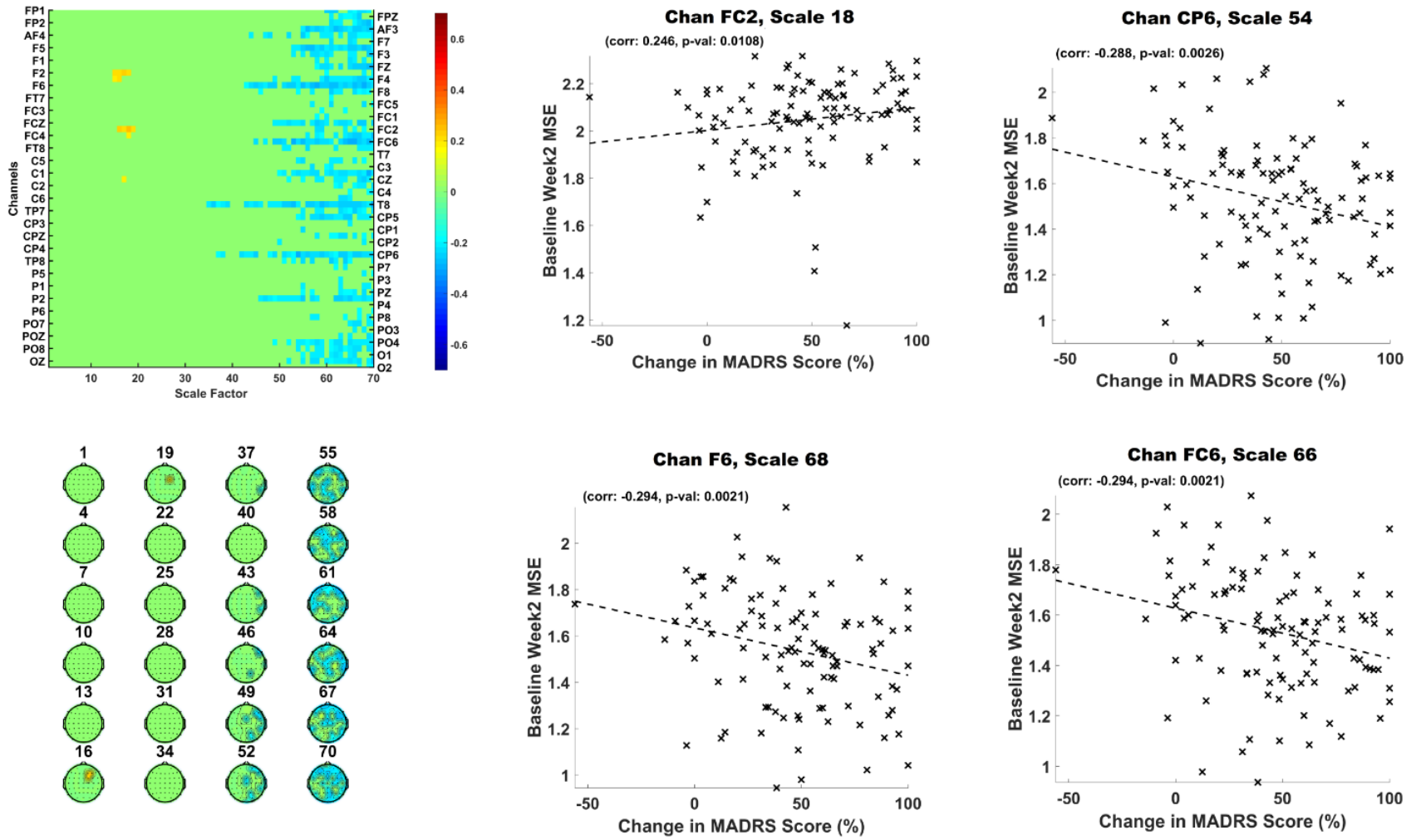
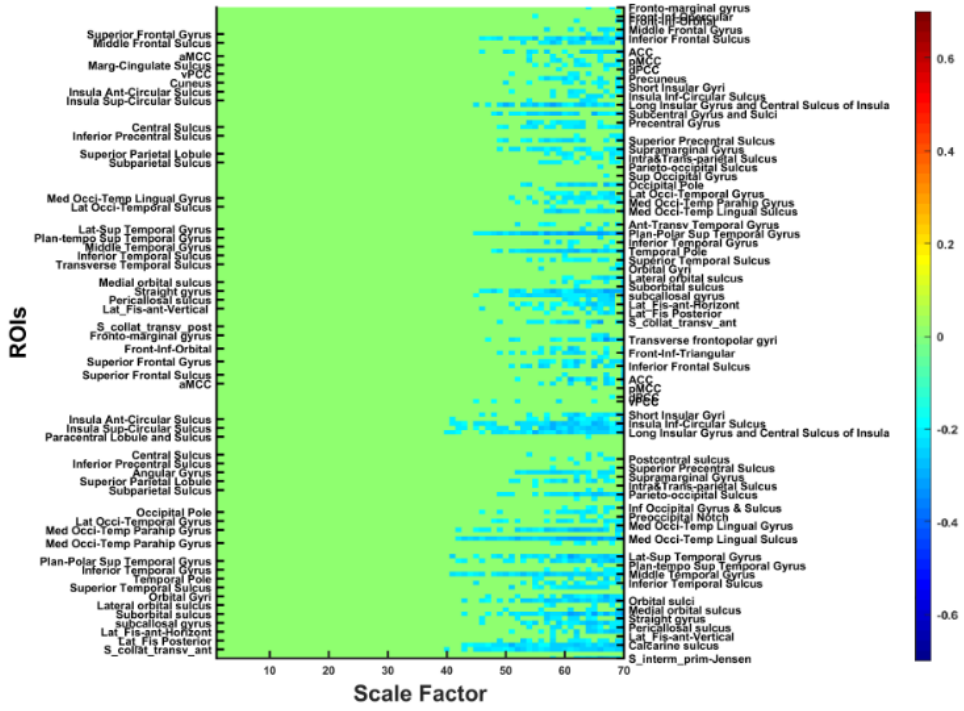
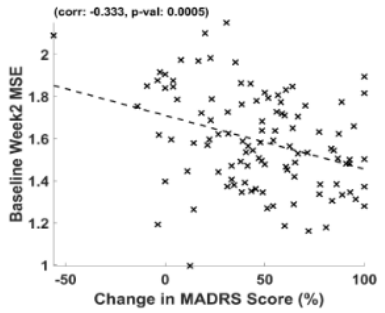


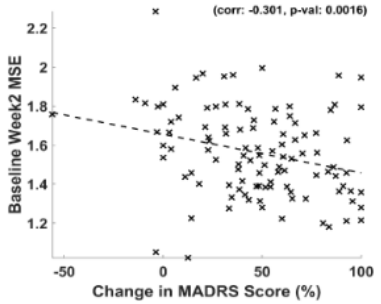
Figure 6.10 – Association between Complexity at Week 2 and Improvement in Mood.
Left. Image shows significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and MSE (at Week 2) in 107 patients receiving escitalopram (blue: negative correlation; red: positive correlation). Cluster-based correction for multiple comparisons was performed. **Right.** Scatter plots highlight the time-scale and region-specific association between MSE at Week 2 (y-axes) and percent change in MADRS (x-axis) revealing highest correlations.



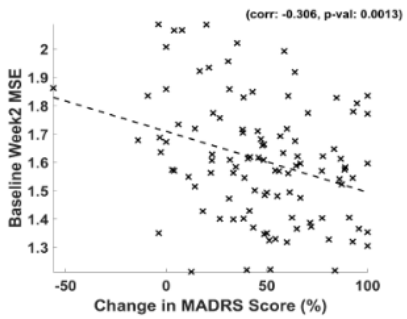
Plan-Polar Sup Temporal Gyrus, Scale 61



Temporal Pole, Scale 65



ACC, Scale 61



Med Occi-Temp_Lingual Sulcus, Scale 66

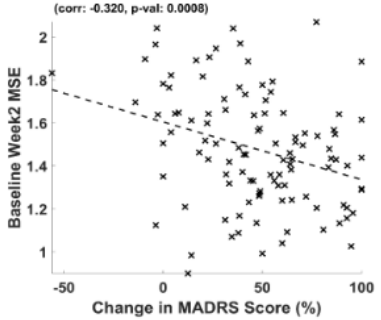


Figure 6.11 - Association between Complexity at Week 2 and Improvement in Mood in Source Space.
Left. Image shows significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and MSE (at Week 2) in 107 patients receiving escitalopram (blue: negative correlation; red: positive correlation). Cluster-based correction for multiple comparisons was performed. **Right.** Scatter plots highlight the time-scale and region-specific association between MSE at Week 2 (y-axes) and percent change in MADRS (x-axis) revealing highest correlations.

6.9 Supplementary Material

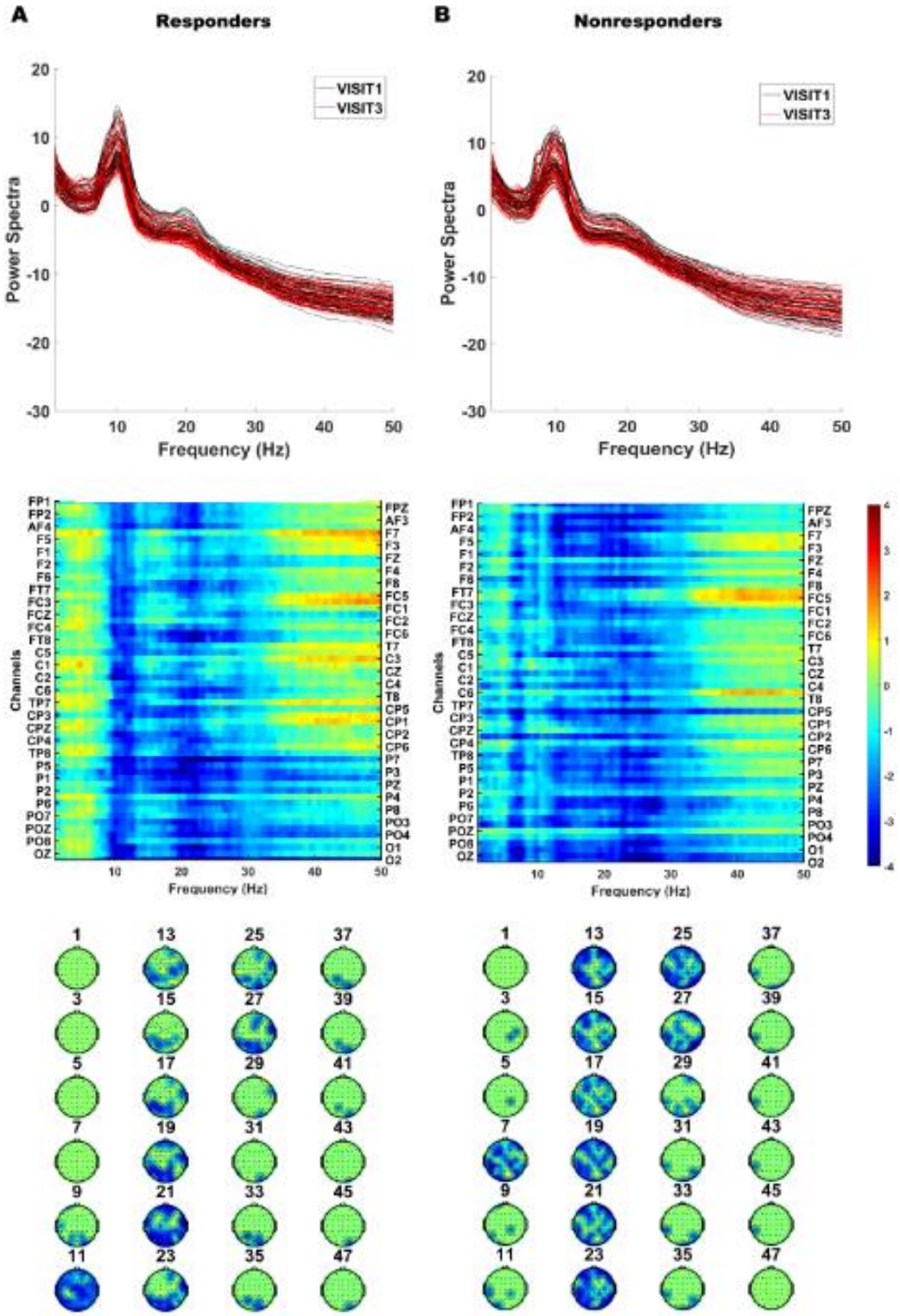


Figure S 6.1 – Effect of Escitalopram on Cortical Oscillations (Baseline to Week 8).
Top. Waveforms depict the absolute power spectrum of resting-state eyes-closed EEG at baseline (black waveforms) and week 8 (red waveforms) into escitalopram treatment in responders **(A)** and non-responders **(B)**. The x-axes are frequency in Hz and the y-axes the relative power in dB. **Middle.** Images show the original post-hoc test statistics maps comparing the relative power across frequency bands (x-axes) and channels (y-axes) post compared to pre-treatment (blue: decreases; red: increases following treatment) for responders and non-responders. **Bottom.** Each topography reflects the significant t-map depicting only the significant clusters $p<0.05$, setting to 0 non-significant pixels.

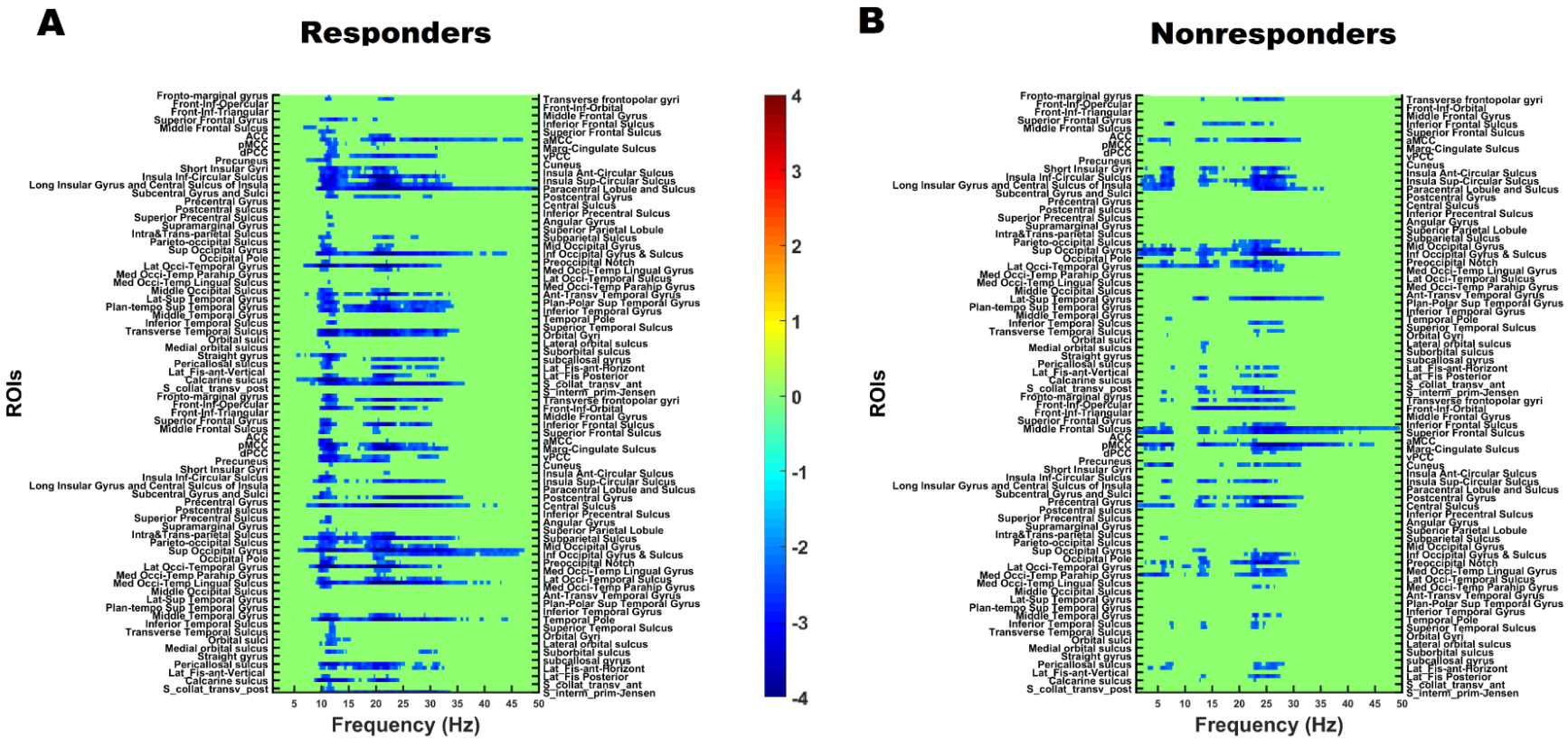


Figure S 6.2 – Effect of Escitalopram on Cortical Oscillations in Source Space (Baseline to Week 8).
In all images, x-axis represents the frequency (1 to 50) in Hertz and y-axis represents Regions of Interest (ROIs) of the Destrieux Atlas (1 to 148). The ROIs are grouped into brain regions in the left (L: the upper half the images) and right (R: the lower half of the images) hemisphere separated by the horizontal black line. Images show the post-hoc independent sample t-test statistics following cluster-based permutation test correction for multiple comparison, depicting only the significant clusters $p<0.05$, labeling only the significant corresponding ROIs, and setting to 0 non-significant pixels. **A.** Image shows the t-test statistics comparing the change in power in responders of escitalopram (red: increase; blue: reduction). **B.** Image shows the t-test statistics comparing the change in power in non-responders of escitalopram (red: increase; blue: reduction).

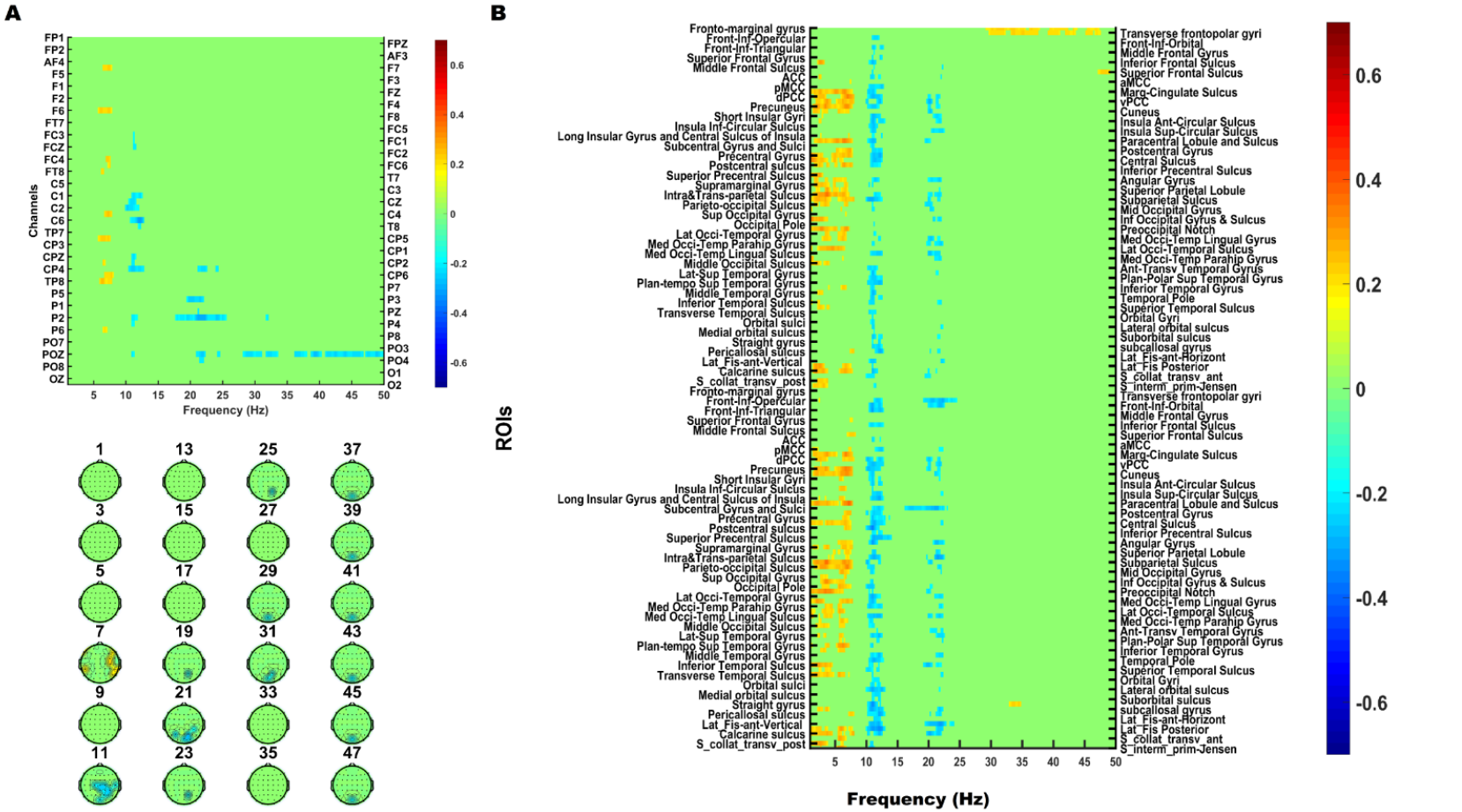


Figure S 6.3 – Association between Modulation of Cortical Oscillations (Baseline to Week 8) and Improvement in Mood in Sensor Space (A) and Source Space (B).
(A) Top. Image shows significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and relative power in 107 patients receiving escitalopram (blue: negative correlation; red: positive correlation). Cluster-based correction for multiple comparisons was performed. **Bottom.** Topographies illustrate all the significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and relative power. **(B)** Image shows significant correlation coefficients in source.

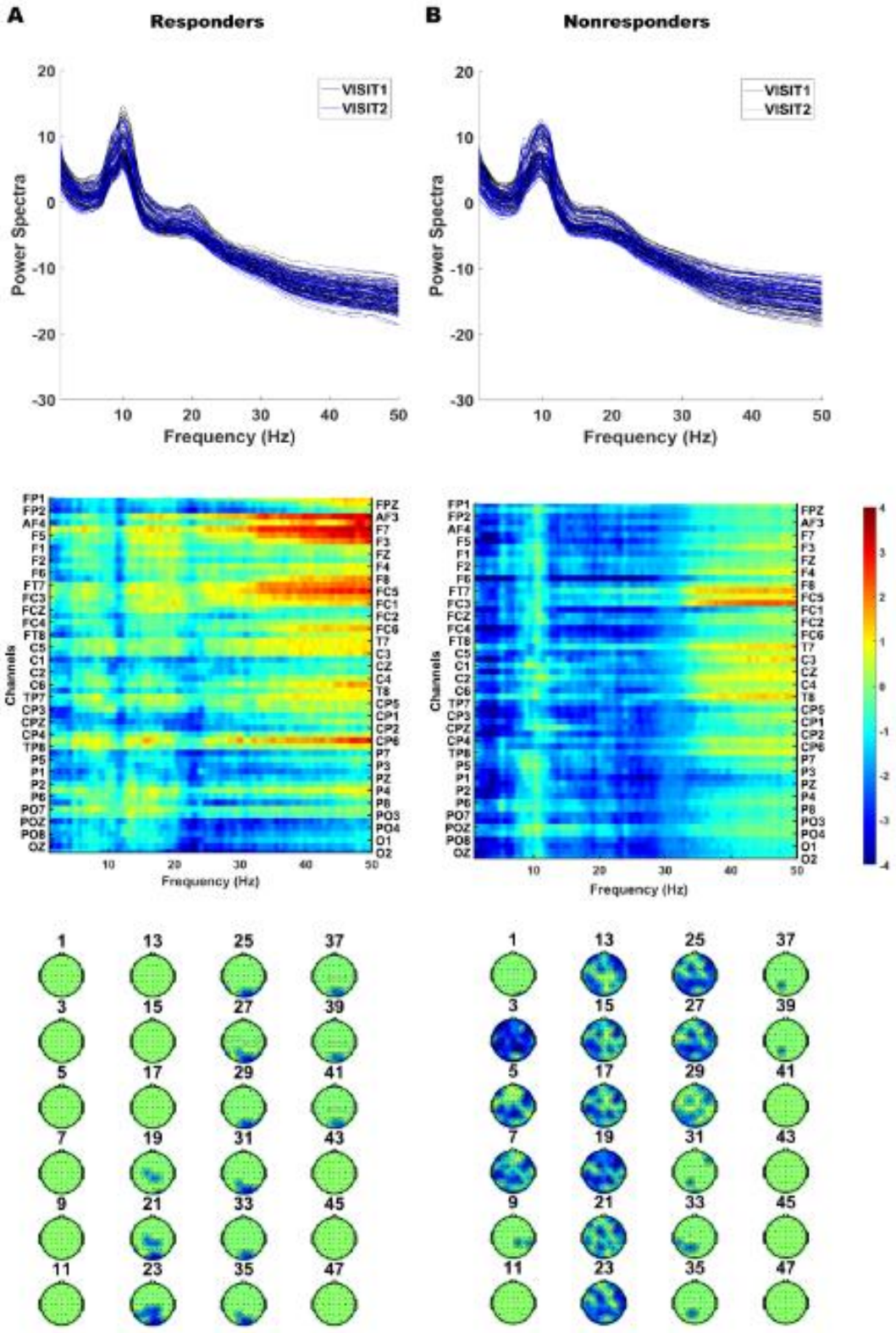

A
Responders
B
Nonresponders
VISIT1
VISIT2
Power Spectra
Frequency (Hz)
Channels

Figure S 6.4 – Effect of Escitalopram on Cortical Oscillations (Week 2 to Week 8).
Top. Waveforms depict the absolute power spectrum of resting-state eyes-closed EEG at week 2 (blue waveforms) and 8 weeks (red waveforms) into escitalopram treatment in responders **(A)** and non-responders **(B)**. The x-axes are frequency in Hz and the y-axes the relative power in dB. **Middle.** Images show the original post-hoc test statistics maps comparing the relative power across frequency bands (x-axes) and channels (y-axes) post compared to pre-treatment (blue: decreases; red: increases following treatment) for responders **(A)** and non-responders **(B)**. **Bottom.** Each topography reflects the significant t-map depicting only the significant clusters $p<0.05$, setting to 0 non-significant pixels.

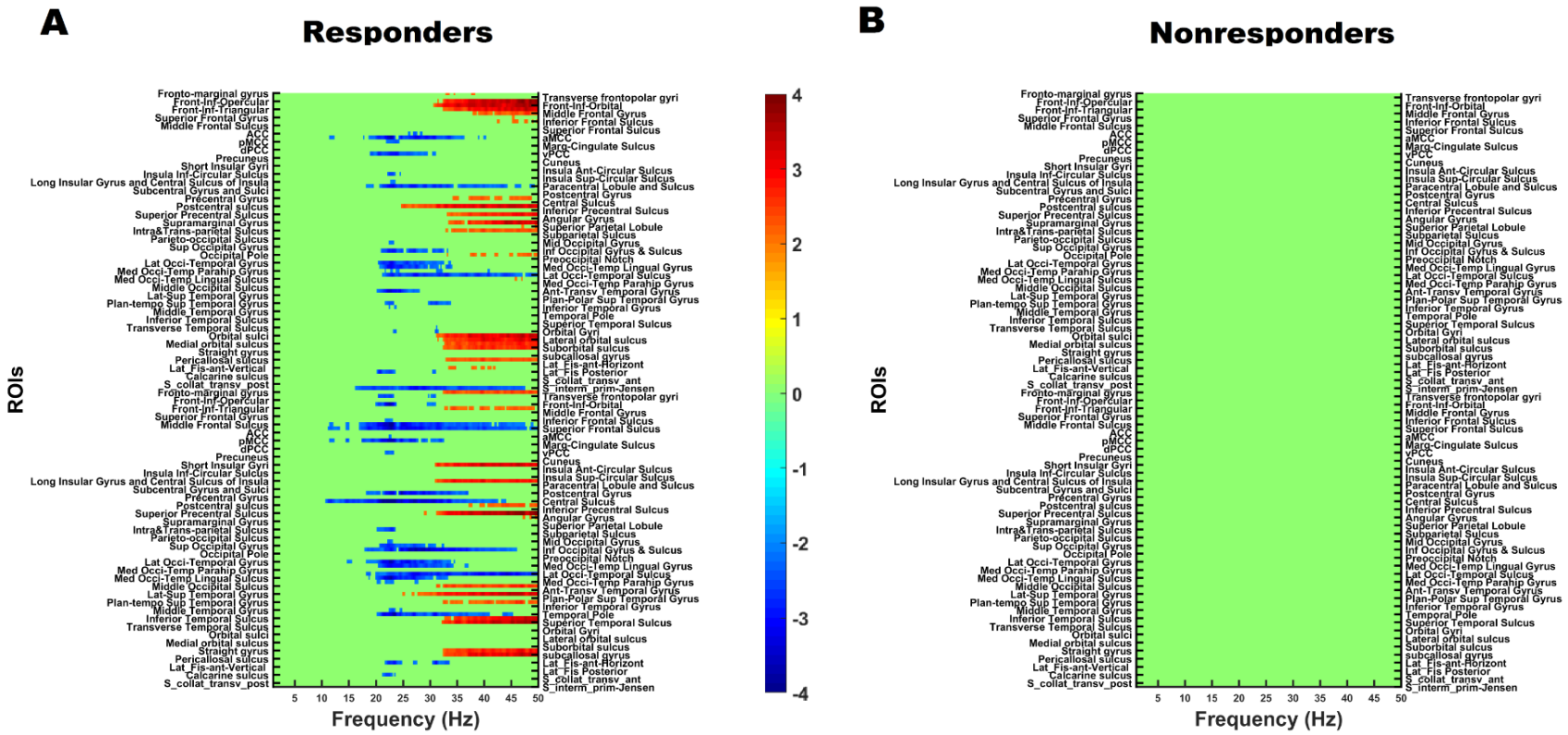


Figure S 6.5 – Effect of Escitalopram on Cortical Oscillations in Source Space (Baseline to Week 2).
In all images, x-axis represents the frequency (1 to 50) in Hertz and y-axis represents Regions of Interest (ROIs) of the Destrieux Atlas (1 to 148). The ROIs are grouped into brain regions in the left (L: the upper half the images) and right (R: the lower half of the images) hemisphere separated by the horizontal black line. Images show the post-hoc independent sample t-test statistics following cluster-based permutation test correction for multiple comparison, depicting only the significant clusters $p<0.05$, labeling only the significant corresponding ROIs, and setting to 0 non-significant pixels. **A.** Image shows the t-test statistics comparing the change in power in responders of escitalopram (red: increase; blue: reduction). **B.** Image shows the t-test statistics comparing the change in power in non-responders of escitalopram (red: increase; blue: reduction).

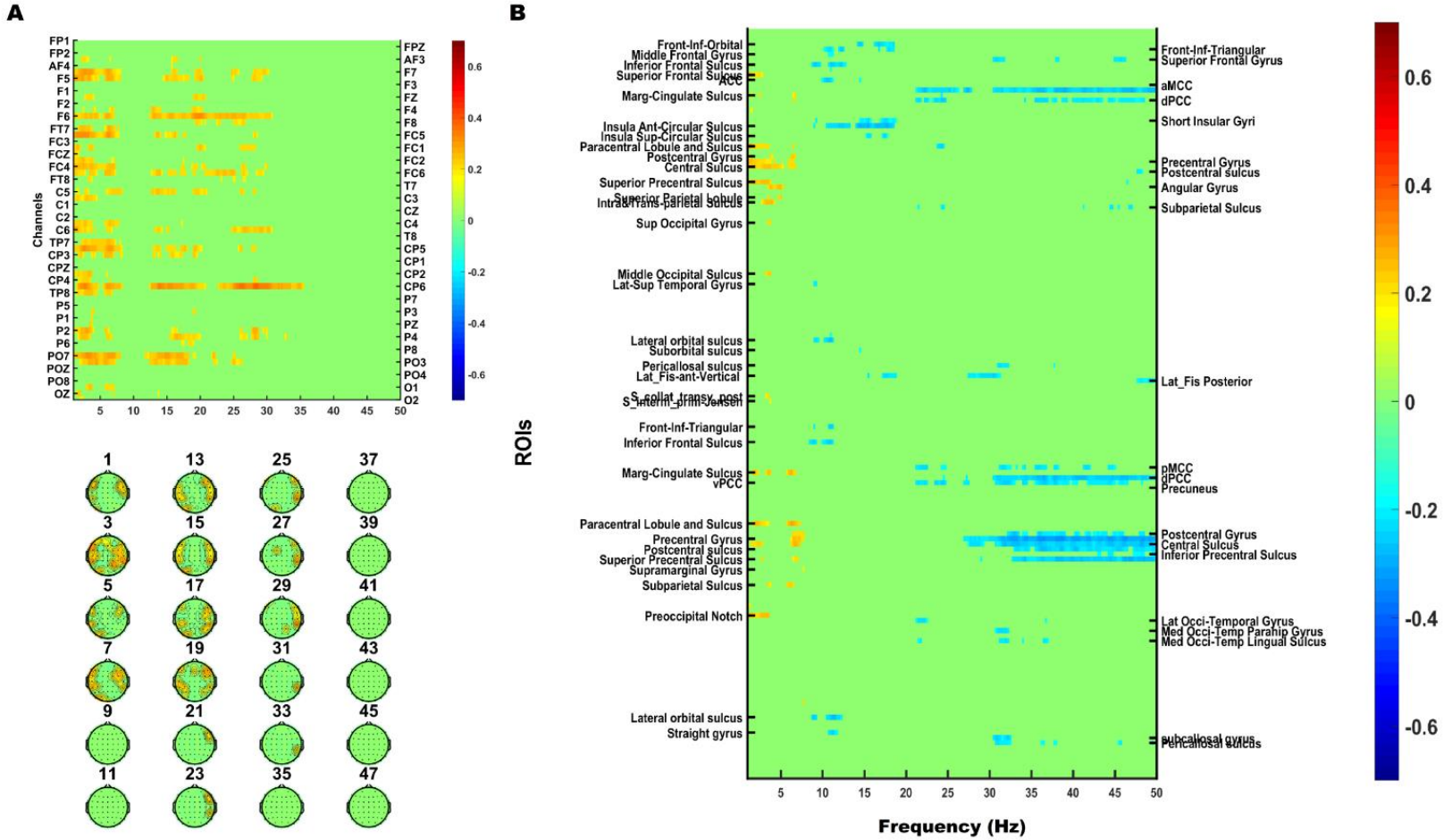


Figure S 6.6 – Association between Modulation of Cortical Oscillations (Baseline to Week 2) and Improvement in Mood in Sensor Space (A) and Source Space (B).

(A) Top. Image shows significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and relative power in 107 patients receiving escitalopram (blue: negative correlation; red: positive correlation). Cluster-based correction for multiple comparisons was performed. **Bottom.** Topographies illustrate all the significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and relative power. **(B)** Image shows significant correlation coefficients in source.

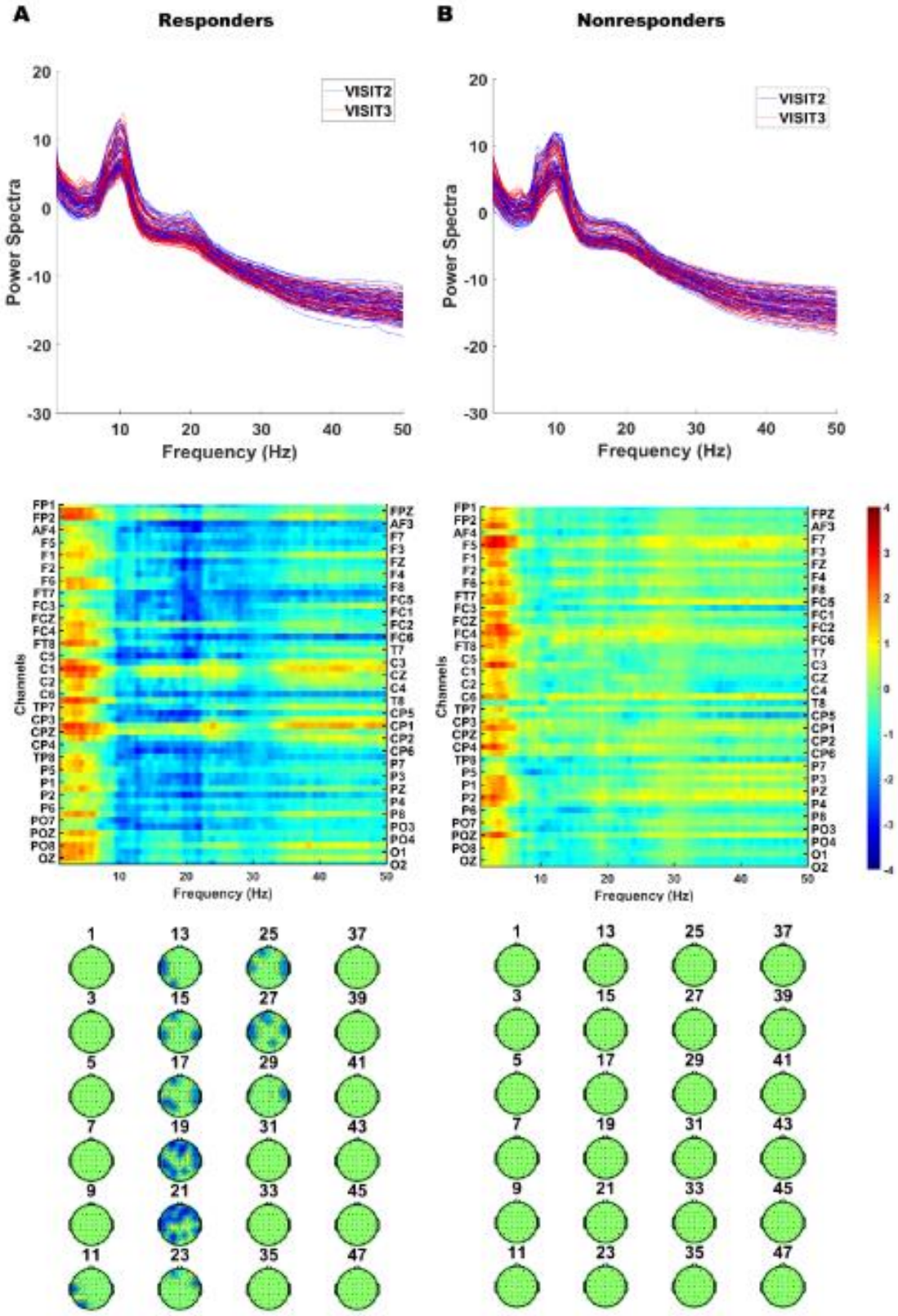

A
Responders
B
Nonresponders
VISIT2
VISIT3
Power Spectra
Frequency (Hz)
Channels

Figure S 6.7 – Effect of Escitalopram on Cortical Oscillations (Week 2 to Week 8).
Top. Waveforms depict the absolute power spectrum of resting-state eyes-closed EEG at week 2 (blue waveforms) and 8 weeks (red waveforms) into escitalopram treatment in responders **(A)** and non-responders **(B)**. The x-axes are frequency in Hz and the y-axes the relative power in dB. **Middle.** Images show the original post-hoc test statistics maps comparing the relative power across frequency bands (x-axes) and channels (y-axes) post compared to pre-treatment (blue: decreases; red: increases following treatment) for responders **(A)** and non-responders **(B)**. **Bottom.** Each topography reflects the significant t-map depicting only the significant clusters $p<0.05$, setting to 0 non-significant pixels.

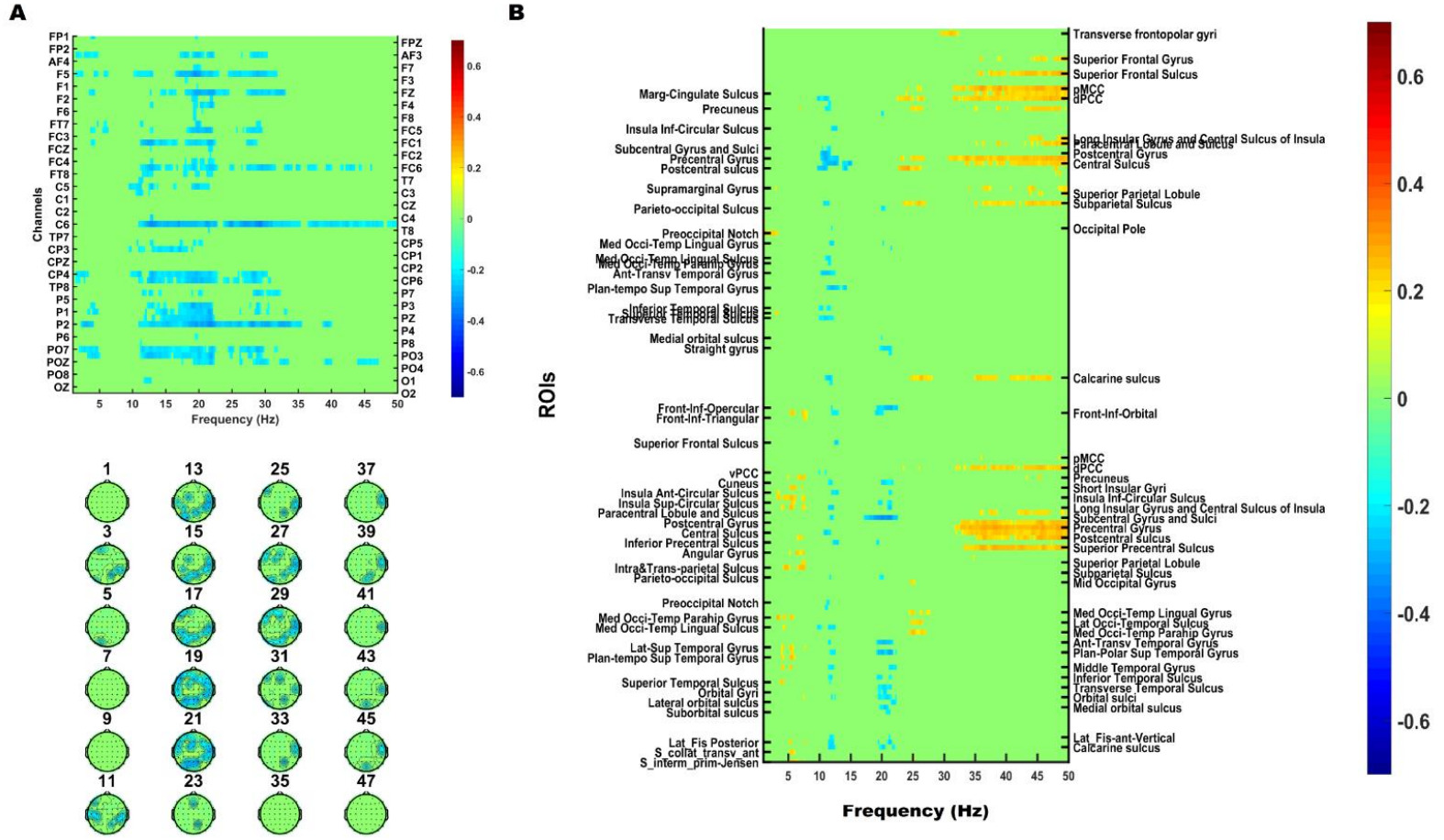


Figure S 6.8 – Association between Modulation of Cortical Oscillations (Week 2 to Week 8) and Improvement in Mood in Sensor (A) and Source Space (B).

(A) Top. Image shows significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and relative power in 107 patients receiving escitalopram (blue: negative correlation; red: positive correlation). Cluster-based correction for multiple comparisons was performed. **Bottom.** Topographies illustrate all the significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and relative power. **(B)** Image shows significant correlation coefficients in source.

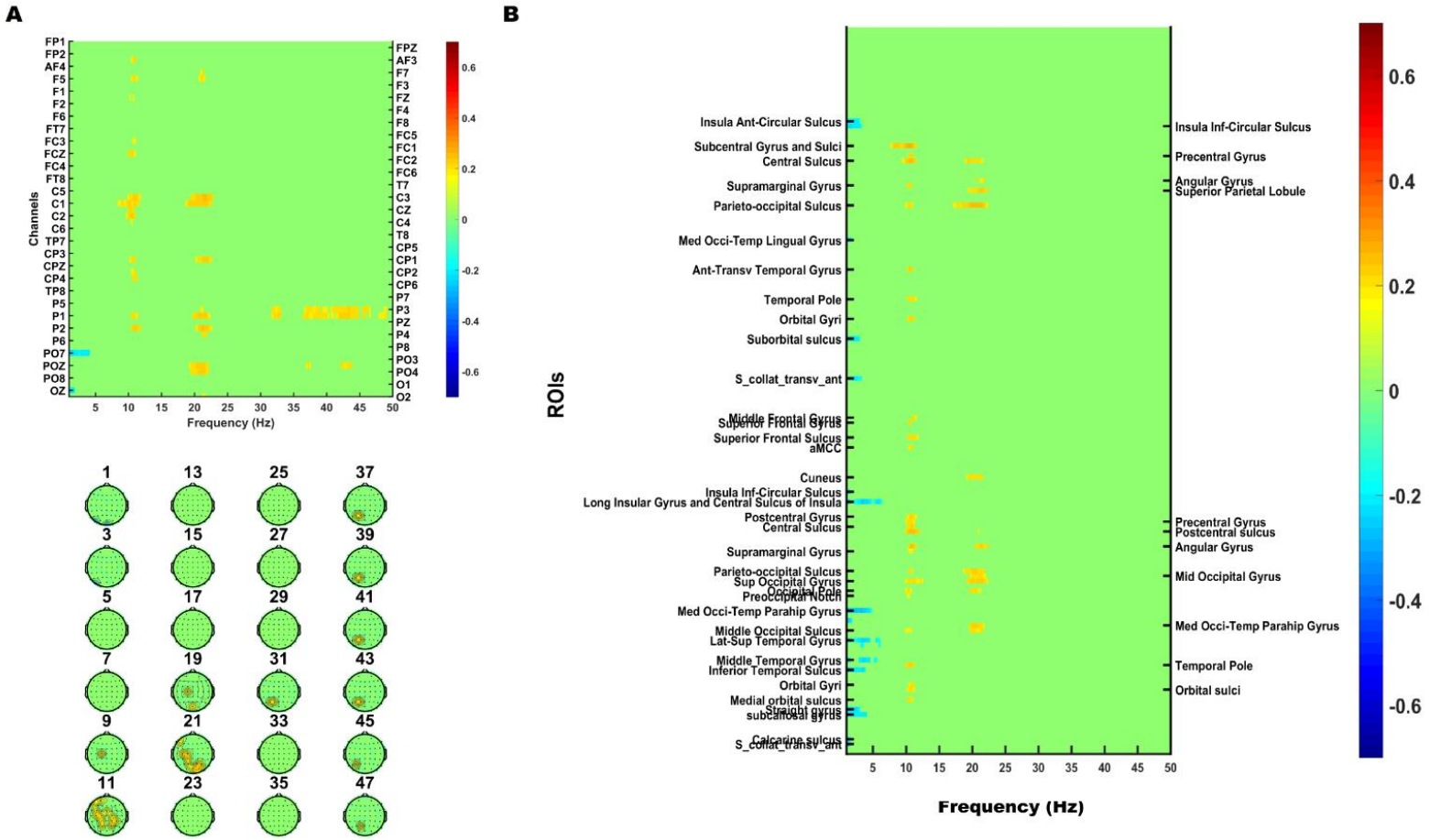


Figure S 6.9 – Association between Baseline Cortical Oscillations and Improvement in Mood in Sensor (A) and Source Space (B).

(A) Top. Image shows significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and baseline relative power in 107 patients receiving escitalopram (blue: negative correlation; red: positive correlation). Cluster-based correction for multiple comparisons was performed. **Bottom.** Topographies illustrate all the significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and baseline relative power. **(B)** Image shows significant correlation coefficients in source.

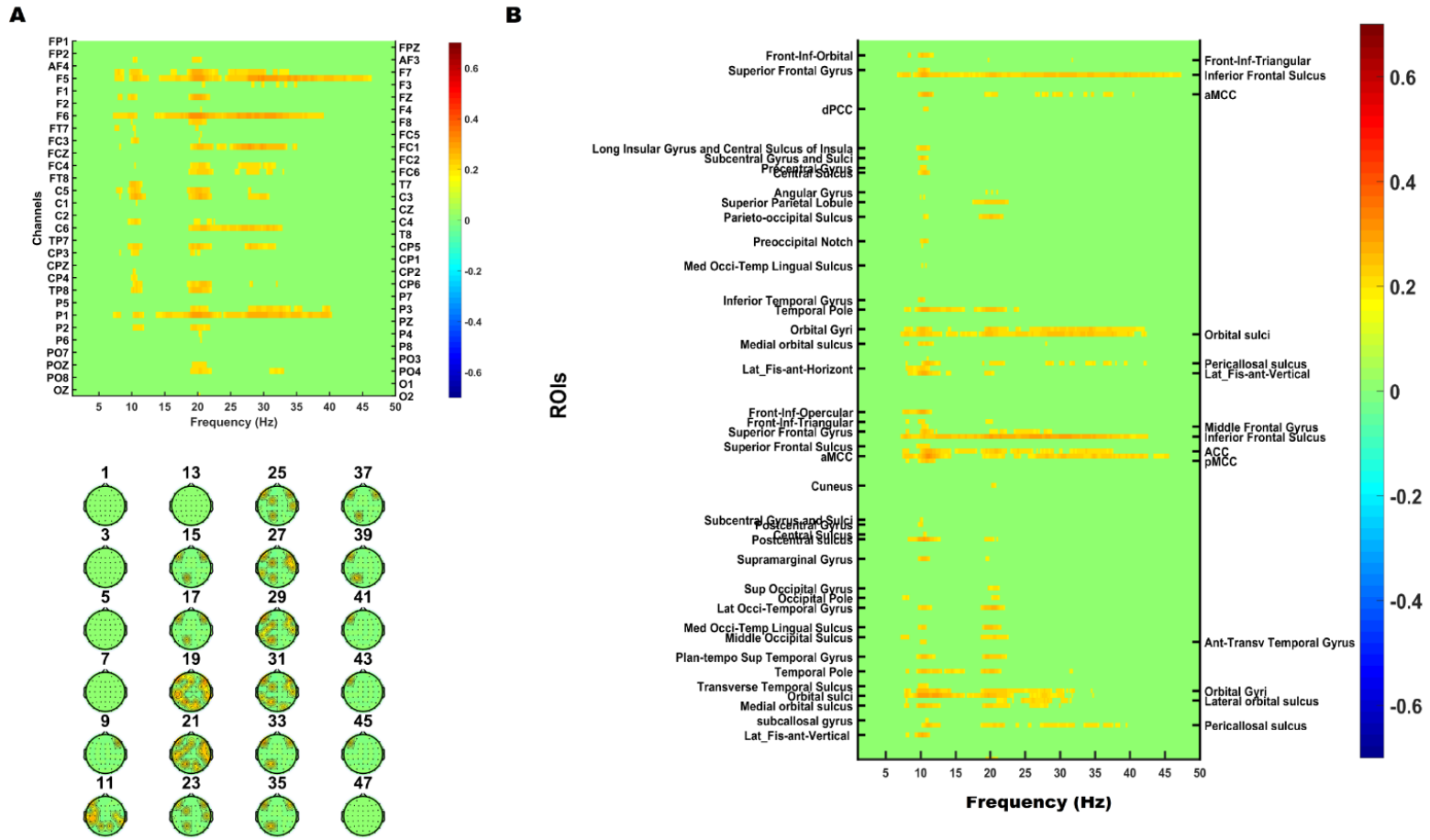


Figure S 6.10 – Association between Cortical Oscillations (at Week 2) and Improvement in Mood in Sensor (A) and Source Space (B).

(A) Top. Image shows significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and relative power at week 2 in 107 patients receiving escitalopram (blue: negative correlation; red: positive correlation). Cluster-based correction for multiple comparisons was performed. **Bottom.** Topographies illustrate all the significant (original $p < 0.05$) spearman correlation coefficients (rho) between percent change in MADRS and relative power at week 2. **(B)** Image shows significant correlation coefficients in source.

Chapter 7 – Characterizing the Modulation of Global-Network Dynamics during Escitalopram Treatment in Major Depressive Disorder: A CAN-BIND-1 Study

In this chapter, we apply EEG microstate analysis to extract global brain-network measures of neural oscillations. We aim to identify whether these measures can provide insight into targets of escitalopram for successful treatment outcome. Introduction section is omitted as it is a repeat of introduction from chapters 5 and 6.

7.1 Methods

Patients and data. In the CAN-BIND-1 study, 211 participants aged 18-60 years who met the DSM-IV requirements for major depressive disorder were recruited and completed the baseline visit. Of this group, 180 patients received 8 weeks of standardized escitalopram treatment (10-20mg). At the 8 week visit, responder or non-responder status was determined as ≥ 50% decrease in Montgomery–Åsberg Depression Rating Scale (MADRS) score from baseline. There were 85 responders and 95 non-responders at week 8 (see (Kennedy, Lam et al. *In Press,* 2018) for further details). Detailed descriptions on the clinical data, research protocol and data acquisition have been published previously (Lam, Milev et al. 2016, Baskaran, Farzan et al. 2017, Farzan, Atluri et al. 2017, Kennedy, Lam et al. *In Press*, 2018). In a subset of patients that were recruited at four sites participating in EEG acquisition, eight minutes of resting-state, eyes-closed EEG data was collected. The participating sites included: University of British Columbia (UBC), Toronto General Hospital (TGH), Queens University (QNS), and the Centre for Addiction and Mental Health (CAMH). The demographic and clinical characteristics are presented in **Table 6.1**.

EEG data was collected at baseline (within 3 days before the start of the treatment trial), at the end of week 2 (i.e., two weeks after the beginning of the trial) and at the end of week 8 (i.e., eight weeks after the beginning of the trial). Of the 180 patients, 124 patients were recruited to participate in EEG assessments. Fifteen patients were excluded because data from one or more of the visits was missing. Two patients were excluded because data from at least one or more visits were noisy. Therefore, this study included resting-state, eyes-closed EEG data collected from 107 patients at baseline, at the end of week 2, and at the end of week 8. Detailed descriptions on the clinical data, research protocol and data acquisition have been published previously (Lam, Milev et al. 2016, Baskaran, Farzan et al. 2017, Farzan, Atluri et al. 2017) and are summarized in Chapters 3, 6 and 8 of this thesis.

Treatment. Escitalopram was administered in an open-label manner, starting at 10 mg daily, which was increased to 20 mg daily at week 2 or later if clinically necessary. For patients who were unable to tolerate the 20 mg dose, the dose could be reduced to 10 mg at the discretion of

the treating psychiatrist. Participants were clinically assessed every 2 weeks throughout the study period (8 weeks) including baseline (before administration of study medication). The primary outcome measure was the change in Montgomery-Åsberg Depression Rating Scale (MADRS) from baseline to week 8 of the study. Response was defined as a ≥50% decrease in MADRS score.

Inter-Site Data Harmonization. All EEG datasets were standardized to the following parameters: 58 electrodes common to all sites, 0.05-100Hz bandpass filter, Cz reference, 512Hz sampling rate. The EEG files were then exported as an EEGLAB (Delorme and Makeig 2004) dataset. Data was standardized using MATLAB R2012b/MATLAB R2016a with the EEGLAB toolbox (v12.0.2.6b). Complete descriptions on the standardization of EEG data across sites in the CANBIND study were recently published (Farzan, Atluri et al. 2017).

Data Preprocessing. Data was analyzed at the Temerty Centre for Therapeutic Brain Stimulation at CAMH. During preprocessing, EEG data was divided into 2-second epochs, bandpass-filtered between 1-80Hz, and notch-filtered at 60Hz. With the removal of eye electrodes and other unused channels, the total number of EEG channels used for analysis was 58. Using EEGLAB, independent component analysis was used to extract eye, muscle and electrode artifacts. Deleted EEG channels were interpolated using spherical spline interpolation (Perrin, Pernier et al. 1989) and data was re-referenced to an average reference. This preprocessing pipeline is currently made available as ERPEEG (http://www.tmseeg.com/multisiteprojects/).

Microstate Analysis. Microstate analysis followed the standard procedure outlined in seminal work (Lehmann, Ozaki et al. 1987, Pascual-Marqui, Michel et al. 1995) and was implemented using CARTOOL (Brunet, Murray et al. 2011). Prior to the application of microstate analysis, 4 minutes of the pre-processed EEG data was bandpass-filtered from 1-30Hz. The topographical maps at the local maxima peaks of the global field power curve are clustered to derive the four prototypical microstate classes (Koenig, Prichep et al. 2002). In this study, the topographical atomize–agglomerate hierarchical clustering algorithm (Tibshirani and Walther 2005) was applied to cluster the data for each visit (baseline or week 2) separately into four states

(microstate maps). Finally, topographical maps at each local maxima point of the global field power curve were assigned to the microstate class of highest correlation using Pearson's spatial product-moment correlation coefficient (Brandeis, Naylor et al. 1992). Three features were calculated for each of the four microstate classes: (i) average duration, (ii) frequency, and (iii) coverage. Average duration is the amount of time a microstate class remains stable when it appears, in milliseconds; frequency refers to the occurrence of each microstate class per second; and coverage is the percent of recording covered by each microstate class.

Statistical Analysis. To examine the effect of treatment on microstate characteristics following escitalopram, a 2x3x4 repeated-measures ANOVA (Lehmann, Faber et al. 2005, Tomescu, Rihs et al. 2014) was conducted for each microstate feature (Duration, Frequency and Coverage) with *RESPONSE* (Responder, Non-responder; Response: ≥50% improvement in MADRS) as a categorical factor, and *TIME* (Baseline, Week 2, Week 8) and *MICROSTATE CLASS* (A, B, C, D) as the repeated-measures factors. *Age* was included as a covariate.

Based on our hypothesis that seizure therapy modulates global neural dynamics, planned comparisons were performed to determine whether changes in microstate characteristics were associated with treatment response. For each of the three microstate characteristics (duration, frequency, coverage), paired t-tests were performed to compare the characteristic before and after treatment for each of the four states. The results were corrected for multiple comparisons using the Bonferroni correction method (for the 4 microstates).

A significance level of $\alpha<0.05$ was used for all statistical tests. Pairwise post-hoc comparisons were performed using Tukey-HSD. All planned comparisons were corrected using the Bonferroni method (4 comparisons for the 4 microstate classes).

Correlation and Predictive Analysis. Associations between baseline (week 0 or week 2) microstate characteristics or an early change (change from week 0 to week 2) in microstate characteristics with clinical assessments (change in MADRS score) were evaluated with multiple regression models. The four microstates (A, B, C and D) and age were included in the model.

Change in microstate characteristics were calculated as (Post-Pre)/Pre*100 where a higher percentage represents an increase in the microstate characteristic. Change in MADRS was calculated as (Pre-Post)/Pre*100 where a higher percentage represents improvement in depressive symptoms.

7.2 Results

A significant change in MADRS score was observed following 8 weeks of escitalopram treatment (t=13.8; *df*=105; *p*<0.0001; Cohen's *d*=2.69).

7.2.1 Effect of Escitalopram on EEG Microstate Characteristics

7.2.1.1 Average Duration

A main effect of *AGE* (F=5.87; *df*=1,98; *p*=0.02; η_p^2=0.057), *TIME* (F=4.87; *df*=2,196; *p*=0.009; η_p^2=0.05), *MAPS* (F=4.09; *df*=3,294; *p*=0.007; η_p^2=0.04), and an interaction effect of *TIME x MAPS* (F=5.41; *df*=6,588; *p*<0.0001; η_p^2=0.05), *MAPS x SITE* (F=6.61; *df*=9,294; *p*<0.0001; η_p^2=0.17), and *TIME x MAPS x SITE* (F=3.78; *df*=18,588; *p*<0.0001; η_p^2=0.10) was seen in the **duration** of microstates. An effect of RESPONSE (≥50% improvement in MADRS) was not observed (F=0.37; *df*=1,98; *p*=0.55; η_p^2=0.004). Post-hoc results are summarized in **Table 7.1**. Planned comparisons were conducted to investigate the effect of response based on our hypotheses. Significant results from planned comparisons are summarized in **Table 7.2**. All significant results are also shown in **Figure 7.1**.

7.2.1.2 Frequency

A main effect of *AGE* (F=5.64; *df*=1,98; *p*=0.02; η_p^2=0.05), *TIME* (F=5.36; *df*=2,196; *p*=0.005; η_p^2=0.05), *MAPS* (F=4.11; *df*=3,294; *p*=0.007; η_p^2=0.04), and an interaction effect of *TIME x MAPS* (F=5.36; *df*=6,588; *p*<0.0001; η_p^2=0.05), *TIME x SITE* (F=2.19; *df*=6,196; *p*=0.04; η_p^2=0.06), *MAPS x SITE* (F=4.77; *df*=9,294; *p*<0.0001; η_p^2=0.13), and *TIME x MAPS x SITE* (F=4.4.; *df*=18,588; *p*<0.0001; η_p^2=0.12) was seen in the **duration** of microstates. An effect of RESPONSE (≥50% improvement in MADRS) was not observed (F=0.08; *df*=1,98; *p*=0.77;

η_p^2=0.0009). Post-hoc results are summarized in **Table 7.3**. Planned comparisons were conducted to investigate the effect of response based on our hypotheses. Significant results from planned comparisons are summarized in **Table 7.4**. All significant results are also shown in **Figure 7.2**.

7.2.1.3 Coverage

A main effect of *AGE* (F=5.33; *df*=1,98; *p*=0.02; η_p^2=0.05), *MAPS* (F=4.30; *df*=3,294; *p*=0.005; η_p^2=0.04), and an interaction effect of *TIME x MAPS* (F=5.45; *df*=6,588; *p*<0.0001; η_p^2=0.05), *MAPS x SITE* (F=6.67; *df*=9,294; *p*<0.0001; η_p^2=0.17) and *TIME x MAPS x SITE* (F=4.36; *df*=18,588; *p*<0.0001; η_p^2=0.12) was seen in the **duration** of microstates. An effect of RESPONSE (≥50% improvement in MADRS) was not observed (F=1.67; *df*=1,98; *p*=0.99; η_p^2=0.01). Post-hoc results are summarized in **Table 7.5**. Planned comparisons were conducted to investigate the effect of response based on our hypotheses. Significant results from planned comparisons are summarized in **Table 7.6**. All significant results are also shown in **Figure 7.3**.

7.2.2 Correlation and Prediction Analysis Results

Change in microstate characteristics from baseline to week 8 did not correlate with improvement in depressive symptoms.

Microstate characteristics at baseline (i.e., week 0) did not correlate with improvement in depressive symptoms. However, at week 2, a long duration and a high frequency of State B were correlated with improvement in depressive symptoms (**Table 7.7**). In addition, an early increase in coverage of State B from baseline to week 2 correlated with improvement in depressive symptoms (**Table 7.7**).

7.3 Discussion

Significant changes in microstate characteristics were observed in both responders and non-responders of escitalopram following 8 weeks of treatment. However, the trajectory of these

changes was observed to be different between the two groups. Responders revealed a gradual change in duration, frequency and coverage of microstates over the entire treatment course (i.e., changes in microstate characteristics from baseline to week 2 were usually in the same direction as changes from week 2 to week 8). On the other hand, non-responders showed a significant early change (i.e., from baseline to week 2) in microstate characteristics that were not significant in responders. Specifically, non-responders revealed an early decrease in the duration, frequency and coverage of State B and an early increase in the duration, frequency and coverage of State D. No significant changes in early microstate characteristics were only specific to responders. If early changes were significant, they were significant in both responders and non-responders or only in non-responders. Results from correlation analysis revealed that a long duration and a high frequency of State B at week 2, and an increase in the coverage of State B after 2 weeks of treatment, may predict treatment outcome. In light of these findings, microstates may be a promising marker for the early evaluation of antidepressant efficacy.

With the highest efficacy for treatment-resistant depression, electroconvulsive therapy (ECT) is perhaps the benchmark to which other treatments can be compared to. Although the exact mechanism of action of ECT is unclear, several studies have made hypotheses based on their findings. In our previous work (Chapter 4), state-specific modulation of global brain dynamics was observed in responders of ECT. Specifically, responders revealed an increase in the duration of State A and a decrease in the frequency of States B, C and D suggesting that ECT may impact global network dynamics for therapeutic efficacy. In comparison, results from this study revealed significant changes in the dynamics of all states (from baseline to week 8) following escitalopram treatment. But, these changes were not specific to responders or non-responders suggesting that the mechanism of action of escitalopram may not include the modulation of global network dynamics for therapeutic effect. However, changes in global-network dynamics by escitalopram may have a negative impact on treatment outcome. As mentioned, early changes (baseline to week 2) in global network dynamics following escitalopram were associated with non-response to treatment.

Investigating the predictive value of microstate characteristics, baseline global brain dynamics were not correlated with response to ECT, but significant correlations were observed with

response to escitalopram. Baseline or early changes in the global dynamics of States B and D were specific to non-responders only. In addition, State B characteristics significantly correlated with improvement in depressive symptoms following the 8-week course of escitalopram. This suggests that, unlike seizure therapy, which potentially shows therapeutic effect by impacting the dynamics of neural networks regardless of the initial configuration of the network, the effectiveness of antidepressant medications is highly dependent on the baseline brain state.

Changes in States B, C and D were observed in non-responders of escitalopram. Previous studies have shown that an increase in the occurrence of microstate B was associated with psychotic symptoms (Lehmann, Faber et al. 2005, Irisawa, Isotani et al. 2006, Nishida, Morishima et al. 2013). With recent progress in the integration of fMRI and EEG data, a few studies have also explored the association between cortical microstate activity and resting-state fMRI networks. State B was associated with the parietal and occipital-parietal areas of spatial-visualization and verbalization. In (Britz, Van De Ville et al. 2010) States C and D were associated with the salience and frontoparietal networks argued to be the facilitators of depressive symptoms. Decreased frequency of occurrence of microstate C was also associated with panic disorder (Kikuchi, Koenig et al. 2011). Taken together, changes in microstate characteristics seen in non-responders of escitalopram may indicate worsening of symptoms in these patients. However, further research is needed to elucidate this hypothesis. Future work could investigate the neurophysiological underpinnings of microstates and how they might be associated with response to escitalopram.

7.4 Conclusions

In this chapter, we provided insight into the effects of escitalopram on resting-state global-brain network dynamics through EEG microstate analysis. Results suggested that antidepressant medications have a significant impact on global brain dynamics but they may not impact global brain dynamics for therapeutic effect. However, initial changes in global brain-network dynamics were associated with non-response to escitalopram. Therefore, medications might potentially cause an early perturbance in global-network dynamics that eventually leads to non-

response. Future studies including data at several time points (i.e., 2 days, 1 week, 4 weeks following treatment) would improve our current understanding of antidepressant medications and their mechanism of action at the network level of biological organization. Finally, mapping global neural dynamics and their modulation by several other antidepressant medications is highly recommended. This would provide useful information on the effects of different medications and may even help guide the treatment selection process for patients with depression.

7.5 Tables

Table 7.1 - Average Duration Feature ANOVA Post-Hoc Results (Tukey-HSD)

Interaction Effect	Post-hoc Result		
	Comparison	**State**	**Stats**
Time x Maps	Week 2 vs Week 0	State C	HSD=-9.66; *df*=588; *p*=0.0005; Cohen's *d*=0.53
	Week 2 vs Week 0	State D	HSD=11.6; *df*=588; *p*<0.0001; Cohen's *d*=0.73
	Week 8 vs Week 2	State A	HSD=-6.47; *df*=588; *p*<0.0001; Cohen's *d*=0.31
	Week 8 vs Week 2	State B	HSD=7.20; *df*=588; *p*<0.0001; Cohen's *d*=0.48
	Week 8 vs Week 0	State A	HSD=-8.35; *df*=588; *p*<0.0001; Cohen's *d*=0.56
	Week 8 vs Week 0	State C	HSD=-11.4; *df*=588; *p*=0.03; Cohen's *d*=0.64
	Week 8 vs Week 0	State D	HSD=9.56; *df*=588; *p*=0.02; Cohen's *d*=0.66
Maps x Site	CAM vs QNS	State C	HSD=3.23; *df*=287.8; *p*=0.04; Cohen's *d*=0.52
	CAM vs TGH	State A	HSD=-5.74; *df*=287.8; *p*=0.007; Cohen's *d*=0.92
	TGH vs UBC	State A	HSD=5.24; *df*=287.8; *p*=0.02; Cohen's *d*=0.45
Time x Maps x Site	TGH vs UBC - Week 0	State A	HSD=7.04; *df*=416.1; *p*=0.0005; Cohen's *d*=1.05

Table 7.2 - Average Duration Feature - Planned Comparisons for Response

Time	Map	Group	Planned Comparison Results
Week 2 vs Week 0	State A	Responders	t = -2.42; *df*=50; Bonferroni-corrected *p*=0.08; Cohen's *d*=0.20
		Non-responders	*N.S.*
	State B	Responders	*N.S.*
		Non-responders	t = -2.88; df=57; Bonferroni-corrected *p*=0.02; Cohen's *d*=0.41
Week 8 vs Week 2	State B	Responders	*N.S.*
		Non-responders	t = 5.05; df=57; Bonferroni-corrected *p*<0.0001; Cohen's *d*=0.83
	State C	Responders	*N.S.*
		Non-responders	t = -2.84; df=57; Bonferroni-corrected *p*=0.02; Cohen's *d*=0.24

Table 7.3 - Frequency Feature ANOVA Post-Hoc Results (Tukey-HSD)

Interaction Effect	Post-hoc		
	Comparison	**State**	**Stats**
Site x Time	none		
Time x Maps	Week 2 vs Week 0	State D	HSD=12.9; *df*=588; *p*<0.0001; Cohen's *d*=0.68
	Week 8 vs Week 0	State A	HSD=-8.75; *df*=588; *p*<0.0001; Cohen's *d*=0.49
	Week 8 vs Week 0	State D	HSD=14.4; *df*=588; *p*<0.0001; Cohen's *d*=0.87
	Week 8 vs Week 2	State A	HSD=-6.05; *df*=588; *p*=0.004; Cohen's *d*=0.36
	Week 8 vs Week 2	State B	HSD=9.90; *df*=588; *p*=0.0002; Cohen's *d*=0.65
Maps x Site	CAM vs QNS	State A	HSD=-3.98; *df*=252.8; *p*=0.01; Cohen's *d*=0.64
	TGH vs UBC	State A	HSD=4.04; *df*=252.8; *p*=0.009; Cohen's *d*=0.33
Time x Maps x Site	none		

Table 7.4 - Frequency Feature - Planned Comparisons for Response

Time	Map	Group	Planned Comparison Results
Week 2 vs Week 0	State B	Responders	*N.S.*
		Non-responders	t = -5.62; df=57; Bonferroni-corrected $p<0.0001$; Cohen's $d=0.57$
	State D	Responders	*N.S.*
		Non-responders	t = 4.72; df=57; Bonferroni-corrected $p<0.0001$; Cohen's $d=0.84$
Week 8 vs Week 2	State C	Responders	*N.S.*
		Non-responders	t = -2.92; df=57; Bonferroni-corrected $p=0.02$; Cohen's $d=0.29$
	State D	Responders	t = 3.57; df=50; Bonferroni-corrected $p=0.003$; Cohen's $d=0.49$
		Non-responders	*N.S.*

Table 7.5 - Coverage Feature ANOVA Post-Hoc Results (Tukey-HSD)

Interaction Effect	Post-hoc		
	Comparison	State	Stats
Time x Maps	Week 2 vs Week 0	State D	HSD=12.5; $df=588$; $p<0.0001$; Cohen's $d=0.92$
	Week 8 vs Week 0	State A	HSD=-9.08; $df=588$; $p<0.0001$; Cohen's $d=0.65$
	Week 8 vs Week 0	State C	HSD=-8.51; $df=588$; $p=0.001$; Cohen's $d=0.62$
	Week 8 vs Week 0	State D	HSD=11.7; $df=588$; $p<0.0001$; Cohen's $d=0.99$
	Week 8 vs Week 2	State A	HSD=-5.30; $df=588$; $p=0.01$; Cohen's $d=0.42$
	Week 8 vs Week 2	State B	HSD=8.33; $df=588$; $p=0.005$; Cohen's $d=0.67$
Maps x Site	CAM vs QNS	State A	HSD=-4.70; $df=294$; $p=0.02$; Cohen's $d=0.72$
	CAM vs QNS	State C	HSD=4.53; $df=294$; $p=0.001$; Cohen's $d=0.73$
	CAM vs TGH	State A	HSD=-7.69; $df=294$; $p<0.0001$; Cohen's $d=1.11$
	CAM vs TGH	State C	HSD=6.20; $df=294$; $p=0.0003$; Cohen's $d=0.94$
	CAM vs UBC	State C	HSD=4.58; $df=294$; $p=0.02$; Cohen's $d=0.68$
	TGH vs UBC	State A	HSD=8.21; $df=294$; $p<0.0001$; Cohen's $d=0.63$
Time x Maps x Site	CAM vs TGH - Week 0	State A	HSD=-6.47; $df=588$; $p=0.005$; Cohen's $d=1.59$
	CAM vs TGH - Week 2	State A	HSD=-5.87; $df=588$; $p=0.02$; Cohen's $d=1.78$
	TGH vs UBC - Week 0	State A	HSD=9.52; $df=588$; $p<0.0001$; Cohen's $d=1.25$
	TGH vs UBC - Week 0	State D	HSD=-5.44; $df=588$; $p=0.06$; Cohen's $d=0.86$
	TGH vs UBC - Week 2	State A	HSD=6.94; $df=588$; $p=0.001$; Cohen's $d=1.13$

Table 7.6 - Coverage Feature - Planned Comparisons for Response

Time	Map	Group	Planned Comparison Results
Week 8 vs Week 0	State B	Responders	t = 2.88; df=50; Bonferroni-corrected p=0.006; Cohen's d=0.59
		Non-responders	*N.S.*
Week 2 vs Week 0	State A	Responders	*N.S.*
		Non-responders	t = -2.53; df=57; Bonferroni-corrected p=0.05; Cohen's d=0.38
	State B	Responders	*N.S.*
		Non-responders	t = -4.00; df=57; Bonferroni-corrected p=0.0007; Cohen's d=0.59
	State D	Responders	*N.S.*
		Non-responders	t = 5.72; df=57; Bonferroni-corrected p<0.0001; Cohen's d=1.15
Week 8 vs Week 2	State B	Responders	*N.S.*
		Non-responders	t = 5.27; df=57; Bonferroni-corrected p<0.0001; Cohen's d=0.94
	State C	Responders	*N.S.*
		Non-responders	t = -3.18; df=57; Bonferroni-corrected p=0.009; Cohen's d=0.32

Table 7.7 - Results from Multiple Regression Model Analysis

	Week 2 ***Average Duration***	**Week 2** ***Frequency***	**Change from Baseline to Week 2** ***Coverage***
R	**0.289**	**0.293**	**0.338**
Adjusted R^2	0.038	0.041	0.070
ANOVA	F(5,101)=1.85, p=0.11	F(5,101) = 1.90, p = 0.10	F(5,101) = 2.60, p = 0.03
State A	b* = -0.002, p = 0.98	b* = -0.16, p = 0.22	b* = -0.03, p = 0.75
State B	b* = 0.27, p = 0.006	b* = 0.28, p = 0.03	b* = 0.25, p = 0.04
State C	b* = -0.006, p = 0.96	b* = -0.07, p = 0.56	b* = -0.14, p = 0.18
State D	b* = -0.066, p = 0.52	b* = -0.15, p = 0.19	b* = -0.14, p = 0.49
Age	b* = -0.05, p = 0.62	b* = -0.09, p = 0.42	b* = -0.09, p = 0.33

7.6 Figures

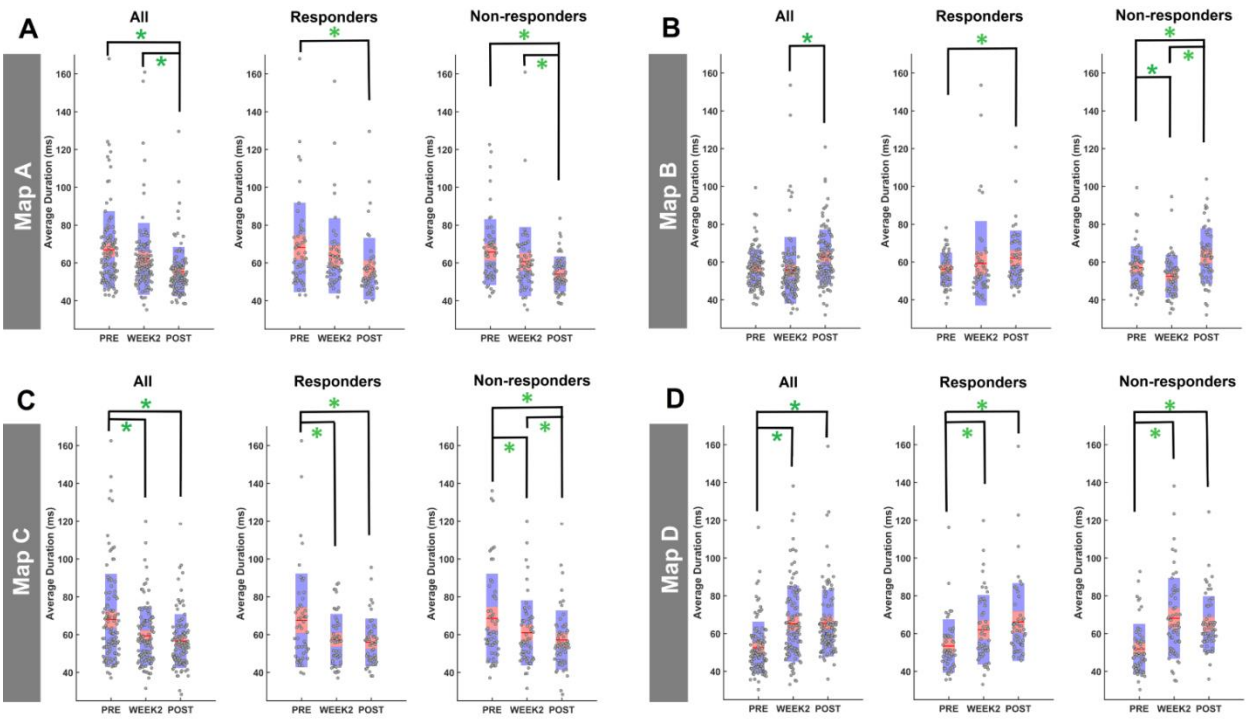


Figure 7.1 - Effect of escitalopram on the duration of all four microstates.
In each subplot, the raw data is plotted on top of a boxplot showing the mean (red line), 95% confidence interval (red area) and 1 standard deviation (blue area). Significant comparisons are marked with a green (*). **(A-D)** Plots of duration of microstates A-D.

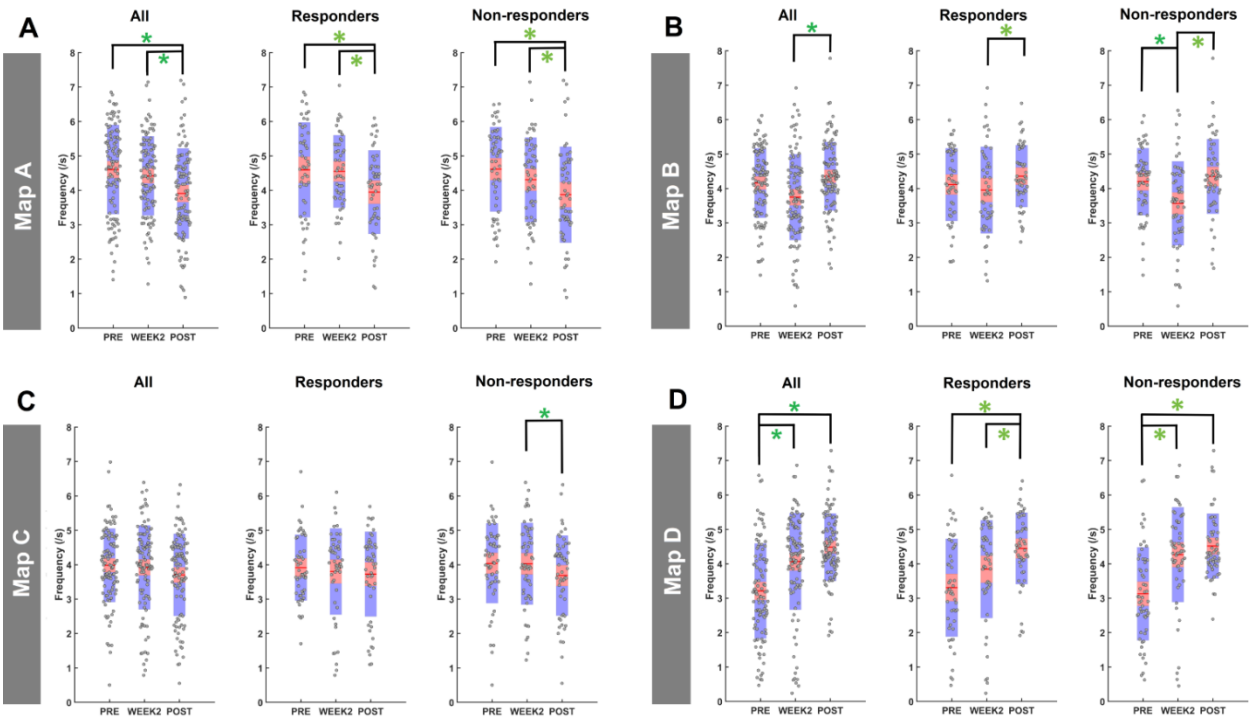


Figure 7.2 - Effect of escitalopram on the frequency of all four microstates.
In each subplot, the raw data is plotted on top of a boxplot showing the mean (red line), 95% confidence interval (red area) and 1 standard deviation (blue area). Significant comparisons are marked with a green (*). **(A-D)** Plots of frequency of microstates A-D.

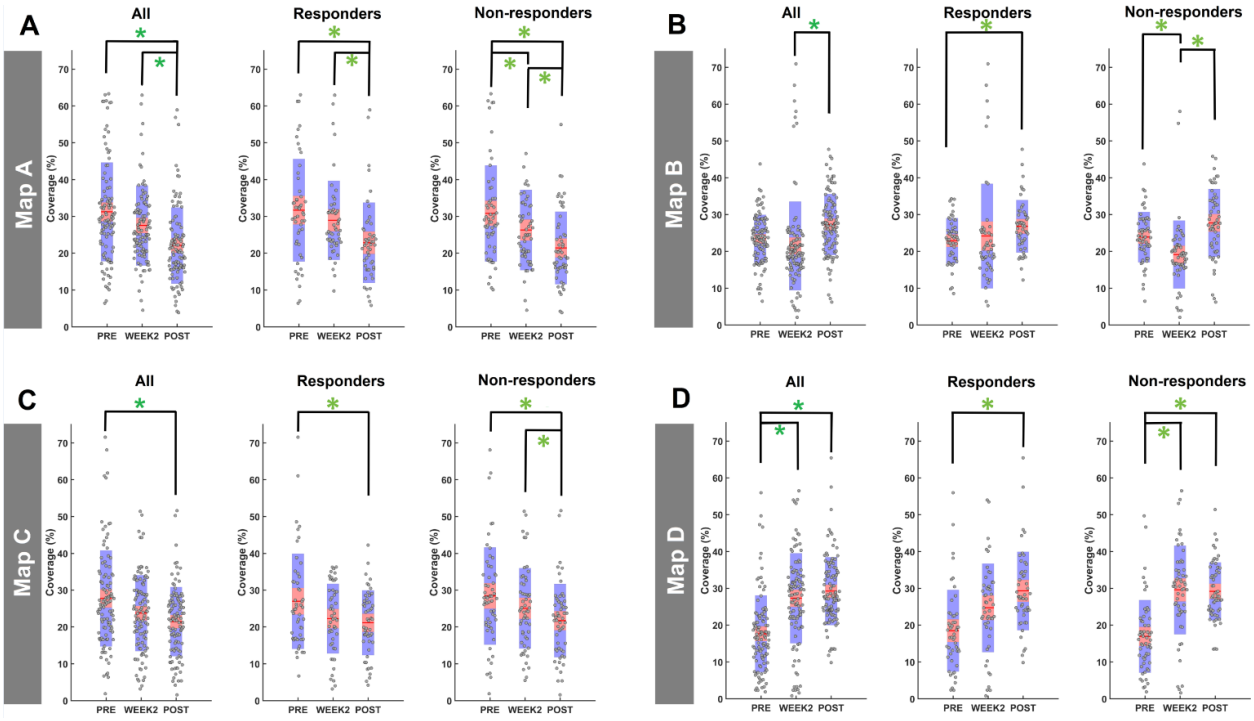


Figure 7.3 - Effect of escitalopram on the coverage of all four microstates.
In each subplot, the raw data is plotted on top of a boxplot showing the mean (red line), 95% confidence interval (red area) and 1 standard deviation (blue area). Significant comparisons are marked with a green (*). **(A-D)** Plots of coverage of microstates A-D.

Section IV: Predicting Response to Pharmacotherapy Using EEG Measures of Neural Dynamics

Chapter 8 – Supervised Machine Learning for the Early Prediction of Response to Escitalopram in Major Depressive Disorder: An EEG Study

In this chapter, we integrate EEG measures of neural dynamics from frequency analysis, multiscale entropy analysis and microstate analysis and evaluate their performance for predicting response to escitalopram.

Contents of this chapter were prepared for publication. Please note that during the publication process, errors may be discovered which could affect the content.

Sravya Atluri, Willy Wong, Andrey Zhdanov, Yasaman Vaghei, Zafiris Daskalakis, Daniel Blumberger, Benicio N. Frey, Peter Giacobbe, Raymond Lam, Glenda M. MacQueen, Roumen Milev, Daniel J Müller, Sagar V. Parikh, Susan Rotzinger, Claudio N. Soares, Colleen A. Brenner, Fidel Vila-Rodriguez, Steven C. Strother, Mary Pat McAndrews, Killian Kleffner, Esther Alonso, Stephen R. Arnott, Rudolf Uher, Jane A. Foster, Sidney H. Kennedy and Faranak Farzan. "Supervised Machine Learning for the Early Prediction of Response to Escitalopram in Major Depressive Disorder: An EEG Study."

8.1 Abstract

One major cause for the social and economic burden of depression is the number of years it can take to find an ideal treatment. Antidepressants are often the first line of treatment for major depression. However, with several different types of medication available, the challenge is in identifying which medication is best suited for each patient. To reduce the time spent in failed trials and avoid the debilitating impact of untreated depression, objective predictors of treatment response must be identified. In this study, machine learning methods were applied to evaluate resting-state neural dynamics for the prediction of escitalopram response. As part of phase 1 of the Canadian Biomarker Integration Network in Depression (CANBIND-1) study, patients received 8 weeks of open-label escitalopram treatment. Resting-state, eyes-closed electroencephalography data was included from 114 patients at baseline, and 107 patients two weeks into the treatment trial. Four prediction models were derived using support vector machines: (i) baseline, (ii) week 2, (iii) 'change from baseline to week 2', and **(iv)** a combined model using data from (i) and (iii). Performance was evaluated with feature sets derived from a feature polling combined with randomized permutation cross-validation. Classification accuracy was 72.8% for the baseline model, 69.2% for the week 2 model, 70.1% for the 'change from baseline to week 2' model, and 83.2% for the combined model. The integration of markers representing baseline neural activity with markers associated with early changes in neural activity was shown to augment the prediction of antidepressant response. With further validation, results from this study can potentially contribute towards the development of a personalized treatment selection tool for patients with depression.

8.2 Introduction

Antidepressants are an effective course of treatment and the first line of treatment for patients with major depressive disorder. Yet, remission rates are around 30% for the first trial and decline progressively with subsequent medication trials (Trivedi, Rush et al. 2006; Rush et al. 2006). Due to the heterogeneity of depression and the absence of tools that can identify the best treatment target for each patient, matching patients to an appropriate treatment has been a

daunting task for clinicians. Currently, a trial-and-error process is implemented to identify the antidepressant best suited for each patient, but this process can take time and patients may spend months to years suffering from symptoms (Solomon, Keller et al. 1997). Although clinical interviews and scales are available to confirm diagnosis and severity of symptoms, they are not sufficient for selecting an appropriate treatment for each patient (Winokur 1997, Chekroud, Zotti et al. 2016). One solution that might help to reduce the time spent in failed trials and avoid the debilitating impact of untreated depression (i.e., poor quality of life, economic burden), is identifying early biological predictors of response to an antidepressant. A personalized tool for the prediction of response to antidepressants can increase treatment efficacy rate and lead to the faster relief of symptoms.

Converging lines of evidence suggest that spontaneous fluctuations in neural activity at baseline, as measured by electroencephalography (EEG), may predict subsequent clinical response to antidepressants (Hunter, Cook et al. 2007, Kemp, Gordon et al. 2008, Iosifescu 2011, Baskaran, Milev et al. 2012, Olbrich and Arns 2013). Studies have also shown that evidence of symptom reduction, as early as two weeks after treatment, may be predictive of treatment outcome (Szegedi, Jansen et al. 2009, de Vries, Roest et al. 2018). For example, several EEG studies have reported that resting-state neural oscillations, especially in the alpha and theta band, may predict response to antidepressants (Ulrich, Renfordt et al. 1986, Ulrich, Renfordt et al. 1986, Knott, Telner et al. 1996, Leuchter, Cook et al. 1997, Cook, Leuchter et al. 2005, Bares, Brunovsky et al. 2008, Rabinoff, Kitchen et al. 2011). Posterior alpha activity has been associated with response to fluoxetine and amitriptyline (Ulrich, Renfordt et al. 1984, Bruder, Sedoruk et al. 2008), theta activity with response to imipramine, venlafaxine and several SSRIs (Knott, Telner et al. 1996, Iosifescu, Nierenberg et al. 2005, Iosifescu, Greenwald et al. 2009), delta activity with response to imipramine and paroxetine (Knott, Telner et al. 1996, Knott, Mahoney et al. 2000), alpha asymmetry with response to fluoxetine (Bruder, Stewart et al. 2001), and increased activity in the rostral anterior cingulate cortex with response to nortriptyline, fluoxetine and venlafaxine (Pizzagalli, Pascual-Marqui et al. 2001, Korb, Hunter et al. 2009). A few studies have also evaluated the association between non-linear measures of EEG signals (e.g., complexity) and response to antidepressants such as citalopram, clomipramine, escitalopram, bupropion and mirtazapine (Thomasson, Pezard et al. 2000, Méndez, Zuluaga et al. 2012,

Jaworska, Wang et al. 2018). However, EEG measures have not been adopted routinely into clinical practice for treatment outcome prediction. This may be because: (i) some measures were identified in small sample sizes and therefore, may not be reproducible, (ii) no evaluation was conducted on generalizability to an independent dataset (either through held-out data or cross-validation), and/or (iii) poor predictive performance was reported. As such, it is not clear whether features derived from resting-state EEG data can be used in clinical practice to predict response to treatment and ultimately match patients to interventions. In this study, we aim to build a model for antidepressant response prediction while addressing these limitations.

The main objective of this study is to investigate whether baseline or early changes (2 weeks into treatment) in resting-state neural dynamics can predict response to escitalopram treatment in patients with major depressive disorder. We hypothesize that the inclusion of a second time point (2 weeks into treatment) and also the integration of multiple EEG measures (linear and non-linear) would increase the predictive value of the model.

Several characteristics of EEG dynamics can be integrated into the prediction model. In addition to traditional frequency analysis, we considered power spectral features in the source domain (Pizzagalli, Pascual-Marqui et al. 2001, Pizzagalli 2011), spatiotemporal complexity (Costa, Goldberger et al. 2005, Farzan, Atluri et al. 2017), and global brain-network dynamics (Pascual-Marqui, Michel et al. 1995), previously shown to have predictive value for antidepressant response. With this approach, the resulting feature set can be extremely large in size, so feature reduction methods were applied to improve the efficiency of the feature selection process and identify the feature set that can best discriminate responders and non-responders. For the machine learning method, support vector machines were used due to their wide use and success in previous neuropsychiatric studies (Orru, Pettersson-Yeo et al. 2012).

8.3 Methods

Participants. In Phase 1 of the CAN-BIND study, 211 participants aged 18-60 years who met the DSM-IV requirements for major depressive disorder were recruited and completed the baseline

visit (Kennedy, Lam et al. *In Press*, 2018). Of this group, 180 patients completed 8 weeks of standardized escitalopram treatment (10-20mg). At the 8 week visit, responder or non-responder status was determined as ≥ 50% decrease in MADRS score from baseline. There were 85 responders and 95 non-responders at week 8 (Kennedy, Lam et al. *In Press*, 2018). Resting-state, eyes-closed EEG data was collected from four of the six sites: University of British Columbia (UBC), Toronto General Hospital (TGH), Queens University (QNS) and the Centre for Addiction and Mental Health (CAMH). At each visit, approximately eight minutes of EEG data was collected. EEG data was collected at baseline (within 3 days before the start of the treatment trial), at the end of week 2 (i.e., two weeks after the beginning of the trial) and at the end of week 8 (i.e., eight weeks after the beginning of the trial). This study included resting-state, eyes-closed EEG data collected from 114 patients at baseline, and 107 patients at the end of week 2 of treatment. Remaining patients either did not complete EEG assessments or the data was too noisy to include. Data was excluded prior to feature extraction. Clinical data for each visit is summarized in **Table 8.1** and **Table 8.2**. For a detailed description of the clinical data, research protocol and data acquisition at each site, see (Lam, Milev et al. 2016, Baskaran, Farzan et al. 2017, Farzan, Atluri et al. 2017, Kennedy, Lam et al. In Press, 2018).

Inter-site Data Harmonization. All EEG datasets were standardized to the following parameters: 58 electrodes common to all sites, 0.05-100Hz bandpass filter, Cz reference, 512Hz sampling rate. The EEG files were then exported as an EEGLAB (Delorme and Makeig 2004) dataset. Data was standardized using MATLAB R2012b-R2016a with the EEGLAB toolbox (v12.0.2.6b). Complete descriptions on the standardization of EEG data across sites in the CAN-BIND study were recently published (Farzan, Atluri et al. 2017).

Data Preprocessing. During pre-processing, EEG data was divided into two second continuous epochs, bandpass-filtered between 1-80Hz (2^{nd} order Butterworth), and notch-filtered at 60Hz (2^{nd} order Butterworth). EEGLAB implementation of independent component analysis was used to extract eye, muscle and electrode artifacts. Deleted EEG channels were interpolated using spherical spline interpolation (Perrin, Pernier et al. 1989) and data was re-referenced to an average reference. This preprocessing pipeline is currently made available as ERPEEG (http://www.tmseeg.com/multisiteprojects/).

Feature Extraction. The EEG data of each of the channels was divided into ten windowed segments of 20 seconds each (i.e., 20 seconds was the longest continuous epoch length we could extract from the data; a minimum of 10 segments were available from the data). The segments were used to calculate each of the metrics below. These results were then averaged across all windows. Averaging reduces variance in the data as well as the effect of outliers. Each EEG channel is represented by a single metric for single-electrode features (e.g., power, complexity) or a single metric may represent all 58 EEG channels for a global feature (e.g., microstates). These features are then passed onto the machine learning method.

(i) Frequency analysis. The EEGLAB function *spectopo* was used to obtain the power spectrum for each electrode. The log-transformed absolute power was obtained for each channel for each of the specified frequency bands: delta (1-3.5Hz); theta (4-8Hz); low-alpha (8.5-10Hz); high-alpha (10.5-12Hz); low-beta (12.5-18Hz); mid-beta (18.5-21Hz); and high-beta (21.5-30Hz). Asymmetry between left and right hemispheres was also considered. The absolute power in the left hemisphere was divided by the absolute power in the right hemisphere for all possible channel pairs in the 58 electrode montage: FP1/FP2, AF3/AF4, F7/F8, F5/F6, F3/F4, F1/F2, FT7/FT8, FC5/FC6, FC3/FC4, FC1/FC2, T7/T8, C5/C6, C3/C4, C1/C2, TP7/TP8, CP5/CP6, CP3/CP4, CP1/CP2, P7/P8, P5/P6, P3/P4, P1/P2, PO7/PO8, PO3/PO4, and O1/O2.

(ii) Current source density analysis. Data was analyzed using the LORETA-KEY software using the eLORETA algorithm (Pascual-Marqui, Lehmann et al. 1999). The transformation matrix (58 channels to 6239 voxels) was derived by co-registering electrode co-ordinates (10-10 international system) to the MNI152 MRI template (i.e., the head model) (Pascual-Marqui 2002) and a relative regularization parameter of 1. The solution space of eLORETA is restricted to cortical and some hippocampal and amygdala grey matter. The MNI152 template brain volume is divided into 6239 cortical gray matter voxels at 5-mm^3 resolution (Pascual-Marqui 2002). From scalp-recorded electrical potential distribution, LORETA computes the three dimensional intracerebral distributions of current density for each of the specified bands of frequency: delta (1-3.5Hz); theta (4-8Hz); low-alpha (8.5-10Hz); high-alpha (10.5-12Hz); low-beta (12.5-18Hz); mid-beta (18.5-21Hz); and high-beta (21.5-30Hz). Regions were selected based on previous

literature: the anterior cingulate cortex (ACC), rostral ACC (rACC), and the medial orbitofrontal cortex (mOFC) (Pizzagalli, Pascual-Marqui et al. 2001, Korb, Hunter et al. 2009).

(iii) Multiscale entropy. Multiscale entropy analysis was performed using the methods outlined in (Costa, Goldberger et al. 2005, Farzan, Atluri et al. 2017). Using the sample entropy equation, multiscale entropy was examined across all 58 electrodes with the coarse-graining process (for 70 scales). Sample entropy quantifies the variability of time series by estimating the predictability of amplitude patterns across a time series. In our analysis, two consecutive data points were used for data matching (m=2) and data points were considered to match if their absolute amplitude difference was less than 15% (i.e., r =0.15) of the standard deviation of the time series (similar to (Costa, Goldberger et al. 2005, Farzan, Atluri et al. 2017)).

Asymmetry between left and right hemispheres was also considered. Multiscale entropy in the left hemisphere was divided by the multiscale entropy in the right hemisphere for the following 25 channel pairs: FP1/FP2, AF3/AF4, F7/F8, F5/F6, F3/F4, F1/F2, FT7/FT8, FC5/FC6, FC3/FC4, FC1/FC2, T7/T8, C5/C6, C3/C4, C1/C2, TP7/TP8, CP5/CP6, CP3/CP4, CP1/CP2, P7/P8, P5/P6, P3/P4, P1/P2, PO7/PO8, PO3/PO4, and O1/O2.

(iv) Microstate analysis. Microstate analysis followed the standard procedure outlined in (Lehmann, Ozaki et al. 1987, Pascual-Marqui, Michel et al. 1995) and was implemented using CARTOOL (Brunet, Murray et al. 2011) . Prior to the application of microstate analysis, the pre-processed EEG data was bandpass-filtered from 1-30 Hz. The topographical maps at the local maxima peaks of the global field power curve are clustered to derive the four prototypical microstate classes (Koenig, Prichep et al. 2002). In this study, the topographical atomize–agglomerate hierarchical clustering algorithm (Tibshirani and Walther 2005) was applied to cluster each individual EEG data into four states (microstate maps). Finally, topographical maps at each local maxima point of the global field power curve were assigned to the microstate class of highest correlation using Pearson's spatial product-moment correlation coefficient (Brandeis, Naylor et al. 1992) . Three features were calculated for each of the four microstate classes: (i) average duration, (ii) frequency, and (iii) coverage. Average duration is the amount of time a microstate class remains stable when it appears, in milliseconds; frequency refers to the

occurrence of each microstate class per second; and coverage is the percent of recording covered by each microstate class.

Machine Learning Algorithm. In this study, support vector machines (SVMs) were used with the radial basis function (RBF) kernel using the LIBSVM toolbox (Chang and Lin 2011). Support vector machines select a hyperplane (linear) or hypersurface (non-linear) that best separates the input data space into two (or more) pre-defined groups (i.e., responder and non-responder) (Hearst, Dumais et al. 1998). An RBF kernel uses nonlinear mapping to transform data into a higher dimensional space and determine an optimal hypersurface for classification (Hsu, Chang et al. 2003). The optimal hypersurface separates two groups with the largest margin (i.e., distance between the hypersurface and the closest data points).

Two model hyperparameters can be specified for the RBF kernel: cost and gamma. To avoid overfitting and ensure the model is well-fit to the given data, a small and restricted range of cost and gamma were used during model optimization. Default value for cost is 1 and gamma is 0. Therefore, the range for cost was specified to be around 1 and for gamma, the range was close to 0. Cost was specified as 2^C, where C = {-3, -1, 1, 3} and gamma was specified as 2^G, where G = {-12, -10, -8, -6}. Model optimization was performed using the grid search method (Chang and Lin 2011) and balanced accuracy was used as the outcome measure. Balanced accuracy (i.e., average of sensitivity and specificity) was used because the number of responders and non-responders was not equal in this study. Therefore, unlike accuracy, balanced accuracy would not be biased to the performance of the group with the larger sample size.

Prediction Models. Four models were created for prediction: (i) baseline model using features derived from the data collected during baseline, (ii) week 2 model using features derived from data collected 2 weeks after treatment, (iii) early change model including change in features from baseline to week 2, and (iv) a combined model using data from (i) and (iii).

Feature Set. For the baseline and week 2 model, the total feature set consisted of 6424 features: 406 features from frequency analysis (7 bands x 58 channels), 175 features from considering asymmetry in frequency features (7 bands x 25 channel pairs between left and right hemisphere),

4060 features from multiscale entropy analysis (70 scales x 58 channels), 1750 features from multiscale entropy across both hemispheres (70 scales x 25 channel pairs between left and right hemisphere), 12 features from microstate analysis (4 maps x 3 features) and 35 features from current source density analysis (7 bands x 3 regions).

Features for the 'change from baseline to week 2' model were calculated as (post-pre)/pre*100 for multiscale entropy, microstate and current source density features and (post-pre) for power features. Asymmetrical features for the change model were calculated as:

$$\frac{(WEEK2_{Left} - BASELINE_{Left})/BASELINE_{Left} * 100}{(WEEK2_{Right} - BASELINE_{Right})/BASELINE_{Right} * 100}$$

Therefore, the total number of features for the change model was also 6424. For the combined model (baseline + change from baseline to week 2), the total number of features was (6424+6424) 12,848.

Feature Selection Method. Identifying the optimal feature set for machine learning is an unsolved problem (Blum and Langley 1997, Guyon and Elisseeff 2003). One common approach is to use wrapper methods like greedy algorithms. However, such methods may be prone to over-fitting and can still be quite computationally intensive (Saeys, Inza et al. 2007). In this study, filter methods are used as a means of removing uninformative features (t-test, F-test and Spearman's correlation). Filter methods can be less prone to over-fitting and are easy to compute (Saeys, Inza et al. 2007). The methods were compared using balanced accuracy (average of sensitivity and specificity).

Model Performance Evaluation. Each of the four models (i.e. baseline model; week 2 model; change from baseline to week 2 model; and combined model using baseline and change from baseline to week 2 features), was evaluated using one of three filtering methods (t-test, F-test or Spearman's correlation). A randomized permutation cross-validation method similar to the Monte-Carlo cross-validation procedure (Molinaro, Simon et al. 2005) was used. We then followed steps 1-5 listed below, repeating them over H iterations. In this study, H=100.

Step 1: For each of the four models, the entire dataset was randomly split into a training set (80% of responders and 80% of non-responders) and a testing set (remaining 20% of responders and 20% of non-responders). In other words, a test set was selected such that it is independent from the training set. The 80:20 ratio was fixed to avoid iterations with an unequal ratio of responders to non-responders. SVMs show poor performance with unequal class sizes.

Step 2: One of the filtering methods was applied to select features from the training set only. A statistical threshold of $\alpha<0.05$ was applied to all methods. The training set was normalized by subtracting each feature by the mean of the feature across subjects and dividing by the standard deviation of each feature across subjects (z-score normalization).

Step 3: Optimal values of cost and gamma parameters for RBF SVM classifier were estimated from the training data using 10-fold cross-validation. Model optimization was conducted using the grid search method (Hsu, Chang et al. 2003). Balanced accuracy was used as the outcome measure for optimization.

Step 4: Only features identified in Step 2 were used for the test set. The test set was normalized using the same mean and standard deviation values as the training set (z-score normalization as in Step 2).

Step 5: The optimum model derived in Step 3 was used to classify the test data.
Steps 1-5 were performed for each of the filtering methods for feature reduction. The filtering method that provided the best performance was chosen for the next steps.

Estimating Model Performance with Reduced Feature Sets. To identify a subset of potentially robust features, we apply the following procedure. During each of the permutations H_i (i = 1,2,…,100), the total feature set (X_j, where j is the feature index from 1 to total number of features for the model) is reduced to a smaller subset using a filtering method (filtering method is only applied to the randomly selected training set – 80% of data). In other words, in each permutation i, feature j may or may not be selected. If the feature j is selected, it gains a vote.

This feature polling method is applied over all 100 permutations. At the end of 100 permutations, the number of votes for each feature may vary from 0 to 100. We evaluated and compared model performance over several thresholds (T = 50, 60, 70, 80, 90). Using features with T or more votes, the final classification model was built on the entire dataset. The entire dataset was normalized together. A 10-fold cross-validation was applied to evaluate the model. Cost and gamma were optimized for maximum balanced accuracy. A low value of cost (close to 1) suggests a lower chance of overfitting to the data. Results from this procedure are summarized in **Table 8.3.**

Classification Evaluation Metrics. The performance metrics reported in this study are accuracy, balanced accuracy, sensitivity (or recall), specificity and precision. The metrics are therefore defined as:

$$Accuracy = \frac{TP + TN}{TotalNumberofSamples}$$

$$BalancedAccuracy = \frac{\frac{TP}{TP + FN} + \frac{TN}{TN + FP}}{2}$$

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Precision = \frac{TP}{TP + FP}$$

Pooled, Within-Site and Between-Site Classification. To evaluate whether pooling data from all sites was valid, classifier performance was evaluated and compared using three tests: (i) **pooled classification**: randomized permutation cross-validation (100 permutations) was performed on data pooled from 3 of the 4 sites. Features with 70 or more votes were chosen for the final model (70 was arbitrarily chosen to ensure all comparisons were consistent). Using this reduced feature set, 10-fold cross validation was performed on the same data from 3 of the 4 sites to evaluate the performance of the model when one site is excluded. (ii) **within-site classification:** randomized permutation cross-validation was performed within sites with larger sample sizes (UBC and

TGH only) to identify features with 70 or more votes. Using this reduced feature set, 10-fold cross validation was performed to evaluate the performance of the model within each site. (iii) **between-site classification:** randomized permutation cross-validation was performed using data from one site and features with 70 or more votes were identified. Using this reduced feature set, 10-fold cross validation was performed to evaluate the performance of the model on another site. These tests were conducted to evaluate the generalizability of the classifier to data collected from different locations, scanners, etc. Methods for these tests were modified from (Rozycki, Satterthwaite et al. 2017).

8.4 Results

8.4.1 Model Evaluation: Comparing Filtering Methods

Using the randomized (100) permutation cross-validation method, models were evaluated with different feature selection (filter) methods. All the filter methods were comparable. However, the t-test method showed slightly better performance for most of the models, specifically in identifying non-response (good specificity). It also provided a low feature-to-subject ratio. Therefore, the t-test method was chosen to be the feature selection method for the final classification model.

8.4.2 Model Performance with Reduced Feature Sets

Each of the four models was evaluated with a reduced feature set derived using the (t-test) feature polling method outlined above. Depending on the threshold selected for the feature polling method (50, 60, 70, 80 or 90 votes), the predictive performance of the model will also vary. In addition, the choice of the threshold may also impact the bias-variance trade-off. Results for these tests are presented in detail in **Table 8.3**. In the following text, we compare model performance at a threshold of 70 or votes.

For the baseline model, features with 70 or more votes yielded an accuracy of 72.8%, sensitivity of 51.0%, specificity of 90.5% and precision of 81.3%. For the week 2 model, features with 70

or more votes yielded an accuracy of 69.2%, sensitivity of 65.3%, specificity of 72.4% and precision of 66.7%. For the change from baseline to week 2 model, features with more than 70 votes yielded an accuracy of 70.1%, sensitivity of 59.2%, specificity of 79.3% and precision of 70.7%. Lastly, for the combined model, features with 70 or more votes yielded an accuracy of 83.2%, sensitivity of 73.5%, specificity of 91.4% and precision of 87.8%. In the next section, we describe the features included in these reduced feature sets that yielded highest performance.

8.4.3 Discriminating Features for Response Classification

For the baseline model, features with 70 or more votes that yielded an accuracy of 72.8% (**Table 8.3**) are illustrated in **Figure 8.1**, where red indicates that the feature value is greater in responders than non-responders and blue indicates that the feature is smaller in responders than non-responders. To summarize, the descriptive features of response were: complexity in mid-timescales (35-45) in central (C1) regions; asymmetry in complexity in fine timescales (21, 31) in frontal (F7/F8) and centro-parietal (CP3/CP4) regions; asymmetry in complexity in coarse timescales (35-70) in frontal (F1/F2), fronto-central (FC1/FC2), central (C3/C4), centro-parietal (CP1/CP2, CP3/CP4), parietal (P3/P4, P5/P6) and temporal regions (T7/T8, TP7/TP8); mid-beta power (18.5-21Hz) in a parieto-occipital region (POZ); and frontal asymmetry (AF3/AF4) in low-beta power (12.5-18Hz). From current source density analysis, features included high-alpha band power (10.5-12Hz) in the ACC. Finally, coverage of Map A was identified from microstate analysis.

For the week 2 model, features with 70 or more votes that yielded an accuracy of 69.2% (**Table 8.3**) are illustrated in **Figure 8.2**. These features included: asymmetry in complexity in fine timescales (5-7) in central (C1/C2) and centro-parietal regions (CP1/CP2); asymmetry in complexity in coarse timescales (40-70) in central (C5/C6), centro-parietal (CP3/CP4) and occipital (O1/O2) regions; mid-beta power (18.5-21Hz) in frontal (F4, F6, F5), fronto-central (FC4), and parieto-occipital regions (POZ); high-beta power (21.5-30Hz) in frontal (F4) regions; and parietal (P3/P4) asymmetry in low-alpha power (8.5-10Hz). No features from current source density and microstate analysis were identified.

For the ‘change from baseline to week 2’ model, features with 70 or more votes that yielded an accuracy of 70.1% (**Table 8.3**) are illustrated in **Figure 8.3**. These included: complexity in fine timescales (5-25) in frontal (F5, F6, F8, FT7), fronto-central (FC2, FC4, FC5) and parietal (P4) regions; asymmetry in complexity in a fine timescale (6) in the temporal (T7/T8) region; asymmetry in complexity in coarse timescales (35, 61, 69) in central (C3/C4) and frontal (F1/F2, F7/F8) regions respectively; delta (1-3.5Hz) and theta (4-8Hz) power in frontal (F7, F5, F4, F6), fronto-central (FC4, FC5) regions; low-beta power (12.5-18Hz) in frontal (F2, F4, F6), fronto-central (FC4, FC5) and centro-parietal (CP6) regions; mid-beta power (18.5-21Hz) in the frontal (F4, F6) and fronto-central (FC4) regions; high-beta power (21.5-30Hz) in the frontal (F4, F6) regions; and mid-beta band power (18.5-21Hz) in the ACC. No features from microstate analysis were identified.

For the combined model, features with 70 or more votes that yielded an accuracy of 83.2% (**Table 8.3**) are illustrated in **Figure 8.4**. From the baseline data (left panel of **Figure 8.4**), these features included: complexity in fine timescales (15-20) in a parietal (P3) region; asymmetry in complexity in fine timescales (1-35) in frontal (F1/F2, F7/F8), fronto-central (FC1/FC2) regions; asymmetry in complexity in coarse timescales (35-70) in fronto-central (FC1/FC2), temporal (TP7/TP8) and parietal (P3/P4, P5/P6) regions; asymmetry in high-alpha power (10.5-12Hz) in parietal (P5/P6) region; asymmetry in low-beta power (12.5-18Hz) in centro-parietal (CP1/CP2, CP5/CP6) regions; and high-alpha power (10.5-12Hz) in the ACC. No features from microstate analysis were identified.

For the combined model again, features from the ‘change from baseline to week 2’ data (right panel of **Figure 8.4**) included: complexity in fine timescales (1-25) in frontal (F5, F6, F8, FT7), fronto-central (FC2, FC4, FC5), and parietal (P4) regions; asymmetry in fine timescales (6, 34) in complexity in temporal (T7/T8) and central (C3/C4) regions; asymmetry in complexity in coarse timescales (59, 69) in frontal (F1/F2, F7/F8) regions; delta (1-3.5Hz) and theta (4-8Hz) power in frontal (F7, F5, F4, F6), fronto-central (FC4, FC5) regions; low-beta power (12.5-18Hz) in frontal (F2, F4, F6), fronto-central (FC4, FC5) and centro-parietal (CP6) regions; mid-beta power (18.5-21Hz) in the frontal (F4, F6) and fronto-central (FC4) regions; high-beta power

(21.5-30Hz) in the frontal (F4, F6) regions; and mid-beta band power (18.5-21Hz) in the ACC. No features from microstate analysis were identified.

8.4.4 Classification is Valid Using Data Pooled across Sites

Pooled, within-site and between-site analysis results are detailed in **Table 8.4**, **Table 8.5, Table 8.6** and **Table 8.7**. Pooled classification with leave-one-site-out (green) was shown to have similar classification accuracy to within-site classification (blue) indicating that pooling across sites is valid. The exclusion of any single site has little effect on results across all models. Between-site analysis results are also reported for completeness (highlighted in red).

8.5 Discussion

The aim of this study was to evaluate the feasibility of predicting response to escitalopram using EEG measures of neural dynamics. Specifically, we integrated several EEG measures of neural dynamics to evaluate their cumulative predictive value. Using randomized permutation cross-validation combined with the feature polling method (threshold of $\geq$ 70), we showed that classification accuracy was 72.8% using baseline data (sensitivity = 51.0%, specificity = 90.5%), 69.2% using week 2 data (sensitivity = 65.3%, specificity = 72.4%) and 70.1% using 'change from baseline to week 2' data (sensitivity = 59.2%, specificity = 79.3%). In addition, as hypothesized, combining baseline neural dynamics with early changes in neural dynamics (change in response to 2 weeks of treatment) increased the accuracy of prediction to 83.2% (sensitivity = 73.5%, specificity = 91.4%). Results from pooled, within-site and between-site analysis also demonstrated that the large-scale analysis of data pooled across multiple sites does not have a significant impact on classifier performance. This is an important finding as it suggests that data can be integrated regardless of scanner and equipment differences, and that the predictive markers identified in this work may have potential for clinical translation.

Our primary discriminatory features, identified through randomized permutation cross-validation combined with the feature polling method, were the asymmetry features (both power and

complexity). Asymmetry in neural activity across the two hemispheres was previously shown to be a marker of depression (Bruder, Stewart et al. 2001). Through our method, we were also able to identify the regions associated with this asymmetry in frequency (power) and time (multiscale entropy). The identified features were specific to the fronto-central and parieto-occipital regions potentially suggesting a relationship to the default-mode network. However, further investigation (e.g., via fMRI data) is needed to better elucidate this relationship.

Among the four models (baseline; week 2; 'change from baseline to week 2'; and the combination of baseline and 'change from baseline to week 2'), the combined model yielded the highest accuracy. This may suggest an early impact of antidepressants on neural circuits. Furthermore, it demonstrates that early changes in neural dynamics with treatment (i.e., 2 weeks) can contribute useful information towards the prediction of an 8-week treatment outcome. When evaluating the effect of study site in all four models, we found that the exclusion of any single site does not have a significant effect on classification performance. Therefore, pooling data across sites is recommended as it can improve the generalizability of the prediction model for this population.

A number of features identified through the feature selection process in this study have been previously shown in the literature to be reliable predictors of antidepressant response in major depressive disorder. These include parieto-occipital alpha (Ulrich, Renfordt et al. 1984, Bruder, Sedoruk et al. 2008, Tenke, Kayser et al. 2011), ACC activity (Mayberg, Brannan et al. 1997, Brody, Saxena et al. 1999, Pizzagalli, Pascual-Marqui et al. 2001, Pizzagalli, Oakes et al. 2003, Saxena, Brody et al. 2003, Korb, Hunter et al. 2009, Pizzagalli 2011, Spronk, Arns et al. 2011), medial and middle frontal cortex activity (Gonul, Kitis et al. 2006, Chen, Ridler et al. 2007), frontal alpha (Suffin and Emory 1995) and theta (Knott, Telner et al. 1996, Iosifescu, Greenwald et al. 2009), and frontal delta (Knott, Telner et al. 1996, Knott, Mahoney et al. 2001). In the literature, prediction accuracy varies for each feature and is between 60-77%. However, since most of these studies did not validate their results and/or used fairly small datasets, accuracies may be inflated and may decrease significantly when applied to a new independent dataset. In our study, we were able to confirm the predictive value of the previously reported features as

well as discover additional features that were not previously reported. Asymmetry in complexity of neural activity between the two hemispheres for example, is a novel marker of antidepressant response identified in this study.

Machine learning studies using resting-state EEG measures for antidepressant response prediction are scarce. In (Knott, Mahoney et al. 2001), resting-state EEG measures were used with a mixture of factor analysis classifier to predict antidepressant response. An accuracy of 87.9% was reported (sensitivity=94.9%; specificity=80.9%). However, the study had some limiting factors: (i) the study was performed with a low sample size (n=22), (ii) responders were defined with ≥30% improvement in clinical scores instead of the usual ≥50%, and (iii) training and testing dataset combined patients on 4 different medications (sertraline, citalopram, fluvoxamine, paroxetine). Accuracy was also high (85-92%) in (Rabinoff, Kitchen et al. 2011) using spectral EEG features with classification and regression tree analysis. The study combined trials for 2 antidepressants (fluoxetine and venlafaxine) to predict response in 51 patients with unipolar depression. The high accuracy however, may be due to overfitting to the data and this is suggested by the 100% specificity in all treatment groups.

Some of the preliminary findings published on the eyes-closed EEG data from the CANBIND phase 1 project were replicated when the threshold for the feature polling method was reduced (i.e., ≥ 50 votes) (Baskaran, Farzan et al. 2017). This included significant differences between responders and non-responders in absolute delta power at week 2. Furthermore, similar to the previous study, there was no difference in theta cordance (Leuchter, Cook et al. 1994) between responders and non-responders. On the other hand, some EEG measures that were not significant in the previous study (for example, alpha power, alpha asymmetry, theta power and theta asymmetry) were shown to have predictive value in this study, potentially due to the higher sample size. Four markers were evaluated in the previous study for their predictive value. Whole-brain absolute delta asymmetry at week 2 yielded an accuracy of 67.8% (sensitivity=77.8%; specificity=57.7%; precision=56.0%), baseline whole-brain absolute alpha asymmetry yielded an accuracy of 70.7% (sensitivity=72.2%; specificity=69.2%; precision=61.9%), baseline parietal alpha asymmetry yielded an accuracy of 77.2%

(sensitivity=88.9%; specificity=65.4%; precision=64.0%) and baseline whole-brain relative delta asymmetry yielded an accuracy of 73.3% (sensitivity=88.9%; specificity=57.7%; precision=59.3%). In the current study, our integrated and multi-feature approach (the combined model) yielded a classification accuracy of 83.2% (sensitivity=73.5%; specificity=91.4%; precision=87.8%) further supporting the integration of markers for more accurate prediction of treatment outcome.

Feature selection is a key step in classification models. In this study, three filtering approaches to feature reduction were applied and compared (t-test, F-test, and correlation). In general, these filtering approaches to feature reduction have advantages and limitations. One main advantage is that they are less prone to overfitting, but this is at the risk of providing low accuracy. Another key advantage is that they are far less computationally expensive than other feature selection methods (Saeys, Inza et al. 2007). One disadvantage of filtering methods, however, is that they may select features that are contaminated by noise (i.e., a feature with high measurement noise or artifacts) and this may lead to reduced generalizability of the model. We aimed to reduce this effect by performing randomized permutation cross-validation.

Validation of performance with an independent dataset is another important aspect of classification models. Yet, it was not implemented in several previous studies. It should be noted that by excluding validation, results on the predictive performance of a marker or a model are potentially inflated and highly unlikely to generalize to a new data set that is independent from the original data set. In this study, cross-validation was applied to select hyperparameters for the machine learning algorithm as well as to estimate the predictive performance of the model. Despite the advantages of cross-validation, it still only provides an estimate on the generalizability of the model. It is important to further validate each model on a larger independent data set for accurate estimates of prediction performance.

Our study has several limitations. First, due to the absence of an independent test data set, reported model performance estimates may not be accurate. To obtain an unbiased estimate of our model’s predictive performance, future studies should validate it against an independent

dataset (i.e., measures of neural dynamics from a different group of patients receiving escitalopram treatment) that was not used for model construction. Second, prediction models reported in this study may not be generalizable to other antidepressants. Model evaluation should be performed independently with several other types of treatments to evaluate the generalizability of our prediction models. This may also provide insight into whether the features identified in this study were specific to escitalopram response or if they can be used to predict general response to antidepressants. Finally, the inclusion of multivariate EEG features should also be explored in future work. Performance may also be improved with the addition of clinical or behavioural variables, genetic measures and other imaging-based measures (fMRI, DTI, etc.) with patient-reported data. These investigations will be conducted by future CAN-BIND studies.

8.6 Conclusions

This work provides a proof-of-concept pipeline for the prediction of escitalopram response and should be further augmented in the future for clinical prediction. For large datasets that include several groups of patients, each receiving a different treatment option (pharmacological and non-pharmacological antidepressants), an approach similar to the one taken by this study may be useful in developing a model that can match each patient to the most effective treatment. The feasibility of such an approach will in part depend on the collection and sharing of large-scale, clinically-reliable data sets, as done by CAN-BIND. These investigations will contribute towards the development of a clinical decision-making tool for data-driven, personalized optimization of antidepressant treatment selection for patients.

8.7 Tables

Table 8.1 - Demographics and Clinical Data for Subjects Included from the Baseline Visit

	UBC	TGH	QNS	CAM	All	All Responders (Week 8)	All Non-responders (Week 8)
N	51	38	18	7	114	51	63
Age (mean,std)	35.4 +/- 11.8	37 +/- 12.8	42.7 +/- 14.4	30.4 +/- 13.0	36.8 +/- 12.8	36.1 +/- 13.0	37.4 +/- 12.7
Gender (M/F)	18 / 33	17 / 21	9 / 9	0 / 7	44/70	18 / 33	26 / 37
MADRS Baseline (mean,std)	28.5 +/- 5.94	32.0 +/- 6.23	30.0 +/- 4.74	28.0 +/- 5.16	29.9 +/- 5.90	29.1 +/- 5.89	30.5 +/- 5.87
MADRS week 2 (mean,std)	21.9 +/- 7.41	24.9 +/- 11.1	23.1 +/- 5.36	21.1 +/- 3.98	23.0 +/- 8.47	19.6 +/- 8.08	25.8 +/- 7.77
MADRS week 8 (mean,std)	14.7 +/- 9.30	19.6 +/- 12.6	18.1 +/- 10.2	15.7 +/- 6.02	16.9 +/- 10.6	7.80 +/- 4.99	24.3 +/- 7.83
Change in MADRS (baseline to week 8) (mean,std) (%)	48.6 +/- 33.2	39.2 +/- 33.7	39.2 +/- 35.9	44.5 +/- 16.9	43.7 +/- 33.0	73.1 +/- 16.1	19.9 +/- 22.2
Responders / Non-responders (After 8 weeks)	24 / 27	17 / 21	7 / 11	3 / 4	51 / 63	-	-

Table 8.2 - Demographics and Clinical Data for Subjects Included from the Week 2 Visit

	UBC	TGH	QNS	CAM	All	All Responders (Week 8)	All Non-responders (Week 8)
N	50	34	17	6	107	49	58
Age (mean,std)	35.7 +/- 11.7	36.3 +/- 12.8	42.9 +/- 14.8	26.0 +/- 6.23	36.5 +/- 12.7	35.4 +/- 12.8	37.4 +/- 12.7
Gender (M/F)	18 / 32	14 / 20	8 / 9	0 / 6	40 / 67	17 / 32	23 / 35
MADRS Baseline (mean,std)	28.5 +/- 6.00	32.4 +/- 6.06	30.3 +/- 4.71	28.8 +/- 5.12	30.1 +/- 5.97	29.3 +/- 5.90	30.7 +/- 6.02
MADRS week 2 (mean,std)	21.9 +/- 7.49	24.9 +/- 11.3	23.0 +/- 5.50	21.8 +/- 3.87	23.0 +/- 8.53	19.7 +/- 8.22	25.8 +/- 7.80
MADRS week 8 (mean,std)	14.3 +/- 8.93	19.2 +/- 12.6	16.9 +/- 9.08	17.0 +/- 5.44	16.4 +/- 10.2	7.73 +/- 5.07	23.7 +/- 7.32
Change in MADRS (baseline to week 8) (mean,std) (%)	50.1 +/- 31.6	41.2 +/- 33.6	44.8 +/- 27.8	41.1 +/- 15.6	45.9 +/- 30.9	73.7 +/- 16.2	22.5 +/- 18.5
Responders / Non-responders (After 8 weeks)	24 / 26	16 / 18	7 / 10	2 / 4	49 / 58	-	-

Table 8.3 - Final Classification Model Performance with Reduced Feature Sets

MODEL	Features With …	T-TEST (Accuracy/Precision/Sensitivity/Specificity)	Meta Data (# Features Selected by Filter Method with ≥X votes; Cost and Gamma for SVM RBF kernel)
Baseline Model	≥ 50 votes	74.6/80.6/56.9/88.9	# Feats = 79; C = 2; Gamma = 0.0039
	≥ 60 votes	75.4/84.8/54.9/92.1	# Feats = 52; C = 2; Gamma = 0.0039
	≥ 70 votes	72.8/81.3/51.0/90.5	# Feats = 35; C = 0.5; Gamma = 0.0156
	≥ 80 votes	72.8/81.3/51.0/90.5	# Feats = 20; C = 2; Gamma = 0.0039
	≥ 90 votes	71.1/71.4/58.8/81.0	# Feats = 10; C = 2; Gamma = 0.0156
Week 2 Model	≥ 50 votes	74.8/77.5/63.3/84.5	# Feats = 57; C = 2; Gamma = 9.77e-4
	≥ 60 votes	72.0/74.4/59.2/82.8	# Feats = 37; C = 2; Gamma = 9.77e-4
	≥ 70 votes	69.2/66.7/65.3/72.4	# Feats = 20; C = 2; Gamma = 0.0156
	≥ 80 votes	70.1/67.3/67.3/72.4	# Feats = 14; C = 2; Gamma = 0.0156
	≥ 90 votes	73.8/70.6/73.5/74.1	# Feats = 4; C = 2; Gamma = 0.0156
Change (Baseline to Week 2) Model	≥ 50 votes	72.9/72.7/65.3/79.3	# Feats = 152; C = 2; Gamma = 9.77e-4
	≥ 60 votes	67.3/65.2/61.2/72.4	# Feats = 111; C =0.5; Gamma = 0.0039
	≥ 70 votes	70.1/70.7/59.2/79.3	# Feats = 79; C = 0.5; Gamma = 0.0039
	≥ 80 votes	65.4/63.6/57.1/72.4	# Feats = 55; C = 2; Gamma = 0.0039
	≥ 90 votes	68.2/68.3/57.1/77.6	# Feats = 33; C = 2; Gamma = 0.0156
Combined Model (Baseline + Change in Baseline to Week 2)	≥ 50 votes	83.2/87.8/73.5/91.4	# Feats = 242; C = 2; Gamma = 0.0039
	≥ 60 votes	85.0/92.3/73.5/94.8	# Feats = 170; C = 2; Gamma = 0.0039
	≥ 70 votes	83.2/87.8/73.5/91.4	# Feats = 110; C = 2; Gamma = 0.0039
	≥ 80 votes	80.4/86.8/67.3/91.4	# Feats = 69; C = 2; Gamma = 0.0156
	≥ 90 votes	81.3/82.2/75.5/86.2	# Feats = 39; C = 2; Gamma = 0.0156

Table 8.4 - Pooled, Within-Site and Between-Site Classification Results for Baseline Model

Train Site[a]	Test Site (Blue and Red) Excluded Site (Green)			
	UBC (n = 51)	**TGH (n = 38)**	**QNS (n = 18)**	**CAM (n = 7)**
UBC	BAC = 79.9% Accu = 80.4% Prec = 85.0% Sens = 70.8% Spec = 88.9%	BAC = 55.7% Accu = 57.9% Prec = 54.5% Sens = 35.3% Spec = 76.2%	BAC = 50% Accu = 61.1% Prec = 0% Sens = 0% Spec = 100%	BAC = 25% Accu = 28.6% Prec = 0% Sens = 0% Spec = 50%
TGH	BAC = 44.7% Accu = 45.1% Prec = 40.9% Sens = 37.5% Spec = 51.9%	BAC = 92.3% Accu = 92.1% Prec = 88.9% Sens = 94.1% Spec = 90.5%	BAC = 67.5% Accu = 66.7% Prec = 55.6% Sens = 71.4% Spec = 63.6%	BAC = 50% Accu = 57.1% Prec = 0% Sens = 0% Spec = 100%
Other sites[b]			BAC = 50% Accu = 61.1% Prec = 0% Sens = 0% Spec = 100%	BAC = 37.5% Accu = 42.9% Prec = 0% Sens = 0% Spec = 75%
Pool UBC and TGH; no test set	BAC = 70.2% Accu = 71.9% Prec = 83.3% Sens = 48.8% Spec = 91.7%			
Pool Train Sites (exclude test site)	BAC = 66.2% Accu = 66.7% Prec = 60.7% Sens = 62.9% Spec = 69.4%	BAC = 75.6% Accu = 77.6% Prec = 90.5% Sens = 55.9% Spec = 95.2%	BAC = 75.9% Accu = 77.1% Prec = 84.4% Sens = 61.4% Spec = 90.4%	BAC = 73.7% Accu = 75.7% Prec = 86.7% Sens = 54.2% Spec = 93.2%

Blue = Within-site classification
Red = Classification with separate train and test set
Green = Classification with pooled data from 3 sites (test set is left-out)

Training: randomized permutation cross-validation with 100 permutations
Testing: Performed using features found to be significant using the T-test Filtering Method and with >=70 votes (from training)

BAC = balanced accuracy**; Accu** = accuracy**; Prec** = precision**; Recall** = sensitivity**; Spec** = **specificity**

Table 8.5 - Pooled, Within-Site and Between-Site Classification Results for Week 2 Model

Train Site[a]	Test Site (Blue and Red) Excluded Site (Green)			
	UBC (n = 51)	**TGH (n = 38)**	**QNS (n = 18)**	**CAM (n = 7)**
UBC	BAC = 69.6% Accu = 70.0% Prec = 73.7% Sens = 58.3% Spec = 80.8%	BAC = 43.8% Accu = 44.1% Prec = 40.0% Sens = 37.5% Spec = 50.0%	BAC = 66.4% Accu = 70.6% Prec = 75.0% Sens = 42.9% Spec = 90.0%	BAC = 50% Accu = 66.7% Prec = 0% Sens = 0% Spec = 100%
TGH	BAC = 34.8% Accu = 36% Prec = 10% Sens = 4.17% Spec = 65.4%	BAC = 78.8% Accu = 79.4% Prec = 84.6% Sens = 68.8% Spec = 88.9%	BAC = 50% Accu = 58.8% Prec = 0% Sens = 0% Spec = 100%	BAC = 50% Accu = 66.7% Prec = 0% Sens = 0% Spec = 100%
Other sites[b]			BAC = 50% Accu = 58.8% Prec = 0% Sens = 0% Spec = 100%	BAC = 50% Accu = 66.7% Prec = 0% Sens = 0% Spec = 100%
Pool UBC and TGH; no test set	BAC = 74.8% Accu = 75% Prec = 75.7% Sens = 70% Spec = 79.5%			
Pool Train Sites (exclude test site)	BAC = 76.0% Accu = 78.9% Prec = 100% Sens = 52% Spec = 100%	BAC = 69.5% Accu = 71.2% Prec = 77.3% Sens = 51.5% Spec = 87.5%	BAC = 73.9% Accu = 74.4% Prec = 75.7% Sens = 66.7% Spec = 81.3%	BAC = 72.9% Accu = 73.3% Prec = 72.7% Sens = 68.1% Spec = 77.8%

Blue = Within-site classification
Red = Classification with separate train and test set
Green = Classification with pooled data from 3 sites (test set is left-out)

Training: randomized permutation cross-validation with 100 permutations
Testing: Performed using features found to be significant using the T-test Filtering Method and with >=70 votes (from training)

BAC = balanced accuracy**; Accu** = accuracy**; Prec** = precision**; Recall** = sensitivity**; Spec** = **specificity**

Table 8.6 - Pooled, Within-Site and Between-Site Classification Results for the 'Change from Baseline to Week 8' Model

Train Site[a]	Test Site (Blue and Red) Excluded Site (Green)			
	UBC (n = 51)	**TGH (n = 38)**	**QNS (n = 18)**	**CAM (n = 7)**
UBC	BAC = 80.1% Accu = 80.0% Prec = 76.9% Sens = 83.3% Spec = 76.9%	BAC = 51.7% Accu = 52.9% Prec = 50.0% Sens = 31.3% Spec = 72.2%	BAC = 50% Accu = 58.8% Prec = 0% Sens = 0% Spec = 100%	BAC = 50% Accu = 66.7% Prec = 0% Sens = 0% Spec = 100%
TGH	BAC = 54% Accu = 54% Prec = 52% Sens = 54.2% Spec = 53.8%	BAC = 84.7% Accu = 85.3% Prec = 92.3% Sens = 75.0% Spec = 94.4%	BAC = 50% Accu = 58.8% Prec = 0% Sens = 0% Spec = 100%	BAC = 50% Accu = 66.7% Prec = 0% Sens = 0% Spec = 100%
Other sites[b]			BAC = 52.1% Accu = 58.8% Prec = 50% Sens = 14.3% Spec = 90%	BAC = 50% Accu = 66.7% Prec = 0% Sens = 0% Spec = 100%
Pool UBC and TGH; no test set	BAC = 79.8% Accu = 79.8% Prec = 78.1% Sens = 80% Spec = 79.5%			
Pool Train Sites (exclude test site)	BAC = 79.8% Accu = 80.7% Prec = 81.8% Sens = 72.0% Spec = 87.5%	BAC = 87.2% Accu = 87.7% Prec = 90.0% Sens = 81.8% Spec = 92.5%	BAC = 80.9% Accu = 81.1% Prec = 80.5% Sens = 78.6% Spec = 83.3%	BAC = 70.5% Accu = 71.3% Prec = 73.7% Sens = 59.6% Spec = 81.5%

Blue = Within-site classification
Red = Classification with separate train and test set
Green = Classification with pooled data from 3 sites (test set is left-out)

Training: randomized permutation cross-validation with 100 permutations
Testing: Performed using features found to be significant using the T-test Filtering Method and with >=70 votes (from training)

BAC = balanced accuracy**; Accu** = accuracy**; Prec** = precision**; Recall** = sensitivity**; Spec** = **specificity**

Table 8.7 - Pooled, Within-Site and Between-Site Classification Results for the Combined Model (Baseline + 'Change from Baseline to Week')

Train Site[a]	Test Site (Blue and Red) Excluded Site (Green)			
	UBC (n = 51)	**TGH (n = 38)**	**QNS (n = 18)**	**CAM (n = 7)**
UBC	BAC = 90.1% Accu = 90% Prec = 88% Sens = 91.7% Spec = 88.5%	BAC = 55.2% Accu = 55.9% Prec = 53.8% Sens = 43.8% Spec = 66.7%	BAC = 50% Accu = 58.8% Prec = 0% Sens = 0% Spec = 100%	BAC = 50% Accu = 66.7% Prec = 0% Sens = 0% Spec = 100%
TGH	BAC = 55.9% Accu = 56% Prec = 54.2% Sens = 54.2% Spec = 57.7%	BAC = 91.3% Accu = 91.2% Prec = 88.2% Sens = 93.8% Spec = 88.9%	BAC = 75.7% Accu = 76.5% Prec = 71.4% Sens = 71.4% Spec = 80%	BAC = 75% Accu = 83.3% Prec = 100% Sens = 50% Spec = 100%
Other sites[b]			BAC = 50% Accu = 58.8% Prec = 0% Sens = 0% Spec = 100%	BAC = 50% Accu = 66.7% Prec = 0% Sens = 0% Spec = 100%
Pool UBC and TGH; no test set	BAC = 86.7% Accu = 86.9% Prec = 89.2% Sens = 82.5% Spec = 90.9%			
Pool Train Sites (exclude test site)	BAC = 82.6% Accu = 82.5% Prec = 77.8% Sens = 84.0% Spec = 81.3%	BAC = 81.6% Accu = 82.2% Prec = 83.3% Sens = 75.8% Spec = 87.5%	BAC = 89.6% Accu = 90.0% Prec = 94.6% Sens = 83.3% Spec = 95.8%	BAC = 87.8% Accu = 88.1% Prec = 90.7% Sens = 82.9% Spec = 92.6%

Blue = Within-site classification
Red = Classification with separate train and test set
Green = Classification with pooled data from 3 sites (test set is left-out)

Training: randomized permutation cross-validation with 100 permutations
Testing: Performed using features found to be significant using the T-test Filtering Method and with >=70 votes (from training)

BAC = balanced accuracy**; Accu** = accuracy**; Prec** = precision**; Recall** = sensitivity**; Spec** = **specificity**

8.8 Figures

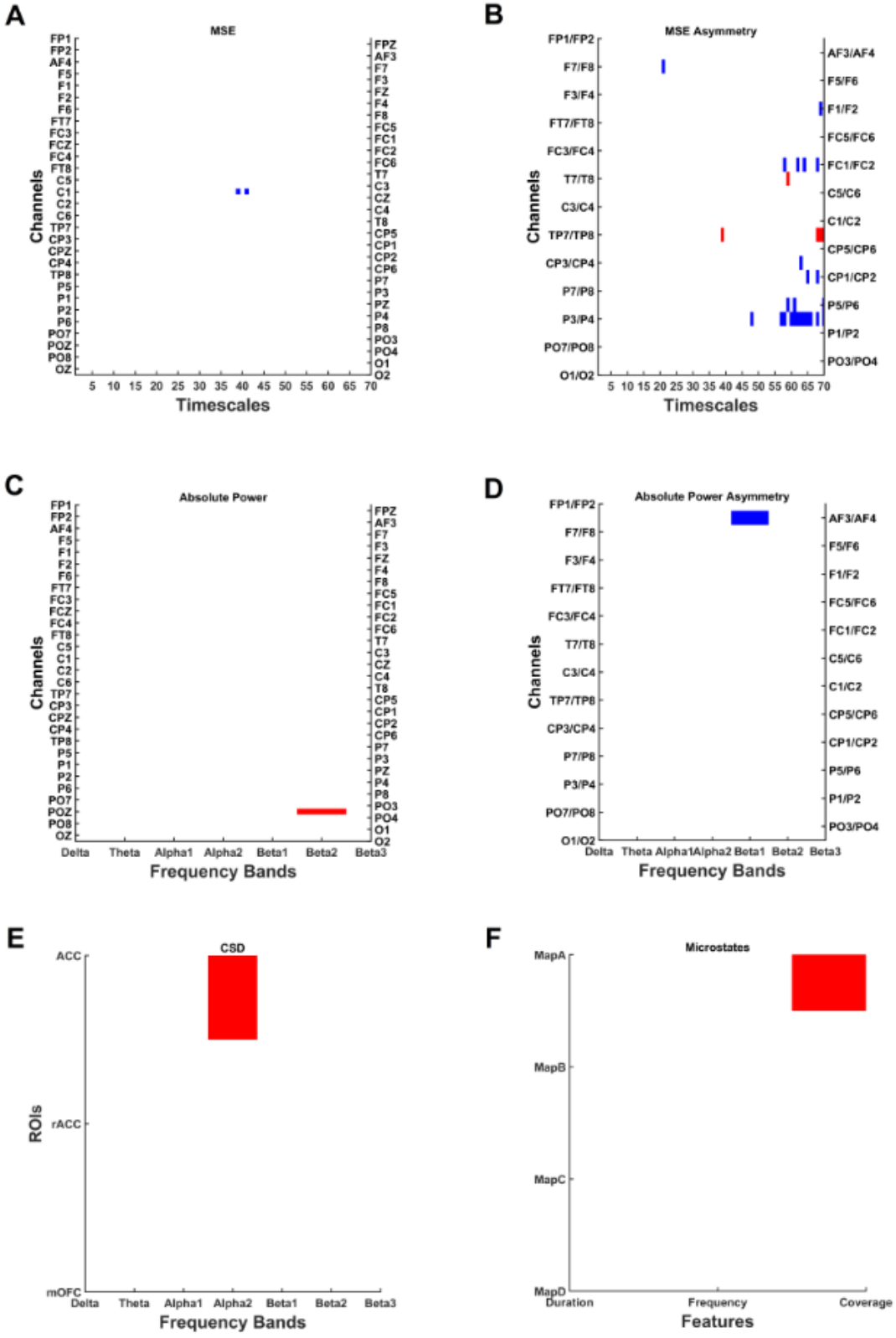

A
MSE
Channels
Timescales
B
MSE Asymmetry
Channels
Timescales
C
Absolute Power
Channels
Delta Theta Alpha1 Alpha2 Beta1 Beta2 Beta3
Frequency Bands
D
Absolute Power Asymmetry
Channels
Delta Theta Alpha1 Alpha2 Beta1 Beta2 Beta3
Frequency Bands
E
CSD
ROIs
ACC
rACC
mOFC
Delta Theta Alpha1 Alpha2 Beta1 Beta2 Beta3
Frequency Bands
F
Microstates
MapA
MapB
MapC
MapD
Duration Frequency Coverage
Features

Figure 8.1 - Baseline features discriminating responders and non-responders with ≥ 70 votes following the randomized permutation cross-validation method.

Red indicates that the feature value is greater in responders than non-responders and blue indicates that the feature is smaller in responders than non-responders. **(A)** Features taken from multiscale entropy analysis. X-axis represents time scales from 1 to 70 and y-axis represents EEG channels. **(B)** Asymmetry features from multiscale entropy analysis. X-axis represents time scales from 1 to 70 and y-axis represents the channels used to calculate asymmetry. **(C)** Features taken from EEG power analysis. X-axis represents frequency: delta (1-3.5Hz); theta (4-8Hz); low-alpha (8.5-10Hz); high-alpha (10.5-12Hz); low-beta (12.5-18Hz); mid-beta (18.5-21Hz); and high-beta (21.5-30Hz). Y-axis represents EEG channels. **(D)** Asymmetry features from power analysis. X-axis represents the frequency bands and y-axis represents the channels used to calculate asymmetry. **(E)** Features taken from current source density analysis. X-axis represents the frequency bands. Y-axis represents the regions of interest: the anterior cingulate cortex (ACC), rostral ACC (rACC), superior frontal gyrus (SFG), middle frontal gyrus (MiddleFG) and the medial frontal gyrus (MedialFG). **(F)** Features taken from microstate analysis. X-axis represents the characteristics of microstates: duration, frequency and coverage. Y-axis represents the microstate map: A, B, C and D.

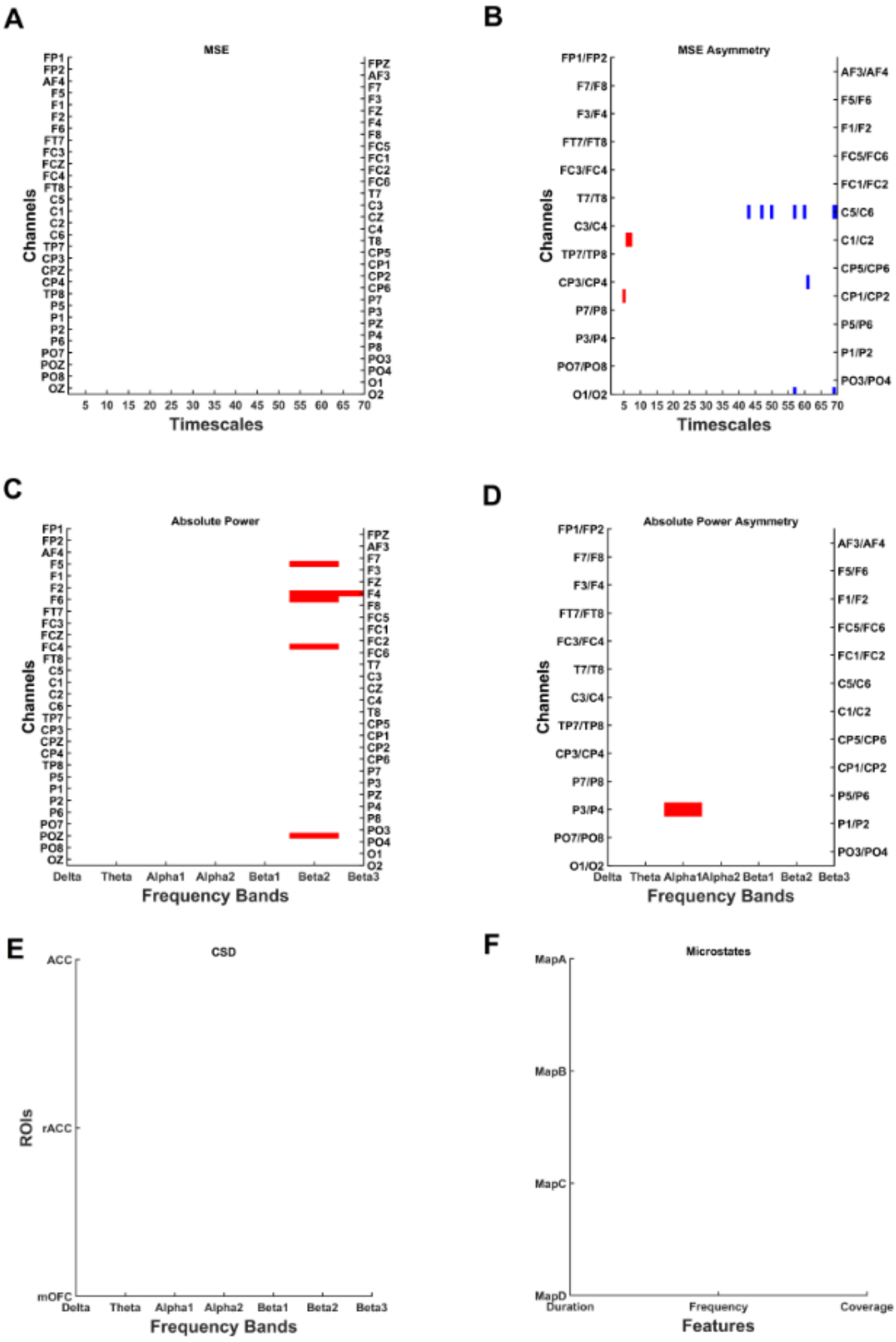

A
MSE
Channels
Timescales
B
MSE Asymmetry
Channels
Timescales
C
Absolute Power
Channels
Delta Theta Alpha1 Alpha2 Beta1 Beta2 Beta3
Frequency Bands
D
Absolute Power Asymmetry
Channels
Delta Theta Alpha1 Alpha2 Beta1 Beta2 Beta3
Frequency Bands
E
CSD
ROIs
ACC
rACC
mOFC
Delta Theta Alpha1 Alpha2 Beta1 Beta2 Beta3
Frequency Bands
F
Microstates
MapA
MapB
MapC
MapD
Duration Frequency Coverage
Features

Figure 8.2 - Features from week 2 discriminating responders and non-responders with ≥ 70 votes following the randomized permutation cross-validation method.

Red indicates that the feature value is greater in responders than non-responders and blue indicates that the feature is smaller in responders than non-responders. **(A)** Features taken from multiscale entropy analysis. X-axis represents time scales from 1 to 70 and y-axis represents EEG channels. **(B)** Asymmetry features from multiscale entropy analysis. X-axis represents time scales from 1 to 70 and y-axis represents the channels used to calculate asymmetry. **(C)** Features taken from EEG power analysis. X-axis represents frequency: delta (1-3.5Hz); theta (4-8Hz); low-alpha (8.5-10Hz); high-alpha (10.5-12Hz); low-beta (12.5-18Hz); mid-beta (18.5-21Hz); and high-beta (21.5-30Hz). Y-axis represents EEG channels. **(D)** Asymmetry features from power analysis. X-axis represents the frequency bands and y-axis represents the channels used to calculate asymmetry. **(E)** Features taken from current source density analysis. X-axis represents the frequency bands. Y-axis represents the regions of interest: the anterior cingulate cortex (ACC), rostral ACC (rACC), superior frontal gyrus (SFG), middle frontal gyrus (MiddleFG) and the medial frontal gyrus (MedialFG). **(F)** Features taken from microstate analysis. X-axis represents the characteristics of microstates: duration, frequency and coverage. Y-axis represents the microstate map: A, B, C and D.

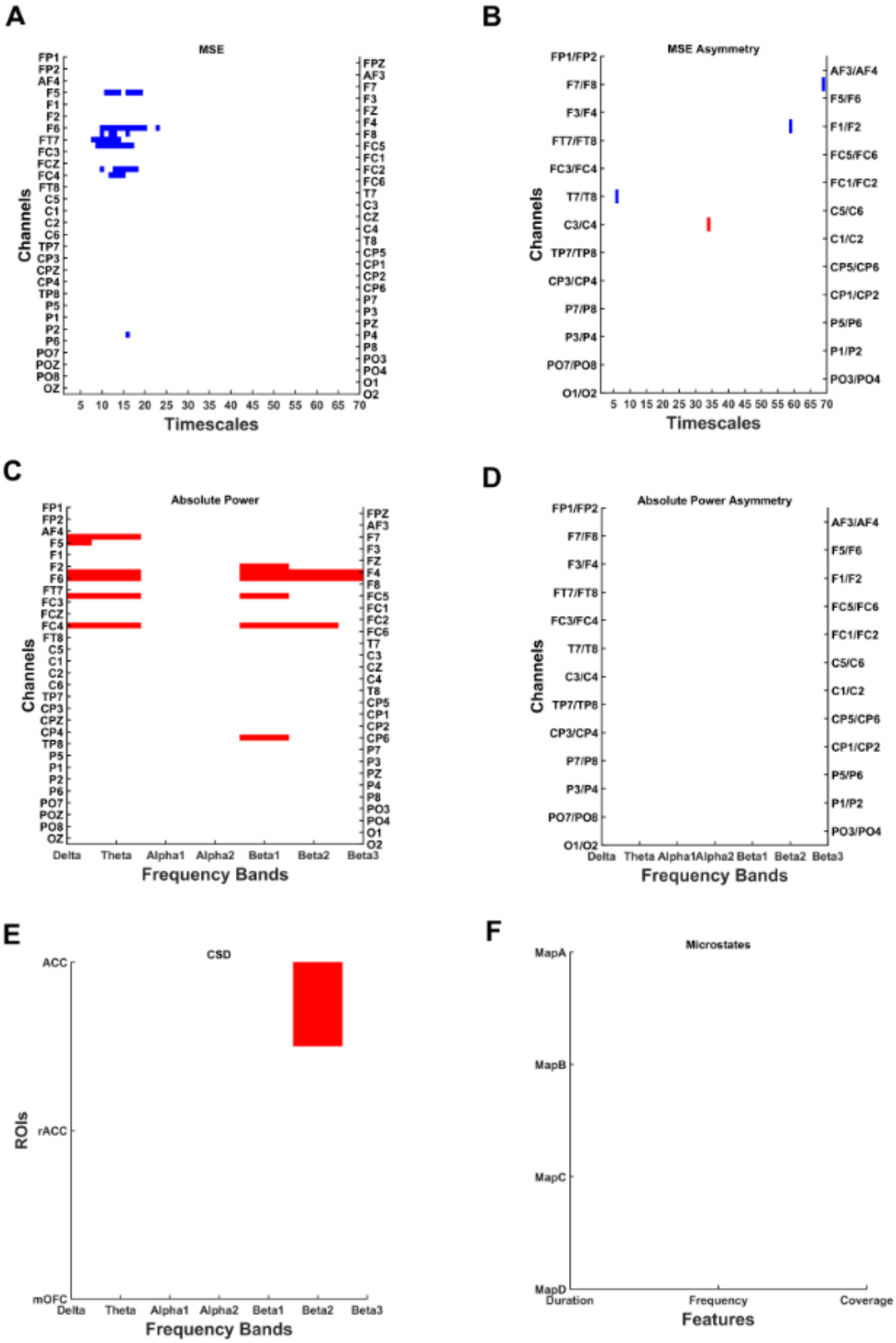

A
MSE
Channels
Timescales
B
MSE Asymmetry
Channels
Timescales
C
Absolute Power
Channels
Frequency Bands
D
Absolute Power Asymmetry
Channels
Frequency Bands
E
CSD
ROIs
Frequency Bands
F
Microstates
Features

Figure 8.3 - 'Change from Baseline to Week 2' features discriminating responders and non-responders with ≥ 70 votes following the randomized permutation cross-validation method.

Red indicates that the feature value is greater in responders than non-responders and blue indicates that the feature is smaller in responders than non-responders. **(A)** Features taken from multiscale entropy analysis. X-axis represents time scales from 1 to 70 and y-axis represents EEG channels. **(B)** Asymmetry features from multiscale entropy analysis. X-axis represents time scales from 1 to 70 and y-axis represents the channels used to calculate asymmetry. **(C)** Features taken from EEG power analysis. X-axis represents frequency: delta (1-3.5Hz); theta (4-8Hz); low-alpha (8.5-10Hz); high-alpha (10.5-12Hz); low-beta (12.5-18Hz); mid-beta (18.5-21Hz); and high-beta (21.5-30Hz). Y-axis represents EEG channels. **(D)** Asymmetry features from power analysis. X-axis represents the frequency bands and y-axis represents the channels used to calculate asymmetry. **(E)** Features taken from current source density analysis. X-axis represents the frequency bands. Y-axis represents the regions of interest: the anterior cingulate cortex (ACC), rostral ACC (rACC), superior frontal gyrus (SFG), middle frontal gyrus (MiddleFG) and the medial frontal gyrus (MedialFG). **(F)** Features taken from microstate analysis. X-axis represents the characteristics of microstates: duration, frequency and coverage. Y-axis represents the microstate map: A, B, C and D.

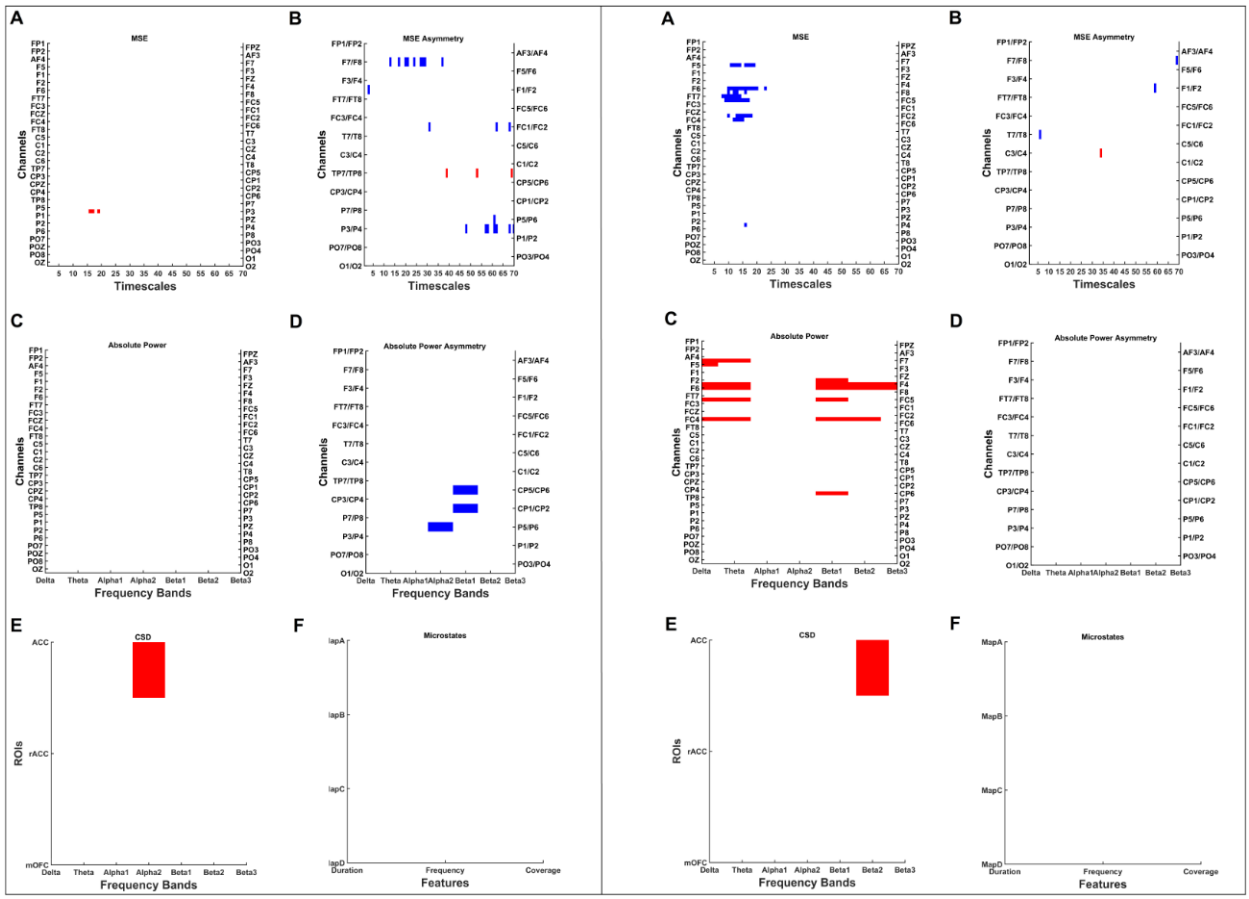


Figure 8.4 - Combined model (Baseline + 'Change from Week 0 to Week 2') features discriminating responders and non-responders selected in ≥ 70 permutations of the randomized permutation cross-validation method.

Red indicates that the feature value is greater in responders than non-responders and blue indicates that the feature is smaller in responders than non-responders. Left panel represents baseline features. Right panel represents 'Change from Week 0 to Week 8') features. **(A)** Features taken from multiscale entropy analysis. X-axis represents time scales from 1 to 70 and y-axis represents EEG channels. **(B)** Asymmetry features from multiscale entropy analysis. X-axis represents time scales from 1 to 70 and y-axis represents the channels used to calculate asymmetry. **(C)** Features taken from EEG power analysis. X-axis represents frequency: delta (1-3.5Hz); theta (4-8Hz); low-alpha (8.5-10Hz); high-alpha (10.5-12Hz); low-beta (12.5-18Hz); mid-beta (18.5-21Hz); and high-beta (21.5-30Hz). Y-axis represents EEG channels. **(D)** Asymmetry features from power analysis. X-axis represents the frequency bands and y-axis represents the channels used to calculate asymmetry. **(E)** Features taken from current source density analysis. X-axis represents the frequency bands. Y-axis represents the regions of interest: the anterior cingulate cortex (ACC), rostral ACC (rACC), superior frontal gyrus (SFG), middle frontal gyrus (MiddleFG) and the medial frontal gyrus (MedialFG). **(F)** Features taken from microstate analysis. X-axis represents the characteristics of microstates: duration, frequency and coverage. Y-axis represents the microstate map: A, B, C and D.

Section V: Discussion and Future Work

Chapter 9 – General discussion and Suggestions for Future Work

There were two main aims in this thesis. One was to investigate the neurophysiological targets of treatments for depression (seizure therapy and antidepressant medications) to understand (a) why seizure therapy is effective for severe, treatment-resistant depression, and (b) why patients might fail to respond to an antidepressant medication, i.e., usually the first line of treatment for depression. The second aim was to build a predictive model to identify responders and non-responders of antidepressant medication, early in the treatment trial. This can reduce the time spent in failed trials and minimize the effect of untreated depression. Studies included in this thesis provided evidence for the modulation of neural dynamics following seizure therapy (electroconvulsive therapy and magnetic seizure therapy – ECT and MST) and antidepressant medication (escitalopram). To characterize neural dynamics, we derived resting-state EEG measures from the frequency domain using power spectral density analysis (*seizure therapy:* Chapter 4; *escitalopram:* Chapter 6), spatio-temporal measures of complexity using multiscale entropy analysis (*seizure therapy:* Chapter 4; *escitalopram:* Chapter 6), and global brain-network measures using microstate analysis (*seizure therapy:* Chapter 5; *escitalopram:* Chapter 7). Prediction models were built to predict response to escitalopram using machine learning methods (Chapter 8).

Results from this thesis suggest that both antidepressant medications and seizure therapy may modulate cerebral oscillatory activity for therapeutic effect. Markers associated with antidepressant efficacy were identified to be distinct from markers of cognitive decline seen with ECT using multiscale entropy analysis. In addition, changes in neural dynamics were widespread in frequency, timescales and regions following ECT. Changes following MST were focal and were observed in high frequency, fine timescales, and fronto-central and parieto-occipital regions. Finally, the trajectory of changes in neural dynamics following eight weeks of escitalopram treatment (i.e., baseline to week 2 and week 2 to week 8) was shown to be different between responders and non-responders of escitalopram. Several measures of neural dynamics at baseline and as early as 2 weeks into treatment, were shown to predict response to escitalopram. The mechanistic and predictive markers identified through this work are summarized in **Appendices II-III.**

This chapter is organized to first summarize and discuss the results of studies included in this thesis that aimed to investigate the targets of treatments for depression (ECT, MST and escitalopram). Then, predictive markers of escitalopram response are summarized and discussed. Next, based on the lessons learned from our work, recommendations for future intervention trials are provided. Finally, limitations, novel findings and implications of this thesis are also discussed.

9.1 Insight into the Mechanism of Action of Treatments for Depression

9.1.1 Summary of Results

Considering the complex pathophysiology of major depression (Tardito, Perez et al. 2006, Belmaker and Agam 2008, Krishnan and Nestler 2010, Duman and Aghajanian 2012, Pandya, Altinay et al. 2012), the mechanism of action of successful treatments is also likely to be complex. For the scope of this thesis, we focused on resting-state neural dynamics at the cortex level, as monitored by EEG, for several reasons. First, EEG neural dynamics are considered to be intermediate measures that can capture changes at both the neuronal and network level of the brain (Zoon, Veth et al. 2013, Leuchter, Hunter et al. 2014). In addition, predictive markers derived from the EEG signal have clinical potential because unlike other neuroimaging methods (i.e., fMRI), EEG is inexpensive and accessible. Finally, seizure therapy is known to impact neural dynamics through the induction of a seizure and this impact was clearly shown with resting-state EEG data (Krystal, Weiner et al. 1993, Krystal, Greenside et al. 1996, Sackeim, Luber et al. 1996, Krystal, Coffey et al. 1998, Perera, Luber et al. 2004, Okazaki, Takahashi et al. 2013). In our work, we investigated whether the modulation of neural dynamics by seizure therapy is associated with its therapeutic effect and if other non-seizure therapies (i.e., antidepressant medications) also show therapeutic effect by modulation of neural dynamics. Results are summarized below. Please see **Appendix II** for a summary of the mechanistic markers of escitalopram, ECT and MST identified through all the studies included in this thesis.

9.1.1.1 Summary of Resting-State Neural Dynamics Associated with Response to Seizure Therapy (Chapters 4-5)

Seizure therapy was shown to have a significant impact on resting-state neural dynamics as measured by EEG. This impact on neural dynamics was also shown to be different between responders of ECT and responders of MST. The widespread changes in frequency of neural oscillations observed following ECT were not specific to responders of ECT. In other words, changes were observed in both responders and non-responders. In comparison, an increase in high frequency oscillations was observed in only responders of MST. However, changes in the frequency of neural oscillations were not correlated with improvement in depressive symptoms. Studying the complexity of neural dynamics, responders of both ECT and MST revealed spatio-temporal specific changes in complexity. A decrease in the complexity of fine timescales (1-30) was observed in responders of ECT and MST, and this decrease in complexity was linked to improvement in depressive symptoms. In addition, an increase in complexity of coarse timescales (35-70) was observed in responders of ECT and this increase in complexity was linked with cognitive decline associated with ECT. Finally, through the examination of global network dynamics, only ECT responders revealed significant network-specific changes associated with treatment response. The effect on global network dynamics by MST was shown to be in the same direction as ECT but the effect was not network-specific and was not seen in responders or non-responders of the treatment.

9.1.1.2 Summary of Resting-State Neural Dynamics Associated with Response to Escitalopram (Chapters 6-7)

Significant changes in resting-state neural dynamics were observed following an eight week course of escitalopram, as measured by EEG. Results from frequency and complexity analysis revealed a distinct pattern of early (baseline to week 2) and late (week 2 to week 8) changes. In responders, early changes in the frequency of neural oscillations were localized to the parieto-occipital regions and were frequency-specific (13-30Hz). Over the remaining course of treatment (week 2 to week 8), responders revealed changes in frequency and complexity that correlated with improvement in depressive symptoms. In non-responders, early changes were seen in several regions, frequencies and timescales and these changes negatively correlated with improvement of depressive symptoms. Following this early change, there was no effect of

treatment on neural dynamics in non-responders from week 2 to week 8. Results from microstate dynamics suggested that escitalopram may have an impact on global brain dynamics. Initial changes in global brain-network dynamics by escitalopram were associated with non-response towards the treatment. Together, results suggested that a potential early perturbance of neural activity and the dynamics of neural activity by escitalopram (that is frequency, spatio-temporal and network specific) may have a positive effect on responders over the course of treatment and an early negative effect on non-responders.

9.1.2 How are Neural Dynamics Related to the Mechanism of Action of Treatments for Depression?

Before speculating on the role of dynamic neural activity in mediating the mechanisms underlying treatments for depression, it is important to highlight the importance of neural dynamics in the complex functions of the brain. Effective neural communication and potentially neuroplasticity rely on the efficient transmission of information through both bottom-up and top-down processes (Rolls, Treves et al. 1998, Buzsáki and Draguhn 2004, Eytan and Marom 2006, Grillner 2006, Rojas 2013). Bottom-up processes are involved in neural communication from the level of neurons to networks, where the summation of individual neurons firing together can generate neural oscillations at the network level. In comparison, top-down processes are involved in neural communication from the level of networks to neurons, where the electric field generated by network oscillations can induce electrical activity at individual neurons (Leuchter, Hunter et al. 2015). Neural dynamics are characteristics of such bi-directional processes responsible for regulating brain functions (Laughlin and Sejnowski 2003, Buzsáki and Draguhn 2004, Schnitzler and Gross 2005) including neuroplasticity (Leuchter, Hunter et al. 2015). Modulation of neuroplasticity is one of the most common findings in studies investigating the impact of treatments for depression on the brain (D'sa and Duman 2002, Brunoni, Lopes et al. 2008, Pittenger and Duman 2008).

Through the results of this work, we suggest that treatments for depression may modulate the dynamics of neural communication. Specifically, we suggest that successful treatments for depression, such as ECT, impact dynamics at the large-scale network level of the brain for

antidepressant response. In a recent meta-analysis (Kaiser, Andrews-Hanna et al. 2015), depression was associated with functional brain network dysfunction, specifically, hypoconnectivity in the frontoparietal network and hyperconnectivity in the default-mode network. In our work, modulation of complexity and frequency of neural oscillations were observed in regions associated with the default-mode network (precuneus, posterior cingulate cortex) as well as the frontoparietal network (anterior cingulate cortex, prefrontal and parietal cortex). Together, these results add to the current line of neuroimaging evidence for the disruption and modulation of large-scale neural networks following treatments of depression for therapeutic efficacy.

Pharmacotherapy may also modulate neural dynamics at the network level for therapeutic efficacy through its effects at the cellular level. Based on the results from our work, this may take more than 2 weeks. In addition, genetic factors may delay the effects of pharmacotherapy at the cellular level (Evans and Relling 1999) and further delay the effect of pharmacotherapy at the network level. In our results, we repeatedly observed that escitalopram does not show an early effect on responders for therapeutic effect. This has also been shown in previous studies, where the effects of pharmacotherapy on depressive symptoms are delayed for 4-6 weeks (Quitkin *et al.*, 1984; Gelenberg and Chesen, 2000; Frazer and Benmansour, 2002). We hypothesize that this delay may be associated with the time it might take cellular effects to translate to the network level.

A novel finding of our work was that pharmacotherapy may have an initial impact in non-responders suggesting that the initial effect of pharmacotherapy may be different between responders and non-responders. One possible explanation may be interpreted from the known effects of pharmacotherapy. The mechanism of action of SSRIs (such as escitalopram) is generally assumed to be mediated by the binding of the drug to serotonin transporters resulting in increased extracellular concentrations of serotonin across the serotonergic pathways (Stahl 1998). However, a recent PET study in humans (Nord, Finnema et al. 2013) reported that a single dose of escitalopram leads to an initial decrease in serotonin concentrations in occipital and temporal regions. Over the course of the medication, there may be a desensitization of inhibitory serotonin receptors and serotonin concentrations increase as expected. We hypothesize

that the neurobiological effects of this initial decrease in serotonin concentrations may differ between responders and non-responders.

At the gene level, brain-derived neurotrophic factor (BDNF) is known to be involved in neural cell proliferation and synaptic plasticity (Katz and Shatz 1996, Duman and Monteggia 2006). Reduced secretion of BDNF by the genetic polymorphism BDNF Val66Met, was previously reported to be associated with severity of depression (Jiang, Xu et al. 2005, Verhagen, Van Der Meij et al. 2010, Molendijk, Bus et al. 2011, Czira, Wersching et al. 2012) and normalization of BDNF levels was associated with antidepressant response (Sen, Duman et al. 2008, Zou, Ye et al. 2010, Molendijk, Bus et al. 2011). In addition, reduced secretion of BDNF was hypothesized to be associated with reduced functional connectivity and oscillatory activity in neuronal systems impaired in depression (Thomason, Yoo et al. 2009). Specifically, polymorphism of BDNF Val66Met was shown to be mediated by parieto-occipital alpha (Gatt, Kuan et al. 2008, Zoon, Veth et al. 2013) and increased parieto-occipital alpha was shown to be higher in patients with depression (Pollock and Schneider 1990, Bruder, Tenke et al. 2005). In our work, responders of escitalopram revealed a reduction in alpha oscillations after 8 weeks of escitalopram treatment while non-responders revealed no changes in alpha power. Based on these results and the previous literature, we hypothesize that non-responders may show impairments in the biological pathway including BDNF *Val66Met*, that are not seen in responders, resulting in differences in their response to escitalopram.

Complexity and oscillatory power are important measures of neural dynamics used in this thesis. Previous studies have suggested that these measures may provide insight into neuroplastic processes. High alpha oscillatory power at rest and during cognitive tasks has been correlated with BDNF *Val66Met* polymorphism that affects BDNF secretion (Gatt, Kuan et al. 2008, Zoon, Veth et al. 2013). Complexity of time series in biological systems is suggested to reflect plasticity (i.e., adaptability) to a changing environment (McIntosh, Vakorin et al. 2013). Recent work suggests that the transfer of information from one neural network to another (i.e., effective neural transmission) requires that the two networks have matching complexities (West and Grigolini 2010, Marmelat and Delignières 2012, Mafahim, Lambert et al. 2015). In general, the

dynamic activity of neurons and neuronal populations, i.e., networks and circuits, plays a key role in regulating neural functions such as neuroplasticity. The observation of such dynamics has proved to be useful in understanding the targets of successful treatments for depression.

If all treatments for depression modulate neural dynamics, why do some treatments show efficacy when others fail?

Due to the heterogeneity of depression, regions affected by depression may vary between each individual patient (Pandya, Altinay et al. 2012). With the current trial-and-error approach to treatment selection, the selected treatment may not be able to target the specific impaired region(s), or the impact on the region(s) affected by depression by the selected treatment may not be sufficient enough to generate a long-lasting change in neuroplasticity required for therapeutic response. For example, ECT has the highest efficacy for patients with severe, treatment-resistant depression (Kellner, Greenberg et al. 2012). It has a very strong impact on global neural dynamics, affecting regions impaired in depression (Farzan, Boutros et al. 2014) but also other regions associated with cognition resulting in cognitive side effects (Devanand, Sobin et al. 1995). On the other hand, MST has a focal effect (Deng, Lisanby et al. 2011). Although MST has the potential to only impact regions impaired in depression (Lisanby, Luber et al. 2003, Spellman, McClintock et al. 2008, Deng, McClintock et al. 2015), selecting an appropriate treatment target (region or network) can be challenging without clear knowledge on the pathophysiology of depression. Finally, antidepressant medications may also have a region-specific effect rather than the widespread effects seen with ECT (Altar, Whitehead et al. 2003). In other words, they may not be able to target impaired connectivity between these regions for therapeutic efficacy.

9.2 Predicting Response to Escitalopram Treatment using Markers of Neural Dynamics

9.2.1 Summary of Results (Chapter 8)

Predictive markers, that can provide insight into therapeutic effectiveness prior to the start of the treatment course, could potentially assist in the clinical decision-making process of treatment selection. In **Appendix III**, the predictive markers of pharmacotherapy (escitalopram) identified

through this thesis are summarized. In the future, the treatment selection process is hoped to be improved through the use of an EEG-based clinical decision-making tool integrating such markers. As a first step towards the development of this tool, we evaluated the feasibility of predicting response to escitalopram treatment using measures of resting-state neural dynamics, as obtained from EEG data.

To recap, prediction models were built using data from patients with major depressive disorder who received 8 weeks of escitalopram treatment. Four models were created for prediction: (1) baseline model using only features derived from the data collected during baseline, (2) week 2 model using only features derived from the data collected during week 2, (3) early change model including change in features from baseline to week 2, and (4) a combined model using data from (1) and (3).

Balanced classification accuracy was approximately 72.8% using baseline data, 69.2% using week 2 data and 70.1% using 'change from baseline to week 2' data. In addition, as hypothesized, combining baseline neural dynamics with early changes in neural dynamics (change in response to 2 weeks of treatment) increased the accuracy of prediction to 83.2%. Please see Chapter 8 for a detailed report.

9.2.2 Comparing Prediction Performance with Previous Studies

Given the large amount of clinical, behavioral and demographic data collected during a treatment trial, there has been interest in whether this data can be used in the clinic to guide treatment decisions. Studies have shown that patient characteristics at baseline may be associated with treatment outcome (Trivedi, Rush et al. 2006, Rush, Wisniewski et al. 2008). However, their predictive performance was shown to be low. The sequential treatment alternatives to relieve depression (STAR*D) was the largest trial to evaluate the predictive value of patient characteristics. In this trial, prediction using a logistic regression model achieved an AUC of 0.71 (Perlis 2013). Prediction using machine learning methods and external validation with an independent dataset yielded an accuracy of 59.6% (Perlis 2013).

In Chapter 1, a table was provided **(Table 1.1**) to summarize studies evaluating the prediction performance of pre-treatment resting-state EEG markers for antidepressant response. Of those studies, only two studies provided accuracy higher than our study (83.2%). An accuracy of 87.9% was reported (sensitivity=94.9%; specificity=80.9%) in (Khodayari-Rostamabad, Reilly et al. 2013). However, this study had several limitations: (i) the study was performed with a low sample size (n=22), (ii) responders were defined with ≥30% improvement in clinical scores rather than the usual ≥50%, and (iii) patients on 4 different medications were combined (sertraline, citalopram, fluvoxamine, paroxetine). Accuracy was also high (85-92%) in (Rabinoff, Kitchen et al. 2011) using spectral EEG features with classification and regression tree analysis. The study combined trials for 2 antidepressants (fluoxetine and venlafaxine) to predict response in 51 patients with unipolar depression. The high accuracy values however, may be due to overfitting to the data and this is suggested by the 100% specificity in all treatment groups. Apart from these studies, most of the literature focused on the evaluation of single marker for predictive value. The accuracy of prediction varies using single features varies between 60-77%. However, since most of these studies did not perform validation, or performed analysis on fairly low sample sizes, accuracies may be inflated and may decrease significantly when applied to a new dataset.

9.3 Importance of Validation for Generalizability of Results

The mechanistic and predictive markers identified in this work as well as the developed predictive models require validation for several reasons. First, the reliability of the markers and the models is yet to be determined. Replication of results would be an important step towards as assessment of reliability. Second, it is important to determine whether the studied changes in neural dynamics are truly associated with treatment or if they may be a result of other confounding effects. For this, future studies should consider the use of a randomized design with a placebo patient group. Studies should also aim to investigate the physiological meaning and implications of these results for depression and for treatment outcome. Finally, the pipeline used to derive prediction models in this work also needs to be validated before its use and application. All these validation tests serve to evaluate the generalizability of the markers and the models.

Validation of performance with an independent dataset is another important aspect of classification models. Yet, it was not implemented in several previous studies. It should be noted that by excluding validation, results on the predictive performance of a marker or a model are potentially inflated and highly unlikely to generalize to a new data set that is independent from the original data set. In this study, cross-validation was applied to select hyperparameters for the machine learning algorithm as well as to estimate the predictive performance of the model. Despite the advantages of cross-validation, it still only provides an estimate on the generalizability of the model. It is important to further validate each model on a larger independent data set for accurate estimates of prediction performance.

One important reason to validate results is the potential sources of bias from the patient sample size. Given the heterogeneity of depression, any size of sample used in research trials suffers from selection bias. Patients may vary in demographics (age, education, previous history of mental illness), severity of illness and/or genetic differences that could potentially affect an individual's response to treatment. In addition, most research trials do not account for the large portion of individuals that do not seek medical attention and therefore, do not receive treatment. Patients could also drop out of the study or not provide consent to participate in neurophysiological assessments. The effect of excluding these subjects can have a severe impact on the generalizability of study results and must be considered when interpreting results.

Therefore, several factors need to be considered in future validation studies. First, efforts should be made for a comprehensive inclusion of the patient group. If possible, patients in the treatment group should be matched with patients in the placebo group. The allocation of patients in either group should be based on a random process and a strict implementation of this random assignment must be followed. Next, the same methodology for data collection, pre-processing and analysis should be followed to ensure there is no bias in the study design. Documentation of deviations to the methodology is highly recommended for transparency.

9.4 Importance of Integrating Markers from Multiple Time Points in Prediction Models

The investigation of predictive markers for escitalopram response over multiple time points proved to be an advantage in our work. In our second project (Chapters 6-7), the trajectory of neurophysiological changes over time was observed to be different between responders and non-responders. This difference in trajectory was also crucial for the identification of early predictors of response and non-response. In our third project (see Chapter 8), prediction models that included measures of changes in neural dynamics from baseline to week 2 were shown to yield higher performance.

In general, frequent neurophysiological assessments are highly recommended to better understand treatments for depression and to improve prediction for treatment outcome in depression. The inclusion of data from several time points would allow for a powerful statistical design that can better detect the effect of time and the interaction between time and other factors, the most important being response. In addition to this advantage, longitudinal studies would also provide a better basis for causal inference. If assessments are performed over several time points during the course of a treatment, we can better understand which effects are specific to treatment and which effects may be associated with confounding factors.

For prediction, it is important to identify the earliest time point at which response or non-response can be detected. As a general guideline and based on the results from this work, neurophysiological assessments should be conducted at least 2 weeks into treatment if the detection of non-response is of priority. If possible, future studies may include assessments 1 week into the treatment course to investigate whether markers of response/non-response can be detected earlier than 2 weeks as suggested by some previous studies (Katz, Koslow et al. 1996, Stassen, Angst et al. 1999). Other studies agree with our results, suggesting that medications may require a minimum of 2-3 weeks to show some improvement (Quitkin, Rabkin et al. 1984, Gelenberg and Chesen 2000). The inclusion of neurophysiological assessments at regular time points after 2 weeks however, could also inform on the gradual changes that occur with treatment. The frequency of these data collection time points would also depend on the

availability of resources (cost and personnel), patient consent and comfort, reliability of data collection procedures, etc.

9.5 Importance of Standardization of Data for Future Work

The feasibility of identifying reliable mechanistic and predictive markers of treatments and the development of a clinical decision-making tool for treatment selection will in part depend on the collection and sharing of large-scale, clinically-reliable data sets, as done by CAN-BIND. Prior to the collection of these large data sets, a systematic procedure must be set in place for proper standardization of data collection, analysis, and handling. With the involvement of several sites, investigators and project initiatives, data may be collected over several modalities as well. Reproducibility of results and combination of data from different modalities (clinical, neuroimaging, genetic) will also largely depend on this standardization procedure as well as the data collection and data sharing procedures. In case the standardization procedure cannot be followed, investigators should at least document all deviations from the standardization protocol. With proper standardization and documentation, results derived from these large-scale studies are more likely to be reproducible and generalizable to a large portion of the patient population.

9.6 Suggestions for Future Intervention Trials

Several recommendations are listed here and are based on the lessons learned from studies included in this thesis.

1) As mentioned, to investigate the targets of successful treatments for depression, clinical and neurophysiological assessments may need to be conducted at frequent time points. It is recommended that for an initial pilot study or for an early portion of a large clinical trial, these assessments should be conducted regularly (e.x., every treatment session). The frequency of assessments may be reduced after initial analysis of the collected data. The main aim would be to identify the earliest time point at which we may be able to detect differences between responders and non-responders. Another aim would be to monitor the variability of neural measures between each time point. This could determine the

maximum time interval that should be set in trials for neurophysiological assessments so that reliable changes associated with treatment outcome can be detected rather than noise.

2) In addition to EEG neurophysiological assessments, it could be highly valuable to conduct additional physiological assessments to include markers from different levels of biological organization (such as genetic, molecular) in future prediction models. It is likely that these factors are linked to different bio-types of patients with depression (Wager and Woo 2017) and how they respond to treatment.

3) If possible, a double-blind, placebo-controlled trial should be considered. The inclusion of a patient group receiving a placebo may reduce the confounding effects of bio-types of depression in evaluating clinical efficacy (Wager and Woo 2017) and also validating the measures of neural dynamics that were associated with clinical outcome in our work.

4) To provide the best treatment for each individual patient and move towards personalized medicine for depression, there is a need to first identify neurophysiological markers that can reliably predict whether a patient will respond to one type of treatment over another. These markers may be different or similar to the ones identified in this work. A future multi-intervention trial must be conducted to elucidate these relationships. Such a trial should also include a patient placebo group.

5) Several steps need to be taken to test the feasibility of translating the predictive markers of pharmacotherapy from this thesis into the clinic. First, as mentioned, results from this thesis should be validated on a large, independent patient dataset and also data collected from a patient placebo group. Then, the test-retest reliability of all these markers must be evaluated with a healthy subject group. Finally, for a portable clinical solution, the test-retest reliability study should also investigate the minimum of number of EEG electrodes, and the best montage of the electrodes, required for a reliable measure of neural dynamics. Based on results from this thesis, a fronto-parietal montage may work best.

9.7 Suggestions for Improving the Performance of Prediction Models for Treatment Outcome

Several future directions can be explored to increase the accuracy of treatment outcome prediction.

1) Future studies should consider integrating measures from several different modalities (i.e., genetic, fMRI, PET) as well as from additional time points during the course of the treatment. In this thesis, we chose to focus on resting-state EEG for its potential towards clinical translation. However, a clinical assessment tool that only uses EEG measures may also have limited accuracy. For example, EEG data has high temporal resolution, but it is limited in spatial resolution. Measures derived from EEG data represent high-level neural activity seen at the cerebral cortex. Source localization methods for EEG can estimate activity at lower levels, but it may also be confounded by other sources of noise. Combining markers from different modalities (each with a unique advantage) may help bring together several lines of evidence to effectively predict treatment response.

2) In this thesis, measures of neural dynamics were found to predict general treatment response to each type of treatment individually: seizure therapy (ECT and MST) or pharmacotherapy (escitalopram). Future studies should consider using these measures, or potentially identify new measures, to predict whether one treatment can provide greater response compared to another. This differential prediction would be the next challenge in the development of personalized tools for treatment selection.

3) Future development of treatment prediction models is strongly suggested to include non-linear measures of neural dynamics (Natarajan, Acharya et al. 2004, Stam 2005). As detailed in previous sections, the neurophysiological mechanisms underlying the efficacy of treatments for depression are highly complex. An understanding of normal or disturbed neural processes responsible for high-level brain functions perhaps cannot be provided solely by reductionist approaches (i.e., linear measures). The successful integration of complexity analysis in our work also provides supports the investigation of additional measures of nonlinear dynamics. Examples may include measures from graph

theory analysis (Smit, Stam et al. 2008, Bullmore and Sporns 2009) or may be based on information theory (Pereda, Quiroga et al. 2005).

9.8 Thesis Limitations

The EEG modality provides excellent temporal resolution needed to record the dynamics of neural activity, on the order of milliseconds to micro-seconds. However, its spatial resolution is poor and potentially not sufficient for differentiating activity between regions or between neural circuits in the brain. Source localization with EEG data also has limitations. The electrical distribution measured over the scalp by EEG, at any given time point, could theoretically be generated by an infinite number of possible sources in the brain. Several methods attempt to solve this inverse problem (including the LORETA method used in this work) but each method has assumptions and limitations (Pascual-Marqui 1999, Grech, Cassar et al. 2008). Therefore, all results in this study should be checked and validated by future studies. An alternative approach is to combine EEG markers with markers from different neuroimaging methods such as PET, fMRI, etc. to compensate for the poor spatial resolution of EEG.

Measures of neural dynamics used in this thesis were informative and successful for the scope of the aims defined in this thesis. However, all of these measures were derived as an average over several segments of the data. Although this provided a good signal-to-noise ratio for each measure, it failed to account for variability over time. In other words, an assumption of stationarity of the neural signal was made for all the analyses. We hope that future work can investigate dynamic transitions in neural states using data collected over a longer length of time.

The dichotomization of patients into responders (i.e., greater than or equal to 50% improvement in a clinical rating scale) or non-responders is also a major limitation of this thesis. In a clinical setting, the binary divide between response and non-response is a common practice and can be useful when defining clinical outcomes but in the study of neurophysiology, defining response as a continuous outcome may be a better approach. Although a bimodal distribution of response was observed in the studies included in this thesis, the lack of a bimodal distribution could lead

to inflated Type-1 errors (Cohen 1983, Irwin and McClelland 2003, Owen and Froman 2005, Van Walraven and Hart 2008, Bennette and Vickers 2012, Barnwell-Ménard, Li et al. 2015). Understanding how minute changes in neurophysiology might be associated with changes in depressive scores is crucial for understanding how treatments for depression might show efficacy.

Another limitation is that much of the work in this thesis only included data collected before and after treatment. For the CAN-BIND study, data from an early time point was also included (i.e., 2 weeks into the treatment trial). To study the neurophysiological targets of treatments, modulation of neural dynamics should be monitored consistently over the course of the treatment (i.e., every week or every 2nd or 3rd treatment session) but unfortunately this was not done in the studies included in this thesis. Future research is necessary to address this limitation.

The work included in this thesis also does not account for confounding effects that may result from different bio-types of depression. A recent resting-state fMRI study suggests that there may be four different bio-types of depression (Wager and Woo 2017), where patients in each bio-type reveal distinct symptoms and neural connectivity. The study also predicted that treatment efficacy is different for each bio-type. In future studies, the inclusion of markers designed to identify these bio-types in treatment outcome prediction models is likely to yield a much improved classification performance.

Our results also do not demonstrate causality. Findings from this work demonstrate correlations between changes in neural dynamics (or baseline neural dynamics) and therapeutic outcome. Studies suggest that the combination of EEG with transcranial brain stimulation might help elucidate causal influences that neural units exert over another (Friston 1994, Massimini, Ferrarelli et al. 2005). Previous studies have used measures of complexity (Casali, Gosseries et al. 2013, Sarasso, Rosanova et al. 2014) or the phase of alpha oscillations (Thut and Miniussi 2009) with TMS-EEG to investigate the functionality of brain regions.

Finally, although machine learning methods have shown success in several clinical applications including diagnosis and prediction of treatment outcome, accuracy of these algorithms can

always be improved by training with larger datasets. Predictive features identified through our machine learning work (see Chapter 8) must also be tested on an independent dataset to assess the generalizability of the models.

9.9 Novel Contributions

There are several novel contributions in this thesis. Results from the seizure therapy project (Chapters 4-5) were the first to show the modulation of neural complexity and global brain-network dynamics, via multiscale entropy and microstate analysis, following a course of ECT or a course of MST. These studies were also the first to show differential impact on neural dynamics following ECT or MST. This is an important novel contribution since MST is still in the clinical trial phase. This thesis also provided evidence towards a common mechanism of action of treatments for depression: the modulation of specific frequency and/or spatio-temporal neural dynamics for treatment response (seizure therapy) or non-response (escitalopram). Variations in therapeutic impact between treatments were hypothesized to arise from potential differences in the modulation of specific frequency, temporal and spatial characteristics of neural dynamics. Additional causational evidence is still required to extend these results and to investigate whether changes in neural dynamics are associated with changes in neuroplasticity.

9.10 Clinical Implications

With further evaluation and development, markers similar to the ones identified in this thesis, may be integrated at several critical decision points of the treatment course. This includes optimizing treatment parameters, validating novel treatment targets, clinical screening of patients for a treatment, and even accelerating the approval process of a treatment by regulating administrations. In addition, the identification of early predictive markers for antidepressant response can contribute towards the development of a clinical decision-making tool for individualized treatment selection and may also provide rationale for optimal combinations of treatments for maximum therapeutic outcome. We hope that the knowledge gained from this work will help guide the development of personalized treatment for depression and potentially other brain disorders.

Appendix I: List of EEG Electrodes Used for Analysis

List of the 60 EEG channels included in the analysis for Study 1	List of the 58 EEG channels included in the analysis for Studies 2 and 3
FP1	FP1
FPZ	FPZ
FP2	FP2
AF3	AF3
AF4	AF4
F7	F7
F5	F5
F3	F3
F1	F1
FZ	FZ
F2	F2
F4	F4
F6	F6
F8	F8
FT7	FT7
FC5	FC5
FC3	FC3
FC1	FC1
FCZ	FCZ
FC2	FC2
FC4	FC4
FC6	FC6
FT8	FT8
T7	T7
C5	C5
C3	C3
C1	C1
CZ	CZ
C2	C2
C4	C4
C6	C6
T8	T8
TP7	TP7
CP5	CP5
CP3	CP3
CP1	CP1
CPZ	CPZ
CP2	CP2
CP4	CP4
CP6	CP6
TP8	TP8
P7	P7
P5	P5
P3	P3
P1	P1

PZ	PZ
P2	P2
P4	P4
P6	P6
P8	P8
PO7	PO7
PO5	PO3
PO3	POZ
POZ	PO4
PO4	PO8
PO6	O1
PO8	OZ
O1	O2
OZ	
O2	

Appendix II: Summary of Mechanistic Markers for Treatments of Depression Studied in this Thesis

	ELECTROCONVULSIVE THERAPY	MAGNETIC SEIZURE THERAPY	ESCITALOPRAM
Frequency Analysis **(Relative Power)**	Responders showed widespread increases in delta and theta oscillations (<8Hz). In addition, widespread decreases were also observed in high alpha (10-12Hz), beta (12-30Hz) and gamma oscillations (30-50Hz).	Responders and non-responders revealed a significant decrease in high beta and gamma oscillations (16-50Hz). ~ Decrease in high beta oscillations was shown to be widespread. ~ Decrease in gamma oscillations was observed in frontal, fronto-central, central and centro-parietal regions.	**Baseline to Week 8** Responders reveal significant widespread increases in delta and theta oscillations (<8Hz). Increase in gamma (33-50Hz) oscillations approached significance (p=0.06). ~ Increase in alpha and beta oscillations was shown to be widespread. ~ Increase in gamma oscillations was observed in left frontal and left fronto-central regions. **Week 2 to Week 8** Responders did not reveal any significant changes.
	Non-responders revealed a widespread increase in delta and theta oscillations (<7Hz) and a widespread decrease in high alpha, beta and low gamma oscillations (10-35Hz).	Non-responders also showed a decrease in high alpha (10-12Hz), beta (16-20Hz; 24-30Hz) and gamma (30-50Hz) oscillations. ~ Decrease in alpha oscillations was	**Baseline to Week 8** Non-responders did not reveal any significant changes. **Week 2 to Week 8** Non-responders did not reveal any

		shown to be widespread ~ Decrease in beta oscillations was shown to be widespread ~ Decrease in low gamma (30-40Hz) oscillations was shown to be widespread. ~ Decrease in high gamma (40-50Hz) oscillations was observed in left frontal, fronto-central and right parieto-occipital regions.	significant changes.
	Correlation analysis revealed that a decrease in gamma oscillations (30-50Hz) was associated with improvement in depressive symptoms following seizure therapy. ~ Decrease in (30-45Hz) gamma oscillations was observed in parieto-occipital regions. ~ Decrease in (45-50Hz) gamma oscillations was observed in fronto-central and parieto-occipital regions. Correlation analysis revealed that a decrease in delta+theta (<9Hz) & gamma oscillations (>40Hz) was associated with improvement in cognition following seizure therapy. ~ Decrease in delta and theta oscillations was global ~ Decrease in gamma oscillations was observed in centro-parietal regions.		**Baseline to Week 8** Correlation analysis revealed that the widespread increases in theta power are associated with improvement in depressive symptoms. In addition, decrease in alpha (11-12Hz) and beta (21-22Hz) oscillations was associated with improvement in symptoms. **Week 2 to Week 8** Correlation analysis revealed that an increase in theta (~7Hz), decrease in alpha (11-12Hz), changes in beta (20-25Hz) and increase in gamma (30-50Hz) oscillations were associated with improvement in symptoms.

			~ Increase in theta oscillations was observed in the right centro-parietal regions. ~ Decrease in alpha oscillations was observed in left central and occipital regions. ~ Decrease in beta oscillations (20-22Hz) was observed in right temporal regions ~ Increase in beta oscillations (23-25Hz) was observed in left central and left centro-parietal regions. ~ Increase in gamma oscillations (30-50Hz) was observed in central and centro-parietal regions.
Complexity Analysis	Responders showed a significant global decrease in fine timescales (<30) and a global increase in coarse timescales (>50).	Responders showed a significant decrease in complexity in fine timescales (<20). ~ This decrease was specific to fronto-central and parieto-occipital regions.	**Baseline to Week 8** Responders did not reveal any significant changes in complexity. **Week 2 to Week 8** Responders revealed a significant increase in complexity in mid-coarse scales (20-70). ~ This effect was observed in fronto-central, temporal, centro-parietal and parieto-occipital regions.
	Non-responders did not reveal any	Non-responders did not reveal any	**Baseline to Week 8** Non-responders did not reveal any

	significant changes in complexity.	significant changes in complexity.	significant changes in complexity. **Week 2 to Week 8** Non-responders did not reveal any significant changes in complexity.
	Correlation analysis showed that a decrease in complexity in fine timescales (<30) following seizure therapy correlated with improvement in depressive symptoms. ~ This effect was observed in fronto-central and parieto-occipital regions. Correlation analysis also showed that an increase in complexity in coarse timescales (>66) following seizure therapy was linked with greater decline in cognition. ~ This effect was spatially global.		**Baseline to Week 8** Correlation analysis revealed that an increase in complexity in mid-high timescales (15-55) correlated with improvement in depressive symptoms. ~ This effect was observed in left parietal and left parieto-occipital regions. **Week 2 to Week 8** Correlation analysis revealed that an increase in complexity in fine (1-10), mid (15-50) and coarse (60-70) timescales was significantly correlated with improvement in depressive symptoms. ~ In fine timescales, this effect was observed in fronto-central, temporal, centro-parietal and parieto-occipital regions. ~ In mid-high timescales, this effect was observed in the centro-parietal, and parieto-occipital regions. ~ In coarser timescales, the effect was observed in the left parietal and left parieto-occipital regions.

Microstate Analysis	Responders revealed a significant increase in the duration of State A and a significant decrease in the frequency of States B, C and D.	An increase in the duration and decrease in the frequency of microstates was observed following MST. However, these effects were not specific to response or to any single microstate. In other words, changes in microstate characteristics were not specific to responders.	**Baseline to Week 8** Responders revealed a significant increase in the coverage of State B. **Week 2 to Week 8** Responders revealed a significant increase in the frequency of State D.
	Non-responders did not reveal any significant changes in microstate characteristics.	Changes in microstate characteristics were not specific to non-responders.	**Baseline to Week 8** Non-responders did not show any significant changes from baseline to week 8. **Week 2 to Week 8** Non-responders revealed a significant: ~ increase in the duration and coverage of State B ~ decrease in the duration, frequency & coverage of State C ~ decrease in the frequency of State D
			Baseline to Week 8

	Correlation analysis revealed that an increase in the coverage of State A and an increase in the duration of State D following ECT is associated with improvement in self-rated depressive symptoms. In addition, the ratio between the "change in duration of State C" and the "change in duration of State D" significantly correlated with improvement in self-rated depressive symptoms. This suggested an interaction between States C and D.	Correlation analysis between change in microstate characteristics (from baseline to end of treatment) and change in clinical scores did not reveal any significant associations.	Correlation analysis did not reveal significant associations. **Week 2 to Week 8** Correlation analysis did not reveal significant associations.

Appendix III: Summary of Predictive Markers for Response to Escitalopram

	Baseline Markers	Week 2 Markers	Early Change Markers (Baseline to Week 2)
Frequency Analysis	Correlation analysis revealed that high alpha (9-12Hz), beta (19-23Hz) and gamma (30-50Hz) oscillations at baseline is linked with improvement in depressive symptoms. ~ In the alpha band, the effect was observed in fronto-central and centro-parietal regions. ~ In the beta band, the effect was observed in frontal-central, centro-parietal, and parieto-occipital regions. ~ In the gamma band, regions were significant in the parietal regions.	Correlation analysis in sensor space revealed that high power in the alpha (8-12Hz), beta (12-30Hz) and gamma (30-45Hz) oscillations correlates with improvement in symptoms. ~ In the alpha band, the effect was seen in fronto-central, temporal and centro-parietal regions. ~ In the low-beta band (12-18Hz), the effect was seen in frontal and parietal regions. ~ In the mid-beta band (18-22Hz), the effect was seen in fronto-central, centro-parietal and parieto-occipital regions. ~ In the high-beta band (22-30Hz), the effect was seen in fronto-central and centro-parietal regions. ~ In the gamma band (30-45Hz), the effect was seen in fronto-central, centro-parietal and parieto-occipital regions.	An increase in delta-theta (1-8Hz) and beta-gamma (18-35Hz) oscillations significantly correlated with improvement in depressive symptoms. ~ For the delta-theta cluster, this effect was observed in fronto-central, centro-parietal, tempro-parietal (TP7, TP8), parietal (P1, P3, PZ, P2, P4), parieto-occipital (PO3, PO7) and occipital (O1, OZ, O2) electrodes. ~ For the beta cluster, this effect was observed in the fronto-central, centro-parietal, and parieto-occipital electrodes. ~ For the low gamma cluster, the effect was seen in right parietal and right centro-parietal regions.
Complexity Analysis	Correlation analysis revealed that a low complexity value at baseline correlates with improvement in depressive symptoms across coarse timescales (30-70).	Correlation analysis revealed that a low complexity value at week 2 in coarser timescales (30-70) and a high complexity value at week 2 in finer timescales (12-17) correlates	Correlation analysis in sensor space did not reveal any significant effects.

	~ This effect was observed in the fronto-central, centro-parietal and parieto-occipital regions.	with improvement in depressive symptoms. ~ The cluster in the coarser timescales was observed in fronto-central, centro-parietal and parieto-occipital regions. ~ The cluster in the finer timescales was observed in fronto-central regions.	
Microstate Analysis	None	A long duration and high frequency of State B at week 2 correlated with improvement in depressive symptoms.	An increase in the coverage of State B (from baseline to week 2) correlated with improvement in depressive symptoms.

References

Abbott, C. C., P. Gallegos, N. Rediske, N. T. Lemke and D. K. Quinn (2014). "A review of longitudinal electroconvulsive therapy: neuroimaging investigations." Journal of geriatric psychiatry and neurology **27**(1): 33-46.

Abrams, R. (2002). Electroconvulsive therapy, Oxford University Press.

Ahmadlou, M., H. Adeli and A. Adeli (2012). "Fractality analysis of frontal brain in major depressive disorder." International Journal of Psychophysiology **85**(2): 206-211.

Alexopoulos, G. S., M. J. Hoptman, D. Kanellopoulos, C. F. Murphy, K. O. Lim and F. M. Gunning (2012). "Functional connectivity in the cognitive control network and the default mode network in late-life depression." Journal of affective disorders **139**(1): 56-65.

Altar, C. A., R. E. Whitehead, R. Chen, G. Wörtwein and T. M. Madsen (2003). "Effects of electroconvulsive seizures and antidepressant drugs on brain-derived neurotrophic factor protein in rat brain." Biological psychiatry **54**(7): 703-709.

American Psychiatric Association (2000). Diagnostic and statistical manual of mental disorders. Arlington, VA, American Psychiatric Publishing.

American Psychiatric Association (2001). The Practice of Electroconvulsive Therapy: Recommendations for Treatment. A Task Force Report of the American Psychiatric Association. Washington, DC, American Psychiatric Association.

American Psychiatric Association (2013). Diagnostic and statistical manual of mental disorders (DSM-5®), American Psychiatric Association.

Amsterdam, J. D. and M. Chopra (2001). "Monoamine oxidase inhibitors revisited." Psychiatric Annals **31**(6): 361-370.

Anand, A., Y. Li, Y. Wang, J. Wu, S. Gao, L. Bukhari, V. P. Mathews, A. Kalnin and M. J. Lowe (2005). "Activity and connectivity of brain mood regulating circuit in depression: a functional magnetic resonance study." Biological psychiatry **57**(10): 1079-1088.

Anand, A., V. Mathews, L. Bukhari, W. Yang, J. Shirley, L. Jeffrey and M. Lowe (2003). Functional connectivity within the mood regulating circuit in major depression-An fMRI study. Biological Psychiatry, ELSEVIER SCIENCE INC 360 PARK AVE SOUTH, NEW YORK, NY 10010-1710 USA.

Anastassiou, C. A., R. Perin, H. Markram and C. Koch (2011). "Ephaptic coupling of cortical neurons." Nature neuroscience **14**(2): 217.

Anderson, I. (1998). "SSRIs versus tricyclic antidepressants in depressed inpatients: A meta-analysis of efficacy and tolerability." Depression and Anxiety **7**(S1): 11-17.

Andrade, L., J. J. Caraveo-Anduaga, P. Berglund, R. V. Bijl, R. D. Graaf, W. Vollebergh, E. Dragomirecka, R. Kohn, M. Keller and R. C. Kessler (2003). "The epidemiology of major depressive episodes: results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys." International journal of methods in psychiatric research **12**(1): 3-21.

Andreasen, N. C., S. Paradiso and D. S. O'leary (1998). "“Cognitive dysmetria” as an integrative theory of schizophrenia: a dysfunction in cortical-subcortical-cerebellar circuitry?" Schizophrenia bulletin **24**(2): 203-218.

Andreescu, C., M. Butters, E. J. Lenze, V. K. Venkatraman, M. Nable, C. F. Reynolds III and H. J. Aizenstein (2009). "fMRI activation in late-life anxious depression: a potential biomarker." International Journal of Geriatric Psychiatry: A journal of the psychiatry of late life and allied sciences **24**(8): 820-828.

Arango, V., M. D. Underwood, A. V. Gubbi and J. J. Mann (1995). "Localized alterations in pre-and postsynaptic serotonin binding sites in the ventrolateral prefrontal cortex of suicide victims." Brain research **688**(1-2): 121-133.

Armitage, R., R. Hoffmann and A. Rush (1999). "Biological rhythm disturbance in depression: temporal coherence of ultradian sleep EEG rhythms." Psychological medicine **29**(6): 1435-1448.

Arns, M., W. H. Drinkenburg, P. B. Fitzgerald and J. L. Kenemans (2012). "Neurophysiological predictors of non-response to rTMS in depression." Brain stimulation **5**(4): 569-576.

Arns, M., A. Etkin, U. Hegerl, L. M. Williams, C. DeBattista, D. M. Palmer, P. B. Fitzgerald, A. Harris and E. Gordon (2015). "Frontal and rostral anterior cingulate (rACC) theta EEG in depression: implications for treatment outcome?" European Neuropsychopharmacology **25**(8): 1190-1200.

Ascher, J. A., J. O. Cole, J.-N. Colin, J. P. Feighner, R. M. Ferris, H. Fibiger, R. N. Golden, P. Martin, W. Z. Potter and E. Richelson (1995). "Bupropion: a review of its mechanism of antidepressant activity." The Journal of clinical psychiatry.

Atluri, S., M. Frehlich, Y. Mei, L. Garcia Dominguez, N. C. Rogasch, W. Wong, Z. J. Daskalakis and F. Farzan (2016). "TMSEEG: A MATLAB-based graphical user interface for processing electrophysiological signals during transcranial magnetic stimulation." Frontiers in neural circuits **10**: 78.

Austin, E. J., I. J. Deary, G. J. Gibson, M. J. McGregor and J. B. Dent (1998). "Individual response spread in self-report scales: Personality correlations and consequences." Personality and Individual Differences **24**(3): 421-438.

Bachmann, M., K. Kalev, A. Suhhova, J. Lass and H. Hinrikus (2015). Lempel Ziv Complexity of EEG in depression. 6th European Conference of the International Federation for Medical and Biological Engineering, Springer.

Bachmann, M., J. Lass, A. Suhhova and H. Hinrikus (2013). "Spectral asymmetry and Higuchi's fractal dimension measures of depression electroencephalogram." Computational and mathematical methods in medicine **2013**.

Backhouse, F. A., Y. Noda, Y. Knyahnytska, F. Farzan, J. Downar, T. K. Rajji, B. H. Mulsant, Z. J. Daskalakis and D. M. Blumberger (2018). "Characteristics of ictal EEG in Magnetic Seizure Therapy at various stimulation frequencies." Clinical Neurophysiology.

Balleine, B. W., M. R. Delgado and O. Hikosaka (2007). "The role of the dorsal striatum in reward and decision-making." Journal of Neuroscience **27**(31): 8161-8165.

Bares, M., M. Brunovsky, M. Kopecek, T. Novak, P. Stopkova, J. Kozeny, P. Sos, V. Krajca and C. Höschl (2008). "Early reduction in prefrontal theta QEEG cordance value predicts response to venlafaxine treatment in patients with resistant depressive disorder." European Psychiatry **23**(5): 350-355.

Bares, M., M. Brunovsky, T. Novak, M. Kopecek, P. Stopkova, P. Sos, V. Krajca and C. Höschl (2010). "The change of prefrontal QEEG theta cordance as a predictor of response to bupropion treatment in patients who had failed to respond to previous antidepressant treatments." European Neuropsychopharmacology **20**(7): 459-466.

Barnwell-Ménard, J. L., Q. Li and A. A. Cohen (2015). "Effects of categorization method, regression type, and variable distribution on the inflation of Type-I error rate when categorizing a confounding variable." Statistics in medicine **34**(6): 936-949.

Barry, R. J., A. R. Clarke, S. J. Johnstone, C. A. Magee and J. A. Rushby (2007). "EEG differences between eyes-closed and eyes-open resting conditions." Clinical Neurophysiology **118**(12): 2765-2773.

Baskaran, A., F. Farzan, R. Milev, C. A. Brenner, S. Alturi, M. P. McAndrews, P. Blier, K. Evans, J. A. Foster and B. N. Frey (2017). "The comparative effectiveness of electroencephalographic indices in predicting response to escitalopram therapy in depression: A pilot study." Journal of affective disorders **227**: 542-549.

Baskaran, A., R. Milev and R. S. McIntyre (2012). "The neurobiology of the EEG biomarker as a predictor of treatment response in depression." Neuropharmacology **63**(4): 507-513.

Bassett, D. S. and E. Bullmore (2006). "Small-world brain networks." The neuroscientist **12**(6): 512-523.

Beck, A. T. (1979). Cognitive therapy of depression, Guilford press.

Beck, A. T., M. Kovacs and A. Weissman (1979). "Assessment of suicidal intention: the Scale for Suicide Ideation." J Consult Clin Psychol **47**(2): 343-352.

Begić, D., V. Popović-Knapić, J. Grubišin, B. Kosanović-Rajačić, I. Filipčić, I. Telarović and M. Jakovljević (2011). "Quantitative electroencephalography in schizophrenia and depression." Psychiatria Danubina **23**(4.): 355-362.

Belmaker, R. and G. Agam (2008). "Major depressive disorder." New England Journal of Medicine **358**(1): 55-68.

Bennette, C. and A. Vickers (2012). "Against quantiles: categorization of continuous variables in epidemiologic research, and its discontents." BMC medical research methodology **12**(1): 21.

Berlim, M. T. and G. Turecki (2007). "Definition, Assessment, and Staging of Treatment—Resistant Refractory Major Depression: A Review of Current Concepts and Methods." The Canadian Journal of Psychiatry **52**(1): 46-54.

Bhagwagar, Z., M. Wylezinska, P. Jezzard, J. Evans, E. Boorman, P. M. Matthews and P. J. Cowen (2008). "Low GABA concentrations in occipital cortex and anterior cingulate cortex in medication-free, recovered depressed patients." International Journal of Neuropsychopharmacology **11**(2): 255-260.

Bismark, A. W., F. A. Moreno, J. L. Stewart, D. N. Towers, J. A. Coan, J. Oas, R. P. Erickson and J. J. Allen (2010). "Polymorphisms of the HTR1a allele are linked to frontal brain electrical asymmetry." Biological psychology **83**(2): 153-158.

Bjørk, M. H., T. Sand, G. Bråthen, O. M. Linaker, G. Morken, B. M. Nilsen and A. E. Vaaler (2008). "Quantitative EEG findings in patients with acute, brief depression combined with other fluctuating psychiatric symptoms: a controlled study from an acute psychiatric department." BMC psychiatry **8**(1): 89.

Blum, A. L. and P. Langley (1997). "Selection of relevant features and examples in machine learning." Artificial intelligence **97**(1-2): 245-271.

Bocchio-Chiavetto, L., R. Zanardini, M. Bortolomasi, M. Abate, M. Segala, M. Giacopuzzi, M. A. Riva, E. Marchina, P. Pasqualetti and J. Perez (2006). "Electroconvulsive Therapy (ECT) increases serum Brain Derived Neurotrophic Factor (BDNF) in drug resistant depressed patients." European Neuropsychopharmacology **16**(8): 620-624.

Bosl, W., A. Tierney, H. Tager-Flusberg and C. Nelson (2011). "EEG complexity as a biomarker for autism spectrum disorder risk." BMC Med **9**: 18.

Bracewell, R. N. and R. N. Bracewell (1986). The Fourier transform and its applications, McGraw-Hill New York.

Brandeis, D., H. Naylor, R. Halliday, E. Callaway and L. Yano (1992). "Scopolamine effects on visual information processing, attention, and event-related potential map latencies." Psychophysiology **29**(3): 315-335.

Bremner, J. D., M. Narayan, E. R. Anderson, L. H. Staib, H. L. Miller and D. S. Charney (2000). "Hippocampal volume reduction in major depression." American Journal of Psychiatry **157**(1): 115-118.

Bremner, J. D., M. Vythilingam, E. Vermetten, A. Nazeer, J. Adil, S. Khan, L. H. Staib and D. S. Charney (2002). "Reduced volume of orbitofrontal cortex in major depression." Biological psychiatry **51**(4): 273-279.

Britz, J., D. Van De Ville and C. M. Michel (2010). "BOLD correlates of EEG topography reveal rapid resting-state network dynamics." Neuroimage **52**(4): 1162-1170.

Brodbeck, V., A. Kuhn, F. von Wegner, A. Morzelewski, E. Tagliazucchi, S. Borisov, C. M. Michel and H. Laufs (2012). "EEG microstates of wakefulness and NREM sleep." Neuroimage **62**(3): 2129-2139.

Brody, A. L., S. Saxena, D. H. Silverman, L. A. Fairbanks, M. E. Phelps, S.-C. Huang, H.-M. Wu, K. Maidment, L. R. Baxter Jr and S. Alborzian (1999). "Brain metabolic changes in major depressive disorder from pre-to post-treatment with paroxetine." Psychiatry Research: Neuroimaging **91**(3): 127-139.

Brown, A. and S. Gershon (1993). "Dopamine and depression." Journal of Neural Transmission/General Section JNT **91**(2-3): 75-109.

Broyd, S. J., C. Demanuele, S. Debener, S. K. Helps, C. J. James and E. J. Sonuga-Barke (2009). "Default-mode brain dysfunction in mental disorders: a systematic review." Neuroscience & biobehavioral reviews **33**(3): 279-296.

Bruder, G. E., J. P. Sedoruk, J. W. Stewart, P. J. McGrath, F. M. Quitkin and C. E. Tenke (2008). "Electroencephalographic alpha measures predict therapeutic response to a selective serotonin reuptake inhibitor antidepressant: pre-and post-treatment findings." Biological psychiatry **63**(12): 1171-1177.

Bruder, G. E., J. W. Stewart, C. E. Tenke, P. J. McGrath, P. Leite, N. Bhattacharya and F. M. Quitkin (2001). "Electroencephalographic and perceptual asymmetry differences between responders and nonresponders to an SSRI antidepressant." Biological Psychiatry **49**(5): 416-425.

Bruder, G. E., C. E. Tenke, V. Warner, Y. Nomura, C. Grillon, J. Hille, P. Leite and M. M. Weissman (2005). "Electroencephalographic measures of regional hemispheric activity in offspring at risk for depressive disorders." Biological psychiatry **57**(4): 328-335.

Brunet, D., M. M. Murray and C. M. Michel (2011). "Spatiotemporal analysis of multichannel EEG: CARTOOL." Computational intelligence and neuroscience **2011**: 2.

Brunoni, A. R., C. Baeken, R. Machado-Vieira, W. F. Gattaz and M.-A. Vanderhasselt (2014). "BDNF blood levels after electroconvulsive therapy in patients with mood disorders: a systematic review and meta-analysis." The World Journal of Biological Psychiatry **15**(5): 411-418.

Brunoni, A. R., M. Lopes and F. Fregni (2008). "A systematic review and meta-analysis of clinical studies on major depression and BDNF levels: implications for the role of neuroplasticity in depression." International Journal of Neuropsychopharmacology **11**(8): 1169-1180.

Buckner, R. L., J. Sepulcre, T. Talukdar, F. M. Krienen, H. Liu, T. Hedden, J. R. Andrews-Hanna, R. A. Sperling and K. A. Johnson (2009). "Cortical hubs revealed by intrinsic functional connectivity: mapping, assessment of stability, and relation to Alzheimer's disease." The Journal of Neuroscience **29**(6): 1860-1873.

Bullmore, E. and O. Sporns (2009). "Complex brain networks: graph theoretical analysis of structural and functional systems." Nature Reviews Neuroscience **10**(3): 186.

Bullmore, E. T., J. Suckling, S. Overmeyer, S. Rabe-Hesketh, E. Taylor and M. J. Brammer (1999). "Global, voxel, and cluster tests, by theory and permutation, for a difference between two groups of structural MR images of the brain." IEEE transactions on medical imaging **18**(1): 32-42.

Buzsáki, G., C. A. Anastassiou and C. Koch (2012). "The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes." Nature reviews neuroscience **13**(6): 407.

Buzsáki, G. and A. Draguhn (2004). "Neuronal oscillations in cortical networks." science **304**(5679): 1926-1929.

Bymaster, F. P., W. Zhang, P. A. Carter, J. Shaw, E. Chernet, L. Phebus, D. T. Wong and K. W. Perry (2002). "Fluoxetine, but not other selective serotonin uptake inhibitors, increases norepinephrine and dopamine extracellular levels in prefrontal cortex." Psychopharmacology **160**(4): 353-361.

Carney, S., P. Cowen, J. Geddes, G. Goodwin, R. Rogers, K. Dearness, A. Tomlin, J. Eastaugh, N. Freemantle and H. Lester (2003). "Efficacy and safety of electroconvulsive therapy in depressive disorders: a systematic review and meta-analysis." The Lancet.

Carpenter, L. L., P. G. Janicak, S. T. Aaronson, T. Boyadjis, D. G. Brock, I. A. Cook, D. L. Dunner, K. Lanocha, H. B. Solvason and M. A. Demitrack (2012). "Transcranial magnetic stimulation (TMS) for major depression: a multisite, naturalistic, observational study of acute treatment outcomes in clinical practice." Depression and anxiety **29**(7): 587-596.

Carvalho, A., H. Moraes, H. Silveira, P. Ribeiro, R. A. Piedade, A. C. Deslandes, J. Laks and M. Versiani (2011). "EEG frontal asymmetry in the depressed and remitted elderly: Is it related to the trait or to the state of depression?" Journal of affective disorders **129**(1-3): 143-148.

Casali, A. G., O. Gosseries, M. Rosanova, M. Boly, S. Sarasso, K. R. Casali, S. Casarotto, M.-A. Bruno, S. Laureys and G. Tononi (2013). "A theoretically based index of consciousness independent of sensory processing and behavior." Science translational medicine **5**(198): 198ra105-198ra105.

Chang, C.-C. and C.-J. Lin (2011). "LIBSVM: a library for support vector machines." ACM transactions on intelligent systems and technology (TIST) **2**(3): 27.

Chang, C. and G. H. Glover (2010). "Time–frequency dynamics of resting-state brain connectivity measured with fMRI." Neuroimage **50**(1): 81-98.

Chebib, M. and G. A. Johnston (1999). "The 'ABC'of GABA receptors: a brief review." Clinical and experimental pharmacology and physiology **26**(11): 937-940.

Chekroud, A. M., R. J. Zotti, Z. Shehzad, R. Gueorguieva, M. K. Johnson, M. H. Trivedi, T. D. Cannon, J. H. Krystal and P. R. Corlett (2016). "Cross-trial prediction of treatment outcome in depression: a machine learning approach." The Lancet Psychiatry **3**(3): 243-250.

Chen, C.-H., K. Ridler, J. Suckling, S. Williams, C. H. Fu, E. Merlo-Pich and E. Bullmore (2007). "Brain imaging correlates of depressive symptom severity and predictors of symptom improvement after antidepressant treatment." Biological psychiatry **62**(5): 407-414.

Chen, Z.-Y., D. Jing, K. G. Bath, A. Ieraci, T. Khan, C.-J. Siao, D. G. Herrera, M. Toth, C. Yang and B. S. McEwen (2006). "Genetic variant BDNF (Val66Met) polymorphism alters anxiety-related behavior." Science **314**(5796): 140-143.

Coffey, C. E., J. Lucke, R. D. Weiner, A. D. Krystal and M. Aque (1995). "Seizure threshold in electroconvulsive therapy (ECT) II. The anticonvulsant effect of ECT." Biological Psychiatry **37**(11): 777-788.

Cohen, J. (1983). "The cost of dichotomization." Applied psychological measurement **7**(3): 249-253.

Cole, M. W., G. Repovš and A. Anticevic (2014). "The frontoparietal control system: a central role in mental health." The Neuroscientist **20**(6): 652-664.

Colwell, C. S. (2011). "Linking neural activity and molecular oscillations in the SCN." Nature Reviews Neuroscience **12**(10): 553.

Connolly, K. R., A. Helmer, M. A. Cristancho, P. Cristancho and J. P. O'Reardon (2012). "Effectiveness of transcranial magnetic stimulation in clinical practice post-FDA approval in the United States: results observed with the first 100 consecutive cases of depression at an academic medical center." J Clin Psychiatry **73**(4): e567-e573.

Cook, I. A. and A. Leuchter (2001). Prefrontal changes and treatment response prediction in depression. Seminars in clinical neuropsychiatry.

Cook, I. A., A. F. Leuchter, M. Morgan, E. Witte, W. F. Stubbeman, M. Abrams, S. Rosenberg and S. H. Uijtdehaage (2002). "Early changes in prefrontal activity characterize clinical responders to antidepressants." Neuropsychopharmacology **27**(1): 120-131.

Cook, I. A., A. F. Leuchter, M. L. Morgan, W. Stubbeman, B. Siegman and M. Abrams (2005). "Changes in prefrontal activity characterize clinical response in SSRI nonresponders: a pilot study." Journal of Psychiatric Research **39**(5): 461-466.

Costa, M., A. L. Goldberger and C.-K. Peng (2005). "Multiscale entropy analysis of biological signals." Physical review **71**(2): 021906.

Cretaz, E., A. R. Brunoni and B. Lafer (2015). "Magnetic seizure therapy for unipolar and bipolar depression: a systematic review." Neural plasticity **2015**.

Creutzfeldt, O. (1974). Electrical Activity from the Neuron to the EEG and EMG: The neuronal generation of the EEG, Elsevier.

Cuijpers, P., A. van Straten, G. Andersson and P. van Oppen (2008). "Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies." Journal of consulting and clinical psychology **76**(6): 909.

Cukic, M., D. Pokrajac, M. Stokic, V. Radivojevic and M. Ljubisavljevic (2018). "EEG machine learning with Higuchi fractal dimension and Sample Entropy as features for successful detection of depression." arXiv preprint arXiv:1803.05985.

Cusin, C., H. Yang, A. Yeung and M. Fava (2009). Rating scales for depression. Handbook of clinical rating scales and assessment in psychiatry and mental health, Springer**:** 7-35.

Czira, M. E., H. Wersching, B. T. Baune and K. Berger (2012). "Brain-derived neurotrophic factor gene polymorphisms, neurotransmitter levels, and depressive symptoms in an elderly population." Age **34**(6): 1529-1541.

D'sa, C. and R. S. Duman (2002). "Antidepressants and neuroplasticity." Bipolar disorders **4**(3): 183-194.

Da Silva, F. L., J. Vos, J. Mooibroek and A. Van Rotterdam (1980). "Relative contributions of intracortical and thalamo-cortical processes in the generation of alpha rhythms, revealed by partial coherence analysis." Electroencephalography and clinical neurophysiology **50**(5-6): 449-456.

Davidson, R. J., W. Irwin, M. J. Anderle and N. H. Kalin (2003). "The neural substrates of affective processing in depressed patients treated with venlafaxine." American Journal of Psychiatry **160**(1): 64-75.

de Kwaasteniet, B. P., M. M. Rive, H. G. Ruhé, A. H. Schene, D. J. Veltman, L. Fellinger, G. A. van Wingen and D. Denys (2015). "Decreased resting-state connectivity between neurocognitive networks in treatment resistant depression." Frontiers in psychiatry **6**.

de Vries, Y. A., A. M. Roest, E. H. Bos, J. G. Burgerhof, H. M. van Loo and P. de Jonge (2018). "Predicting antidepressant response by monitoring early improvement of individual symptoms of depression: individual patient data meta-analysis." The British Journal of Psychiatry: 1-7.

Deakin, J. (1991). "Depression and 5HT." International Clinical Psychopharmacology.

Dell'Osso, B., M. Buoli, D. S. Baldwin and A. C. Altamura (2010). "Serotonin norepinephrine reuptake inhibitors (SNRIs) in anxiety disorders: a comprehensive review of their clinical efficacy." Human Psychopharmacology: Clinical and Experimental **25**(1): 17-29.

Delorme, A. and S. Makeig (2004). "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis." Journal of neuroscience methods **134**(1): 9-21.

Deng, Z.-D., S. H. Lisanby and A. V. Peterchev (2011). "Electric field strength and focality in electroconvulsive therapy and magnetic seizure therapy: a finite element simulation study." Journal of neural engineering **8**(1): 016007.

Deng, Z.-D., S. H. Lisanby and A. V. Peterchev (2013). "Controlling stimulation strength and focality in electroconvulsive therapy via current amplitude and electrode size and spacing: comparison with magnetic seizure therapy." The journal of ECT **29**(4): 325.

Deng, Z.-D., S. M. McClintock, N. E. Oey, B. Luber and S. H. Lisanby (2015). "Neuromodulation for mood and memory: from the engineering bench to the patient bedside." Current opinion in neurobiology **30**: 38-43.

Desmyter, S., C. Van Heeringen and K. Audenaert (2011). "Structural and functional neuroimaging studies of the suicidal brain." Progress in neuro-psychopharmacology and biological psychiatry **35**(4): 796-808.

Destrieux, C., B. Fischl, A. Dale and E. Halgren (2010). "Automatic parcellation of human cortical gyri and sulci using standard anatomical nomenclature." Neuroimage **53**(1): 1-15.

Devanand, D., C. Sobin and H. Sackeim (1995). "Predictors of retrograde amnesia following ECT." Am J Psychiatry **152**(7): 995-1001.

Dosenbach, N. U., D. A. Fair, A. L. Cohen, B. L. Schlaggar and S. E. Petersen (2008). "A dual-networks architecture of top-down control." Trends in cognitive sciences **12**(3): 99-105.

Drevets, W. C., J. L. Price, J. R. Simpson Jr, R. D. Todd, T. Reich, M. Vannier and M. E. Raichle (1997). "Subgenual prefrontal cortex abnormalities in mood disorders." Nature **386**(6627): 824.

Drevets, W. C., T. O. Videen, J. L. Price, S. H. Preskorn, S. T. Carmichael and M. E. Raichle (1992). "A functional anatomical study of unipolar depression." Journal of Neuroscience **12**(9): 3628-3641.

Duman, R. S. and G. K. Aghajanian (2012). "Synaptic dysfunction in depression: potential therapeutic targets." science **338**(6103): 68-72.

Duman, R. S. and L. M. Monteggia (2006). "A neurotrophic model for stress-related mood disorders." Biological psychiatry **59**(12): 1116-1127.

Duval, F., B. D. Lebowitz and J.-P. Macher (2006). "Treatments in depression." Dialogues in clinical neuroscience **8**(2): 191.

Earnheart, J. C., C. Schweizer, F. Crestani, T. Iwasato, S. Itohara, H. Mohler and B. Lüscher (2007). "GABAergic control of adult hippocampal neurogenesis in relation to behavior indicative of trait anxiety and depression states." Journal of Neuroscience **27**(14): 3845-3854.

Engel, A. K., P. Fries and W. Singer (2001). "Dynamic predictions: oscillations and synchrony in top–down processing." Nature Reviews Neuroscience **2**(10): 704.

Evans, D. L., J. Davidson and D. Raft (1982). "Early and late side effects of phenelzine." Journal of clinical psychopharmacology **2**(3): 208-210.

Evans, W. E. and M. V. Relling (1999). "Pharmacogenomics: translating functional genomics into rational therapeutics." science **286**(5439): 487-491.

Eytan, D. and S. Marom (2006). "Dynamics and effective topology underlying synchronization in networks of cortical neurons." Journal of Neuroscience **26**(33): 8465-8476.

Fagiolini, A. and D. J. Kupfer (2003). "Is treatment-resistant depression a unique subtype of depression?" Biological Psychiatry **53**(8): 640-648.

Fallon, B., B. Foote, B. Walsh and S. Roose (1988). "'Spontaneous' hypertensive episodes with monoamine oxidase inhibitors." The Journal of clinical psychiatry **49**(4): 163-165.

Farzan, F., S. Atluri, M. Frehlich, P. Dhami, K. Kleffner, R. Price, R. W. Lam, B. N. Frey, R. Milev and A. Ravindran (2017). "Standardization of electroencephalography for multi-site, multi-platform and multi-investigator studies: insights from the canadian biomarker integration network in depression." Scientific reports **7**(1): 7473.

Farzan, F., S. Atluri, Y. Mei, S. Moreno, A. J. Levinson, D. M. Blumberger and Z. J. Daskalakis (2017). "Brain temporal complexity in explaining the therapeutic and cognitive effects of seizure therapy." Brain **140**(4): 1011-1025.

Farzan, F., N. N. Boutros, D. M. Blumberger and Z. J. Daskalakis (2014). "What does the electroencephalogram tell us about the mechanisms of action of ECT in major depressive disorders?" The journal of ECT **30**(2): 98-106.

Farzan, F., A. Pascual-Leone, J. D. Schmahmann and M. Halko (2016). "Enhancing the Temporal Complexity of Distributed Brain Networks with Patterned Cerebellar Stimulation." Scientific Reports **6**: 23599.

Fava, M. (2000). "Weight gain and antidepressants." The Journal of clinical psychiatry **61**: 37-41.

Fava, M. (2003). "Diagnosis and definition of treatment-resistant depression." Biological psychiatry **53**(8): 649-659.

Fava, M. and K. S. Kendler (2000). "Major depressive disorder." Neuron **28**(2): 335-341.

Fekadu, A., S. C. Wooderson, K. Markopoulo, C. Donaldson, A. Papadopoulos and A. J. Cleare (2009). "What happens to patients with treatment-resistant depression? A systematic review of medium to long term outcome studies." Journal of affective disorders **116**(1-2): 4-11.

Fingelkurts, A. A., A. A. Fingelkurts, H. Rytsälä, K. Suominen, E. Isometsä and S. Kähkönen (2006). "Composition of brain oscillations in ongoing EEG during major depression disorder." Neuroscience research **56**(2): 133-144.

Fingelkurts, A. A., A. A. Fingelkurts, H. Rytsälä, K. Suominen, E. Isometsä and S. Kähkönen (2007). "Impaired functional connectivity at EEG alpha and theta frequency bands in major depression." Human brain mapping **28**(3): 247-261.

Fink, M. and R. L. Kahn (1957). "Relation of electroencephalographic delta activity to behavioral response in electroshock; quantitative serial studies." AMA Arch Neurol Psychiatry **78**(5): 516-525.

Fischer, A. S., C. J. Keller and A. Etkin (2016). "The clinical applicability of functional connectivity in depression: Pathways toward more targeted intervention." Biological Psychiatry: Cognitive Neuroscience and Neuroimaging **1**(3): 262-270.

Fitzgerald, P. B., K. E. Hoy, S. E. Herring, A. M. Clinton, G. Downey and Z. J. Daskalakis (2013). "Pilot study of the clinical and cognitive effects of high-frequency magnetic seizure therapy in major depressive disorder." Depression and anxiety **30**(2): 129-136.

Flor-Henry, P. (1979). "Neurophysiological studies of schizophrenia, mania and depression." Hemisphere asymmetries of function in psychopathology: 182-222.

Food and Drug Administration (2013). "Antidepressant Medications: U.S. Food and Drug Administration-Approved Indications and Dosages for Use in Adults." from www.cms.gov/Medicare-Medicaid-Coordination/Fraud-Prevention/Medicaid-Integrity-Education/Pharmacy-Education-Materials/Downloads/ad-adult-dosingchart.pdf.

Fox, M. D., R. L. Buckner, H. Liu, M. M. Chakravarty, A. M. Lozano and A. Pascual-Leone (2014). "Resting-state networks link invasive and noninvasive brain stimulation across diverse psychiatric and neurological diseases." Proceedings of the National Academy of Sciences **111**(41): E4367-E4375.

Fox, M. D., R. L. Buckner, M. P. White, M. D. Greicius and A. Pascual-Leone (2012). "Efficacy of transcranial magnetic stimulation targets for depression is related to intrinsic functional connectivity with the subgenual cingulate." Biological psychiatry **72**(7): 595-603.

Fox, M. D., A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen and M. E. Raichle (2005). "The human brain is intrinsically organized into dynamic, anticorrelated functional networks." Proceedings of the National Academy of Sciences **102**(27): 9673-9678.

Fransson, P. and G. Marrelec (2008). "The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis." Neuroimage **42**(3): 1178-1184.

Frazer, A. and S. Benmansour (2002). "Delayed pharmacological effects of antidepressants." Molecular psychiatry **7**(S1): S23.

Frederick, A., J. Bourget-Murray, C. A. Chapman, S. Amir and R. Courtemanche (2014). "Diurnal influences on electrophysiological oscillations and coupling in the dorsal striatum and cerebellar cortex of the anesthetized rat." Frontiers in systems neuroscience **8**: 145.

Frehlich, M., L. G. Dominguez, S. Atluri, N. Radhu, Y. Sun, Z. J. Daskalakis and F. Farzan (2016). "Unbiased cluster estimation of electrophysiological brain response." Journal of neuroscience methods **271**: 43-49.

Friston, K. (2002). "Beyond phrenology: what can neuroimaging tell us about distributed circuitry?" Annual review of neuroscience **25**(1): 221-250.

Friston, K. J. (1994). "Functional and effective connectivity in neuroimaging: a synthesis." Human brain mapping **2**(1-2): 56-78.

Frodl, T., A. L. Bokde, J. Scheuerecker, D. Lisiecka, V. Schoepf, H. Hampel, H.-J. Möller, H. Brückmann, M. Wiesmann and E. Meisenzahl (2010). "Functional connectivity bias of the orbitofrontal cortex in drug-free patients with major depression." Biological psychiatry **67**(2): 161-167.

Fu, C. H., H. Steiner and S. G. Costafreda (2013). "Predictive neural biomarkers of clinical response in depression: a meta-analysis of functional and structural neuroimaging studies of pharmacological and psychological therapies." Neurobiology of disease **52**: 75-83.

Fuggetta, G. and N. A. Noh (2013). "A neurophysiological insight into the potential link between transcranial magnetic stimulation, thalamocortical dysrhythmia and neuropsychiatric disorders." Experimental neurology **245**: 87-95.

Garrity, A. G., G. D. Pearlson, K. McKiernan, D. Lloyd, K. A. Kiehl and V. D. Calhoun (2007). "Aberrant “default mode” functional connectivity in schizophrenia." American journal of psychiatry **164**(3): 450-457.

Gatt, J. M., S. A. Kuan, C. Dobson-Stone, R. H. Paul, R. T. Joffe, A. H. Kemp, E. Gordon, P. R. Schofield and L. M. Williams (2008). "Association between BDNF Val66Met polymorphism and trait depression is mediated via resting EEG alpha band activity." Biological psychology **79**(2): 275-284.

Gelenberg, A. J. and C. L. Chesen (2000). "How fast are antidepressants?" The Journal of clinical psychiatry.

George, M. S. (2010). "Transcranial magnetic stimulation for the treatment of depression." Expert review of neurotherapeutics **10**(11): 1761-1772.

George, M. S., E. M. Wassermann, T. A. Kimbrell, J. T. Little, W. E. Williams, A. L. Danielson, B. D. Greenberg, M. Hallett and R. M. Post (1997). "Mood improvement following daily left prefrontal repetitive transcranial magnetic stimulation in patients with depression: a placebo-controlled crossover trial." American Journal of Psychiatry **154**(12): 1752-1756.

George, M. S., E. M. Wassermann, W. A. Williams, A. Callahan, T. A. Ketter, P. Basser, M. Hallett and R. M. Post (1995). "Daily repetitive transcranial magnetic stimulation (rTMS) improves mood in depression." Neuroreport: An International Journal for the Rapid Communication of Research in Neuroscience.

Gillman, P. (2007). "Tricyclic antidepressant pharmacology and therapeutic drug interactions updated." British journal of pharmacology **151**(6): 737-748.

Goddard, A. W., S. G. Ball, J. Martinez, M. J. Robinson, C. R. Yang, J. M. Russell and A. Shekhar (2010). "Current perspectives of the roles of the central norepinephrine system in anxiety and depression." Depression and anxiety **27**(4): 339-350.

Gold, C., J. Fachner and J. Erkikila (2013). "Validity and reliability of electroencephalographic frontal alpha asymmetry and frontal midline theta as biomarkers for depression." Scandinavian Journal of Psychology **54**(2): 118-126.

Goldstein, B. J. and P. J. Goodnick (1998). "Selective serotonin reuptake inhibitors in the treatment of affective disorders—III. Tolerability, safety and pharmacoeconomics." Journal of Psychopharmacology **12**(4_suppl): 55-S87.

Gonul, A. S., O. Kitis, E. Ozan, F. Akdeniz, C. Eker, O. D. Eker and S. Vahip (2006). "The effect of antidepressant treatment on N-acetyl aspartate levels of medial frontal cortex in drug-free depressed patients." Progress in Neuro-Psychopharmacology and Biological Psychiatry **30**(1): 120-125.

Goodwin, G. M. (1996). "How do antidepressants affect serotonin receptors? The role of serotonin receptors in the therapeutic and side effect profile of the SSRIs." The Journal of clinical psychiatry.

Gramfort, A., T. Papadopoulo, E. Olivi and M. Clerc (2010). "OpenMEEG: opensource software for quasistatic bioelectromagnetics." Biomedical engineering online **9**(1): 45.

Grech, R., T. Cassar, J. Muscat, K. P. Camilleri, S. G. Fabri, M. Zervakis, P. Xanthopoulos, V. Sakkalis and B. Vanrumste (2008). "Review on solving the inverse problem in EEG source analysis." Journal of neuroengineering and rehabilitation **5**(1): 25.

Greicius, M. (2008). "Resting-state functional connectivity in neuropsychiatric disorders." Current opinion in neurology **21**(4): 424-430.

Greicius, M. D., B. H. Flores, V. Menon, G. H. Glover, H. B. Solvason, H. Kenna, A. L. Reiss and A. F. Schatzberg (2007). "Resting-state functional connectivity in major depression: abnormally increased contributions from subgenual cingulate cortex and thalamus." Biological psychiatry **62**(5): 429-437.

Greicius, M. D., B. Krasnow, A. L. Reiss and V. Menon (2003). "Functional connectivity in the resting brain: a network analysis of the default mode hypothesis." Proceedings of the National Academy of Sciences **100**(1): 253-258.

Greicius, M. D., K. Supekar, V. Menon and R. F. Dougherty (2009). "Resting-state functional connectivity reflects structural connectivity in the default mode network." Cerebral cortex **19**(1): 72-78.

Grieve, S. M., M. S. Korgaonkar, S. H. Koslow, E. Gordon and L. M. Williams (2013). "Widespread reductions in gray matter volume in depression." Neuroimage Clin **3**: 332-339.

Grillner, S. (2006). "Biological pattern generation: the cellular and computational logic of networks in motion." Neuron **52**(5): 751-766.

Grin-Yatsenko, V. A., I. Baas, V. A. Ponomarev and J. D. Kropotov (2009). "EEG power spectra at early stages of depressive disorders." Journal of Clinical Neurophysiology **26**(6): 401-406.

Grin-Yatsenko, V. A., I. Baas, V. A. Ponomarev and J. D. Kropotov (2010). "Independent component approach to the analysis of EEG recordings at early stages of depressive disorders." Clinical Neurophysiology **121**(3): 281-289.

Groppe, D. M., T. P. Urbach and M. Kutas (2011). "Mass univariate analysis of event-related brain potentials/fields II: Simulation studies." Psychophysiology **48**(12): 1726-1737.

Guo, W.-b., X.-l. Sun, L. Liu, Q. Xu, R.-r. Wu, Z.-n. Liu, C.-l. Tan, H.-f. Chen and J.-P. Zhao (2011). "Disrupted regional homogeneity in treatment-resistant depression: a resting-state fMRI study." Progress in Neuro-Psychopharmacology and Biological Psychiatry **35**(5): 1297-1302.

Guyon, I. and A. Elisseeff (2003). "An introduction to variable and feature selection." Journal of machine learning research **3**(Mar): 1157-1182.

Hamilton, J. P., D. J. Furman, C. Chang, M. E. Thomason, E. Dennis and I. H. Gotlib (2011). "Default-mode and task-positive network activity in major depressive disorder: implications for adaptive and maladaptive rumination." Biological psychiatry **70**(4): 327-333.

Hardeveld, F., J. Spijker, R. De Graaf, W. Nolen and A. Beekman (2010). "Prevalence and predictors of recurrence of major depressive disorder in the adult population." Acta Psychiatrica Scandinavica **122**(3): 184-191.

Harmer, C. J., G. M. Goodwin and P. J. Cowen (2009). "Why do antidepressants take so long to work? A cognitive neuropsychological model of antidepressant drug action." The British Journal of Psychiatry **195**(2): 102-108.

Harrison, B. J., J. Pujol, H. Ortiz, A. Fornito, C. Pantelis and M. Yücel (2008). "Modulation of brain resting-state networks by sad mood induction." PLoS One **3**(3): e1794.

Hasler, G., J. W. van der Veen, T. Tumonis, N. Meyers, J. Shen and W. C. Drevets (2007). "Reduced prefrontal glutamate/glutamine and γ-aminobutyric acid levels in major depression determined using proton magnetic resonance spectroscopy." Archives of general psychiatry **64**(2): 193-200.

Hearst, M. A., S. T. Dumais, E. Osuna, J. Platt and B. Scholkopf (1998). "Support vector machines." IEEE Intelligent Systems and their applications **13**(4): 18-28.

Heijnen, W. T., T. K. Birkenhäger, A. I. Wierdsma and W. W. van den Broek (2010). "Antidepressant pharmacotherapy failure and response to subsequent electroconvulsive therapy: a meta-analysis." Journal of clinical psychopharmacology **30**(5): 616-619.

Henriques, J. B. and R. J. Davidson (1990). "Regional brain electrical asymmetries discriminate between previously depressed and healthy control subjects." Journal of abnormal psychology **99**(1): 22.

Henriques, J. B. and R. J. Davidson (1991). "Left frontal hypoactivation in depression." Journal of abnormal psychology **100**(4): 535.

Honey, C. J., R. Kötter, M. Breakspear and O. Sporns (2007). "Network structure of cerebral cortex shapes functional connectivity on multiple time scales." Proceedings of the National Academy of Sciences **104**(24): 10240-10245.

Hosseinifard, B., M. H. Moradi and R. Rostami (2013). "Classifying depression patients and normal subjects using machine learning techniques and nonlinear features from EEG signal." Computer methods and programs in biomedicine **109**(3): 339-345.

Howland, R. H., M. G. Wilson, S. G. Kornstein, A. H. Clayton, M. H. Trivedi, M. M. Wohlreich and M. Fava (2008). "Factors predicting reduced antidepressant response: experience with the SNRI duloxetine in patients with major depression." Annals of Clinical Psychiatry **20**(4): 209-218.

Hoy, K. E. and P. B. Fitzgerald (2010). "Introducing magnetic seizure therapy: A novel therapy for treatment resistant depression." Aust N Z J Psychiatry **44**(7): 591-598.

Hoy, K. E., R. H. Thomson, M. Cherk, K. S. Yap, Z. J. Daskalakis and P. B. Fitzgerald (2013). "Effect of magnetic seizure therapy on regional brain glucose metabolism in major depression." Psychiatry Research: Neuroimaging **211**(2): 169-175.

Hsu, C.-W., C.-C. Chang and C.-J. Lin (2003). "A practical guide to support vector classification."

Hunter, A. M., I. A. Cook and A. F. Leuchter (2007). "The promise of the quantitative electroencephalogram as a predictor of antidepressant treatment outcomes in major depressive disorder." Psychiatric Clinics of North America **30**(1): 105-124.

Hunter, A. M., A. S. Korb, I. A. Cook and A. F. Leuchter (2013). "Rostral anterior cingulate activity in major depressive disorder: state or trait marker of responsiveness to medication?" The Journal of neuropsychiatry and clinical neurosciences **25**(2): 126-133.

Hutcheon, B. and Y. Yarom (2000). "Resonance, oscillation and the intrinsic frequency preferences of neurons." Trends in neurosciences **23**(5): 216-222.

Hutchison, R. M., T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. Della Penna, J. H. Duyn, G. H. Glover and J. Gonzalez-Castillo (2013). "Dynamic functional connectivity: promise, issues, and interpretations." Neuroimage **80**: 360-378.

Iosifescu, D. V. (2011). "Electroencephalography-derived biomarkers of antidepressant response." Harvard review of psychiatry **19**(3): 144-154.

Iosifescu, D. V., S. Greenwald, P. Devlin, D. Mischoulon, J. W. Denninger, J. E. Alpert and M. Fava (2009). "Frontal EEG predictors of treatment outcome in major depressive disorder." European Neuropsychopharmacology **19**(11): 772-777.

Iosifescu, D. V., A. A. Nierenberg, D. Mischoulon, R. H. Perlis, G. I. Papakostas, J. L. Ryan, J. E. Alpert and M. Fava (2005). "An open study of triiodothyronine augmentation of selective serotonin reuptake inhibitors in treatment-resistant major depressive disorder." The Journal of clinical psychiatry **66**(8): 1038-1042.

Irisawa, S., T. Isotani, T. Yagyu, S. Morita, K. Nishida, K. Yamada, M. Yoshimura, G. Okugawa, K. Nobuhara and T. Kinoshita (2006). "Increased omega complexity and decreased microstate duration in nonmedicated schizophrenic patients." Neuropsychobiology **54**(2): 134-139.

Irwin, J. R. and G. H. McClelland (2003). "Negative consequences of dichotomizing continuous predictor variables." Journal of Marketing Research **40**(3): 366-371.

Jann, K., M. Kottlow, T. Dierks, C. Boesch and T. Koenig (2010). "Topographic electrophysiological signatures of fMRI resting state networks." PloS one **5**(9): e12945.

Jaworska, N., P. Blier, W. Fusee and V. Knott (2012). "Alpha power, alpha asymmetry and anterior cingulate cortex activity in depressed males and females." Journal of psychiatric research **46**(11): 1483-1491.

Jaworska, N., H. Wang, D. M. Smith, P. Blier, V. Knott and A. B. Protzner (2018). "Pre-treatment EEG signal variability is associated with treatment success in depression." NeuroImage: Clinical **17**: 368-377.

Jiang, X., K. Xu, J. Hoberman, F. Tian, A. J. Marko, J. F. Waheed, C. R. Harris, A. M. Marini, M.-A. Enoch and R. H. Lipsky (2005). "BDNF variation and mood disorders: a novel functional promoter polymorphism and Val66Met are associated with anxiety but have opposing effects." Neuropsychopharmacology **30**(7): 1353.

John, E. R., L. Prichep, J. Fridman and P. Easton (1988). "Neurometrics: computer-assisted differential diagnosis of brain dysfunctions." Science **239**(4836): 162-169.

Kaiser, D. A. (2007). "What is quantitative EEG?" Journal of Neurotherapy **10**(4): 37-52.

Kaiser, R. H., J. R. Andrews-Hanna, T. D. Wager and D. A. Pizzagalli (2015). "Large-scale network dysfunction in major depressive disorder: a meta-analysis of resting-state functional connectivity." JAMA psychiatry **72**(6): 603-611.

Kaiser, R. H. and D. A. Pizzagalli (2015). "Dysfunctional Connectivity in the Depressed Adolescent Brain." Biol Psychiatry **78**(9): 594-595.

Kalueff, A. V. and D. J. Nutt (2007). "Role of GABA in anxiety and depression." Depression and anxiety **24**(7): 495-517.

Kandel, E. R., J. H. Schwartz, T. M. Jessell, S. A. Siegelbaum and A. J. Hudspeth (2012). Principles of neural science, McGraw-Hill Education / Medical.

Karege, F., G. Perret, G. Bondolfi, M. Schwald, G. Bertschy and J.-M. Aubry (2002). "Decreased serum brain-derived neurotrophic factor levels in major depressed patients." Psychiatry research **109**(2): 143-148.

Katz, L. C. and C. J. Shatz (1996). "Synaptic activity and the construction of cortical circuits." Science **274**(5290): 1133-1138.

Katz, M. M., S. H. Koslow and A. Frazer (1996). "Onset of antidepressant activity: reexamining the structure of depression and multiple actions of drugs." Depression and anxiety **4**(6): 257-267.

Kayser, S., B. Bewernick, A. Matusch, R. Hurlemann, M. Soehle and T. E. Schlaepfer (2015). "Magnetic seizure therapy in treatment-resistant depression: clinical, neuropsychological and metabolic effects." Psychological medicine **45**(5): 1073-1092.

Kayser, S., B. H. Bewernick, C. Grubert, B. L. Hadrysiewicz, N. Axmacher and T. E. Schlaepfer (2011). "Antidepressant effects, of magnetic seizure therapy and electroconvulsive therapy, in treatment-resistant depression." Journal of psychiatric research **45**(5): 569-576.

Kayser, S., B. H. Bewernick, R. Hurlemann, M. Soehle and T. E. Schlaepfer (2013). "Comparable seizure characteristics in magnetic seizure therapy and electroconvulsive therapy for major depression." European Neuropsychopharmacology **23**(11): 1541-1550.

Keller, M. B. and R. J. Boland (1998). "Implications of failing to achieve successful long-term maintenance treatment of recurrent unipolar major depression." Biological psychiatry **44**(5): 348-360.

Kellner, C. H., R. M. Greenberg, J. W. Murrough, E. O. Bryson, M. C. Briggs and R. M. Pasculli (2012). "ECT in treatment-resistant depression." American Journal of Psychiatry **169**(12): 1238-1244.

Kemp, A. H., E. Gordon, A. J. Rush and L. M. Williams (2008). "Improving the prediction of treatment response in depression: integration of clinical, cognitive, psychophysiological, neuroimaging, and genetic measures." CNS spectrums **13**(12): 1066-1086.

Kennedy, S. H., K. R. Evans, S. Krüger, H. S. Mayberg, J. H. Meyer, S. McCann, A. I. Arifuzzman, S. Houle and F. J. Vaccarino (2001). "Changes in regional brain glucose metabolism measured with positron emission tomography after paroxetine treatment of major depression." American Journal of Psychiatry **158**(6): 899-905.

Kennedy, S. H., R. W. Lam, S. Rotzinger, R. V. Milev, P. Blier, J. Downar, K. R. Evans, F. Farzan, J. A. Foster, B. N. Frey, P. Giacobbe, G. B. Hall, K. L. Harkness, S. Hassel, Z. Ismail, F. Leri, S. McInerney, G. M.

MacQueen, L. Minuzzi, D. J. Muller, S. V. Parikh, P. F. M, L. C. Quilty, A. Ravindran, R. B. Sassi, C. N. Soares, S. C. Strother, G. Turecki, A. L. Vaccarino, F. Vila-Rodriguez, J. Yu, R. Uher and CANBIND-Investigator-Team. (In Press, 2018). "Symptomatic and functional outcomes and early prediction of response to escitalopram monotherapy and sequential adjunctive aripiprazole therapy in patients with major depressive disorder: A CAN-BIND-1 report." Journal of Clinical Psychology.

Kessler, R. C., P. Berglund, O. Demler, R. Jin, D. Koretz, K. R. Merikangas, A. J. Rush, E. E. Walters and P. S. Wang (2003). "The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R)." Jama **289**(23): 3095-3105.

Khanna, A., A. Pascual-Leone and F. Farzan (2014). "Reliability of resting-state microstate features in electroencephalography." PloS one **9**(12): e114163.

Kho, K. H., M. F. van Vreeswijk, S. Simpson and A. H. Zwinderman (2003). "A meta-analysis of electroconvulsive therapy efficacy in depression." The journal of ECT **19**(3): 139-147.

Khodayari-Rostamabad, A., J. P. Reilly, G. M. Hasey, H. de Bruin and D. J. MacCrimmon (2013). "A machine learning approach using EEG data to predict response to SSRI treatment for major depressive disorder." Clinical Neurophysiology **124**(10): 1975-1985.

Kikuchi, M., T. Koenig, T. Munesue, A. Hanaoka, W. Strik, T. Dierks, Y. Koshino and Y. Minabe (2011). "EEG microstate analysis in drug-naive patients with panic disorder." PLoS One **6**(7): e22912.

Kikuchi, M., T. Koenig, Y. Wada, M. Higashima, Y. Koshino, W. Strik and T. Dierks (2007). "Native EEG and treatment effects in neuroleptic-naive schizophrenic patients: time and frequency domain approaches." Schizophrenia research **97**(1): 163-172.

Kimbrell, T. A., T. A. Ketter, M. S. George, J. T. Little, B. E. Benson, M. W. Willis, P. Herscovitch and R. M. Post (2002). "Regional cerebral glucose utilization in patients with a range of severities of unipolar depression." Biological psychiatry **51**(3): 237-252.

Kinoshita, T., W. Strik, C. Michel, T. Yagyu, M. Saito and D. Lehmann (1995). "Microstate segmentation of spontaneous multichannel EEG map series under diazepam and sulpiride." Pharmacopsychiatry **28**(02): 51-55.

Klerman, G. L. and M. M. Weissman (1994). Interpersonal psychotherapy of depression: A brief, focused, specific strategy, Jason Aronson, Incorporated.

Klimesch, W. (1999). "EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis." Brain research reviews **29**(2-3): 169-195.

Knott, V., C. Mahoney, S. Kennedy and K. Evans (2000). "Pre-treatment EEG and it's relationship to depression severity and paroxetine treatment outcome." Pharmacopsychiatry **33**(06): 201-205.

Knott, V., C. Mahoney, S. Kennedy and K. Evans (2001). "EEG power, frequency, asymmetry and coherence in male depression." Psychiatry Research: Neuroimaging **106**(2): 123-140.

Knott, V., C. Mahoney, S. Kennedy and K. Evans (2002). "EEG correlates of acute and chronic paroxetine treatment in depression." Journal of affective disorders **69**(1-3): 241-249.

Knott, V. J. and Y. D. Lapierre (1987). "Computerized EEG correlates of depression and antidepressant treatment." Progress in neuro-psychopharmacology & biological psychiatry.

Knott, V. J., J. I. Telner, Y. D. Lapierre, M. Browne and E. R. Horn (1996). "Quantitative EEG in the prediction of antidepressant response to imipramine." Journal of Affective Disorders **39**(3): 175-184.

Knutson, B., J. P. Bhanji, R. E. Cooney, L. Y. Atlas and I. H. Gotlib (2008). "Neural responses to monetary incentives in major depression." Biological psychiatry **63**(7): 686-692.

Koch, K. and C. C. Schultz (2014). "Clinical and pathogenetic implications of occipital bending in depression." Brain **137**(6): 1576-1578.

Koenig, T., L. Prichep, D. Lehmann, P. V. Sosa, E. Braeker, H. Kleinlogel, R. Isenhart and E. R. John (2002). "Millisecond by millisecond, year by year: normative EEG microstates and developmental stages." Neuroimage **16**(1): 41-48.

Kok, A. (1997). "Event-related-potential (ERP) reflections of mental resources: a review and synthesis." Biological psychology **45**(1-3): 19-56.

Korb, A. S., I. A. Cook, A. M. Hunter and A. F. Leuchter (2008). "Brain electrical source differences between depressed subjects and healthy controls." Brain topography **21**(2): 138-146.

Korb, A. S., A. M. Hunter, I. A. Cook and A. F. Leuchter (2009). "Rostral anterior cingulate cortex theta current density and response to antidepressants and placebo in major depression." Clinical Neurophysiology **120**(7): 1313-1319.

Korb, A. S., A. M. Hunter, I. A. Cook and A. F. Leuchter (2011). "Rostral anterior cingulate cortex activity and early symptom improvement during treatment for major depressive disorder." Psychiatry Research: Neuroimaging **192**(3): 188-194.

Krishnan, V. and E. J. Nestler (2008). "The molecular neurobiology of depression." Nature **455**(7215): 894.

Krishnan, V. and E. J. Nestler (2010). "Linking molecules to mood: new insight into the biology of depression." American Journal of Psychiatry **167**(11): 1305-1320.

Kriss, A., A. Halliday, E. Halliday and R. Pratt (1978). "EEG immediately after unilateral ECT." Acta Psychiatrica Scandinavica **58**(3): 231-244.

Krystal, A. D., C. E. Coffey, R. D. Weiner and T. Holsinger (1998). "Changes in seizure threshold over the course of electroconvulsive therapy affect therapeutic response and are detected by ictal EEG ratings." The Journal of neuropsychiatry and clinical neurosciences **10**(2): 178-186.

Krystal, A. D., H. S. Greenside, R. D. Weiner and D. Gassert (1996). "A comparison of EEG signal dynamics in waking, after anesthesia induction and during electroconvulsive therapy seizures." Clinical Neurophysiology **99**(2): 129-140.

Krystal, A. D., R. D. Weiner, W. V. McCall, F. E. Shelp, R. Arias and P. Smith (1993). "The effects of ECT stimulus dose and electrode placement on the ictal electroencephalogram: an intraindividual crossover study." Biological Psychiatry **34**(11): 759-767.

Kumari, V., M. T. Mitterschiffthaler, J. D. Teasdale, G. S. Malhi, R. G. Brown, V. Giampietro, M. J. Brammer, L. Poon, A. Simmons and S. C. Williams (2003). "Neural abnormalities during cognitive generation of affect in treatment-resistant depression." Biological psychiatry **54**(8): 777-791.

Kwon, J. S., T. Youn and H. Y. Jung (1996). "Right hemisphere abnormalities in major depression: quantitative electroencephalographic findings before and after treatment." Journal of affective disorders **40**(3): 169-173.

Lam, R. W., R. Milev, S. Rotzinger, A. C. Andreazza, P. Blier, C. Brenner, Z. J. Daskalakis, M. Dharsee, J. Downar and K. R. Evans (2016). "Discovering biomarkers for antidepressant response: protocol from the Canadian biomarker integration network in depression (CAN-BIND) and clinical characteristics of the first patient cohort." BMC psychiatry **16**(1): 105.

Lamm, C. and T. Singer (2010). "The role of anterior insular cortex in social emotions." Brain Structure and Function **214**(5-6): 579-591.

Lanzenberger, R., P. Baldinger, A. Hahn, J. Ungersboeck, M. Mitterhauser, D. Winkler, Z. Micskei, P. Stein, G. Karanikas and W. Wadsak (2013). "Global decrease of serotonin-1A receptor binding after electroconvulsive therapy in major depression measured by PET." Molecular psychiatry **18**(1): 93.

Laughlin, S. B. and T. J. Sejnowski (2003). "Communication in neuronal networks." Science **301**(5641): 1870-1874.

Lawrie, S. M., C. Buechel, H. C. Whalley, C. D. Frith, K. J. Friston and E. C. Johnstone (2002). "Reduced frontotemporal functional connectivity in schizophrenia associated with auditory hallucinations." Biological psychiatry **51**(12): 1008-1011.

Lee, B.-H. and Y.-K. Kim (2010). "The roles of BDNF in the pathophysiology of major depression and in antidepressant treatment." Psychiatry investigation **7**(4): 231-235.

Lee, J.-S., B.-H. Yang, J.-H. Lee, J.-H. Choi, I.-G. Choi and S.-B. Kim (2007). "Detrended fluctuation analysis of resting EEG in depressed outpatients and healthy controls." Clinical Neurophysiology **118**(11): 2489-2496.

Lehmann, D., P. L. Faber, S. Galderisi, W. M. Herrmann, T. Kinoshita, M. Koukkou, A. Mucci, R. D. Pascual-Marqui, N. Saito and J. Wackermann (2005). "EEG microstate duration and syntax in acute, medication-naive, first-episode schizophrenia: a multi-center study." Psychiatry Research: Neuroimaging **138**(2): 141-156.

Lehmann, D., H. Ozaki and I. Pal (1987). "EEG alpha map series: brain micro-states by space-oriented adaptive segmentation." Electroencephalography and clinical neurophysiology **67**(3): 271-288.

Lerner, D. and R. M. Henke (2008). "What does research tell us about depression, job performance, and work productivity?" Journal of Occupational and Environmental Medicine **50**(4): 401-410.

Leuchter, A. F., I. A. Cook, Y. Jin and B. Phillips (2013). "The relationship between brain oscillatory activity and therapeutic effectiveness of transcranial magnetic stimulation in the treatment of major depressive disorder." Frontiers in human neuroscience **7**: 37.

Leuchter, A. F., I. A. Cook, R. B. Lufkin, J. Dunkin, T. F. Newton, J. L. Cummings, J. K. Mackey and D. O. Walter (1994). "Cordance: a new method for assessment of cerebral perfusion and metabolism using quantitative electroencephalography." Neuroimage **1**(3): 208-219.

Leuchter, A. F., I. A. Cook, L. B. Marangell, W. S. Gilmer, K. S. Burgoyne, R. H. Howland, M. H. Trivedi, S. Zisook, R. Jain and J. T. McCracken (2009). "Comparative effectiveness of biomarkers and clinical indicators for predicting outcomes of SSRI treatment in Major Depressive Disorder: results of the BRITE-MD study." Psychiatry Research **169**(2): 124-131.

Leuchter, A. F., I. A. Cook, S. Uijtdehaage, J. Dunkin, R. B. Lufkin, C. Anderson-Hanley, M. Abrams, S. Rosenberg-Thompson, R. O'Hara and S. L. Simon (1997). "Brain structure and function and the outcomes of treatment for depression." The Journal of clinical psychiatry **58**: 22-31.

Leuchter, A. F., A. M. Hunter, D. E. Krantz and I. A. Cook (2014). "Intermediate phenotypes and biomarkers of treatment outcome in major depressive disorder." Dialogues in clinical neuroscience **16**(4): 525.

Leuchter, A. F., A. M. Hunter, D. E. Krantz and I. A. Cook (2015). "Rhythms and blues: modulation of oscillatory synchrony and the mechanism of action of antidepressant treatments." Annals of the New York Academy of Sciences **1344**(1): 78-91.

Leuchter, A. F., S. H. Uijtdehaage, I. A. Cook, R. O'Hara and M. Mandelkern (1999). "Relationship between brain electrical activity and cortical perfusion in normal subjects." Psychiatry Research: Neuroimaging **90**(2): 125-140.

Li, B., L. Liu, K. J. Friston, H. Shen, L. Wang, L.-L. Zeng and D. Hu (2013). "A treatment-resistant default mode subnetwork in major depression." Biological psychiatry **74**(1): 48-54.

Li, Y., Y. Li, S. Tong, Y. Tang and Y. Zhu (2007). More normal EEGs of depression patients during mental arithmetic than rest. Noninvasive Functional Source Imaging of the Brain and Heart and the International Conference on Functional Biomedical Imaging, 2007. NFSI-ICFBI 2007. Joint Meeting of the 6th International Symposium on, IEEE.

Li, Y., S. Tong, D. Liu, Y. Gai, X. Wang, J. Wang, Y. Qiu and Y. Zhu (2008). "Abnormal EEG complexity in patients with schizophrenia and depression." Clinical Neurophysiology **119**(6): 1232-1241.

Lieber, A. L. and L. S. Prichep (1988). "Diagnosis and subtyping of depressive disorders by quantitative electroencephalography: I. Discriminant analysis of selected variables in untreated depressives." Hillside Journal of Clinical Psychiatry.

Lindvall, O. and A. Björklund (1978). "Anatomy of the dopaminergic neuron systems in the rat brain." Advances in biochemical psychopharmacology **19**: 1-23.

Lisanby, S. H., B. Luber, T. E. Schlaepfer and H. A. Sackeim (2003). "Safety and feasibility of magnetic seizure therapy (MST) in major depression: randomized within-subject comparison with electroconvulsive therapy." Neuropsychopharmacology **28**(10): 1852.

Lisanby, S. H., J. H. Maddox, J. Prudic, D. P. Devanand and H. A. Sackeim (2000). "The effects of electroconvulsive therapy on memory of autobiographical and public events." Arch Gen Psychiatry **57**(6): 581-590.

Livingston, M. G. and H. M. Livingston (1996). "Monoamine oxidase inhibitors." Drug safety **14**(4): 219-227.

Llinás, R. R. (1988). "The intrinsic electrophysiological properties of mammalian neurons: insights into central nervous system function." Science **242**(4886): 1654-1664.

Lohoff, F. W. (2010). "Overview of the genetics of major depressive disorder." Current psychiatry reports **12**(6): 539-546.

López-Muñoz, F. and C. Alamo (2009). "Monoaminergic neurotransmission: the history of the discovery of antidepressants from 1950s until today." Current pharmaceutical design **15**(14): 1563-1586.

Lorenzetti, V., N. B. Allen, A. Fornito and M. Yücel (2009). "Structural brain abnormalities in major depressive disorder: a selective review of recent MRI studies." Journal of affective disorders **117**(1-2): 1-17.

Lubar, J. F., M. Congedo and J. H. Askew (2003). "Low-resolution electromagnetic tomography (LORETA) of cerebral activity in chronic depressive disorder." International Journal of Psychophysiology **49**(3): 175-185.

Lucki, I. (1998). "The spectrum of behaviors influenced by serotonin." Biological psychiatry.

MacQueen, G. M., K. Yucel, V. H. Taylor, K. Macdonald and R. Joffe (2008). "Posterior hippocampal volumes are associated with remission rates in patients with major depressive disorder." Biological psychiatry **64**(10): 880-883.

Mafahim, J. U., D. Lambert, M. Zare and P. Grigolini (2015). "Complexity matching in neural networks." New Journal of Physics **17**(1): 015003.

Maller, J. J., R. H. Thomson, J. V. Rosenfeld, R. Anderson, Z. J. Daskalakis and P. B. Fitzgerald (2014). "Occipital bending in depression." Brain **137**(6): 1830-1837.

Maller, J. J., R. H. Thomson, J. V. Rosenfeld, R. Anderson, Z. J. Daskalakis and P. B. Fitzgerald (2015). "Reply: occipital bending in depression." Brain **138**(Pt 1): e318.

Malmivuo, P., J. Malmivuo and R. Plonsey (1995). Bioelectromagnetism: principles and applications of bioelectric and biomagnetic fields, Oxford University Press, USA.

Mann, J. J. (1999). "Role of the serotonergic system in the pathogenesis of major depression and suicidal behavior." Neuropsychopharmacology **21**(2): 99S-105S.

Manoliu, A., C. Meng, F. Brandl, A. Doll, M. Tahmasian, M. Scherr, D. Schwerthöffer, C. Zimmer, H. Förstl and J. Bäuml (2014). "Insular dysfunction within the salience network is associated with severity of symptoms and aberrant inter-network connectivity in major depressive disorder." Frontiers in human neuroscience **7**: 930.

Maris, E. and R. Oostenveld (2007). "Nonparametric statistical testing of EEG-and MEG-data." Journal of neuroscience methods **164**(1): 177-190.

Marmelat, V. and D. Delignières (2012). "Strong anticipation: complexity matching in interpersonal coordination." Experimental Brain Research **222**(1-2): 137-148.

Masand, P. S. and S. Gupta (2002). "Long-term side effects of newer-generation antidepressants: SSRIS, venlafaxine, nefazodone, bupropion, and mirtazapine." Annals of Clinical Psychiatry **14**(3): 175-182.

Massimini, M., F. Ferrarelli, R. Huber, S. K. Esser, H. Singh and G. Tononi (2005). "Breakdown of cortical effective connectivity during sleep." Science **309**(5744): 2228-2232.

Mayberg, H. S. (2003). "Modulating dysfunctional limbic-cortical circuits in depression: towards development of brain-based algorithms for diagnosis and optimised treatment." British medical bulletin **65**(1): 193-207.

Mayberg, H. S., S. K. Brannan, R. K. Mahurin, P. A. Jerabek, J. S. Brickman, J. L. Tekell, J. A. Silva, S. McGinnis, T. G. Glass and C. C. Martin (1997). "Cingulate function in depression: a potential predictor of treatment response." Neuroreport **8**(4): 1057-1061.

Mayberg, H. S., M. Liotti, S. K. Brannan, S. McGinnis, R. K. Mahurin, P. A. Jerabek, J. A. Silva, J. L. Tekell, C. C. Martin and J. L. Lancaster (1999). "Reciprocal limbic-cortical function and negative mood: converging PET findings in depression and normal sadness." American Journal of Psychiatry **156**(5): 675-682.

McAllister-Williams, R. and A. Massey (2003). "EEG effects of buspirone and pindolol: a method of examining 5-HT 1A receptor function in humans." Psychopharmacology **166**(3): 284-293.

McClintock, S. M., J. Choi, Z. D. Deng, L. G. Appelbaum, A. D. Krystal and S. H. Lisanby (2014). "Multifactorial determinants of the neurocognitive effects of electroconvulsive therapy." J ECT **30**(2): 165-176.

McClintock, S. M., N. K. DeWind, M. M. Husain, S. B. Rowny, T. J. Spellman, H. Terrace and S. H. Lisanby (2013). "Disruption of component processes of spatial working memory by electroconvulsive shock but not magnetic seizure therapy." Int J Neuropsychopharmacol **16**(1): 177-187.

McCormick, L. M., T. Yamada, M. Yeh, M. C. Brumm and R. W. Thatcher (2009). "Antipsychotic effect of electroconvulsive therapy is related to normalization of subgenual cingulate theta activity in psychotic depression." Journal of psychiatric research **43**(5): 553-560.

McDonough, I. M. and K. Nashiro (2014). "Network complexity as a measure of information processing across resting-state networks: evidence from the Human Connectome Project." Frontiers in human neuroscience **8**: 409.

McIntosh, A. and F. Gonzalez-Lima (1998). "Large-scale functional connectivity in associative learning: interrelations of the rat auditory, visual, and limbic systems." Journal of Neurophysiology **80**(6): 3148-3162.

McIntosh, A., V. Vakorin, N. Kovacevic, H. Wang, A. Diaconescu and A. Protzner (2013). "Spatiotemporal dependency of age-related changes in brain signal variability." Cerebral Cortex **24**(7): 1806-1817.

McIntosh, A. R., V. Vakorin, N. Kovacevic, H. Wang, A. Diaconescu and A. B. Protzner (2014). "Spatiotemporal dependency of age-related changes in brain signal variability." Cereb Cortex **24**(7): 1806-1817.

McNally, K. A. and H. Blumenfeld (2004). "Focal network involvement in generalized seizures: new insights from electroconvulsive therapy." Epilepsy & Behavior **5**(1): 3-12.

Meltzer, H. Y. (1990). "Role of Serotonin in Depression a." Annals of the New York Academy of Sciences **600**(1): 486-499.

Méndez, M. A., P. Zuluaga, R. Hornero, C. Gómez, J. Escudero, A. Rodríguez-Palancas, T. Ortiz and A. Fernández (2012). "Complexity analysis of spontaneous brain activity: effects of depression and antidepressant treatment." Journal of Psychopharmacology **26**(5): 636-643.

Meng, C., F. Brandl, M. Tahmasian, J. Shao, A. Manoliu, M. Scherr, D. Schwerthoffer, J. Bauml, H. Forstl, C. Zimmer, A. M. Wohlschlager, V. Riedl and C. Sorg (2014). "Aberrant topology of striatum's connectivity is associated with the number of episodes in depression." Brain **137**(Pt 2): 598-609.

Menon, V. and L. Q. Uddin (2010). "Saliency, switching, attention and control: a network model of insula function." Brain Structure and Function **214**(5-6): 655-667.

Mensen, A. and R. Khatami (2013). "Advanced EEG analysis using threshold-free cluster-enhancement and non-parametric statistics." Neuroimage **67**: 111-118.

Michel, C. M. and T. Koenig (2017). "EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: A review." NeuroImage.

Mientus, S., J. Gallinat, Y. Wuebben, R. D. Pascual-Marqui, C. Mulert, K. Frick, H. Dorn, W. M. Herrmann and G. Winterer (2002). "Cortical hypoactivation during resting EEG in schizophrenics but not in depressives and schizotypal subjects as revealed by low resolution electromagnetic tomography (LORETA)." Psychiatry Research: Neuroimaging **116**(1-2): 95-111.

Milz, P., P. L. Faber, D. Lehmann, T. Koenig, K. Kochi and R. D. Pascual-Marqui (2016). "The functional significance of EEG microstates—associations with modalities of thinking." Neuroimage **125**: 643-656.

Milz, P., R. D. Pascual-Marqui, D. Lehmann and P. L. Faber (2016). "Modalities of Thinking: State and Trait Effects on Cross-Frequency Functional Independent Brain Networks." Brain topography **29**(3): 477-490.

Misic, B., T. Mills, M. J. Taylor and A. R. McIntosh (2010). "Brain noise is task dependent and region specific." J Neurophysiol **104**(5): 2667-2676.

Mizuno, T., T. Takahashi, R. Y. Cho, M. Kikuchi, T. Murata, K. Takahashi and Y. Wada (2010). "Assessment of EEG dynamical complexity in Alzheimer's disease using multiscale entropy." Clin Neurophysiol **121**(9): 1438-1446.

Modinos, G., J. Ormel and A. Aleman (2009). "Activation of anterior insula during self-reflection." PloS one **4**(2): e4618.

Molendijk, M. L., B. A. Bus, P. Spinhoven, B. W. Penninx, G. Kenis, J. Prickaerts, R. O. Voshaar and B. M. Elzinga (2011). "Serum levels of brain-derived neurotrophic factor in major depressive disorder: state–trait issues, clinical features and pharmacological treatment." Molecular psychiatry **16**(11): 1088.

Molinaro, A. M., R. Simon and R. M. Pfeiffer (2005). "Prediction error estimation: a comparison of resampling methods." Bioinformatics **21**(15): 3301-3307.

Moret, C. and M. Briley (2011). "The importance of norepinephrine in depression." Neuropsychiatric disease and treatment **7**(Suppl 1): 9.

Moscrip, T. D., H. S. Terrace, H. A. Sackeim and S. H. Lisanby (2006). "Randomized controlled trial of the cognitive side-effects of magnetic seizure therapy (MST) and electroconvulsive shock (ECS)." The The International Journal of Neuropsychopharmacology **9**(1): 1-11.

Mulders, P. C., P. F. van Eijndhoven, A. H. Schene, C. F. Beckmann and I. Tendolkar (2015). "Resting-state functional connectivity in major depressive disorder: a review." Neuroscience & Biobehavioral Reviews **56**: 330-344.

Mulert, C., G. Juckel, M. Brunnmeier, S. Karch, G. Leicht, R. Mergl, H.-J. Möller, U. Hegerl and O. Pogarell (2007). "Prediction of treatment response in major depression: integration of concepts." Journal of affective disorders **98**(3): 215-225.

Mulert, C., G. Juckel, M. Brunnmeier, S. Karch, G. Leicht, R. Mergl, H.-J. Möller, U. Hegerl and O. Pogarell (2007). "Rostral anterior cingulate cortex activity in the theta band predicts response to antidepressive medication." Clinical EEG and neuroscience **38**(2): 78-81.

Murray, C. J., A. D. Lopez and W. H. Organization (1996). "The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020: summary."

Murray, C. J. L. and A. D. Lopez (1996). The Global Burden of Disease: A Comprehensive Assessment of Mortality and Disability from Diseases, Injuries, and Risk Factors in 1990 and Projected to 2020 Cambridge, MA, Harvard University Press

Musso, F., J. Brinkmeyer, A. Mobascher, T. Warbrick and G. Winterer (2010). "Spontaneous brain activity and EEG microstates. A novel EEG/fMRI analysis approach to explore resting-state networks." Neuroimage **52**(4): 1149-1161.

Nandrino, J.-L., L. Pezard, J. Martinerie, F. El Massioui and J.-F. Allilaireº (1994). "Decrease of complexity in EEG as a symptom of." Neuroreport **5**: 530.

Natarajan, K., R. Acharya, F. Alias, T. Tiboleng and S. K. Puthusserypady (2004). "Nonlinear analysis of EEG signals at different mental states." BioMedical Engineering OnLine **3**(1): 7.

Niedermeyer, E. and F. L. da Silva (2005). Electroencephalography: basic principles, clinical applications, and related fields, Lippincott Williams & Wilkins.

Nishida, K., Y. Morishima, M. Yoshimura, T. Isotani, S. Irisawa, K. Jann, T. Dierks, W. Strik, T. Kinoshita and T. Koenig (2013). "EEG microstates associated with salience and frontoparietal networks in frontotemporal dementia, schizophrenia and Alzheimer's disease." Clinical Neurophysiology **124**(6): 1106-1114.

Nobler, M. S., M. A. Oquendo, L. S. Kegeles, K. M. Malone, C. Campbell, H. A. Sackeim and J. J. Mann (2001). "Decreased regional brain metabolism after ECT." American Journal of Psychiatry **158**(2): 305-308.

Nobler, M. S. and H. A. Sackeim (2008). "Neurobiological correlates of the cognitive side effects of electroconvulsive therapy." J ECT **24**(1): 40-45.

Nord, M., S. J. Finnema, C. Halldin and L. Farde (2013). "Effect of a single dose of escitalopram on serotonin concentration in the non-human and human primate brain." International Journal of Neuropsychopharmacology **16**(7): 1577-1586.

Nunez, P. L., R. B. Silberstein, Z. Shi, M. R. Carpenter, R. Srinivasan, D. M. Tucker, S. M. Doran, P. J. Cadusch and R. S. Wijesinghe (1999). "EEG coherency II: experimental comparisons of multiple measures." Clinical Neurophysiology **110**(3): 469-486.

Nunez, P. L., R. Srinivasan, A. F. Westdorp, R. S. Wijesinghe, D. M. Tucker, R. B. Silberstein and P. J. Cadusch (1997). "EEG coherency: I: statistics, reference electrode, volume conduction, Laplacians, cortical imaging, and interpretation at multiple scales." Electroencephalography and clinical neurophysiology **103**(5): 499-515.

Nystrom, C., M. Matousek and T. Hallstrom (1986). "Relationships between EEG and clinical characteristics in major depressive disorder." Acta psychiatrica scandinavica **73**(4): 390-394.

Okazaki, R., T. Takahashi, K. Ueno, K. Takahashi, M. Higashima and Y. Wada (2013). "Effects of electroconvulsive therapy on neural complexity in patients with depression: report of three cases." Journal of affective disorders **150**(2): 389-392.

Olbrich, S. and M. Arns (2013). "EEG biomarkers in major depressive disorder: discriminative power and prediction of treatment response." International Review of Psychiatry **25**(5): 604-618.

Olchanski, N., M. M. Myers, M. Halseth, P. L. Cyr, L. Bockstedt, T. F. Goss and R. H. Howland (2013). "The economic burden of treatment-resistant depression." Clinical therapeutics **35**(4): 512-522.

Orru, G., W. Pettersson-Yeo, A. F. Marquand, G. Sartori and A. Mechelli (2012). "Using support vector machine to identify imaging biomarkers of neurological and psychiatric disease: a critical review." Neuroscience & Biobehavioral Reviews **36**(4): 1140-1152.

Owen, S. V. and R. D. Froman (2005). "Why carve up your continuous data?" Research in nursing & health **28**(6): 496-503.

Owens, M. J. and C. B. Nemeroff (1994). "Role of serotonin in the pathophysiology of depression: focus on the serotonin transporter." Clinical chemistry **40**(2): 288-295.

Pandya, M., M. Altinay, D. A. Malone and A. Anand (2012). "Where in the brain is depression?" Current psychiatry reports **14**(6): 634-642.

Parsey, R. V., R. S. Hastings, M. A. Oquendo, Y.-y. Huang, N. Simpson, J. Arcement, Y. Huang, R. T. Ogden, R. L. Van Heertum and V. Arango (2006). "Lower serotonin transporter binding potential in the human brain during major depressive episodes." American Journal of Psychiatry **163**(1): 52-58.

Pascual-Marqui, R. D. (1999). "Review of methods for solving the EEG inverse problem." International journal of bioelectromagnetism **1**(1): 75-86.

Pascual-Marqui, R. D. (2002). "Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details." Methods Find Exp Clin Pharmacol **24**(Suppl D): 5-12.

Pascual-Marqui, R. D., D. Lehmann, P. Faber, P. Milz, K. Kochi, M. Yoshimura, K. Nishida, T. Isotani and T. Kinoshita (2014). "The resting microstate networks (RMN): cortical distributions, dynamics, and frequency specific information flow." arXiv preprint arXiv:1411.1949.

Pascual-Marqui, R. D., D. Lehmann, T. Koenig, K. Kochi, M. C. Merlo, D. Hell and M. Koukkou (1999). "Low resolution brain electromagnetic tomography (LORETA) functional imaging in acute, neuroleptic-naive, first-episode, productive schizophrenia." Psychiatry Research: Neuroimaging **90**(3): 169-179.

Pascual-Marqui, R. D., C. M. Michel and D. Lehmann (1995). "Segmentation of brain electrical activity into microstates: model estimation and validation." Biomedical Engineering, IEEE Transactions on **42**(7): 658-665.

Paus, T. (2000). Functional anatomy of arousal and attention systems in the human brain. Progress in brain research, Elsevier. **126:** 65-77.

Paus, T., P. Sipila and A. Strafella (2001). "Synchronization of neuronal activity in the human primary motor cortex by transcranial magnetic stimulation: an EEG study." Journal of Neurophysiology **86**(4): 1983-1990.

Paykel, E., R. Ramana, Z. Cooper, H. Hayhurst, J. Kerr and A. Barocka (1995). "Residual symptoms after partial remission: an important outcome in depression." Psychological medicine **25**(6): 1171-1180.

Peng, D., E. B. Liddle, S. J. Iwabuchi, C. Zhang, Z. Wu, J. Liu, K. Jiang, L. Xu, P. F. Liddle and L. Palaniyappan (2015). "Dissociated large-scale functional connectivity networks of the precuneus in medication-naive first-episode depression." Psychiatry Research: Neuroimaging **232**(3): 250-256.

Pereda, E., R. Q. Quiroga and J. Bhattacharya (2005). "Nonlinear multivariate analysis of neurophysiological signals." Progress in neurobiology **77**(1-2): 1-37.

Perera, T. D., B. Luber, M. S. Nobler, J. Prudic, C. Anderson and H. A. Sackeim (2004). "Seizure expression during electroconvulsive therapy: relationships with clinical outcome and cognitive side effects." Neuropsychopharmacology **29**(4): 813.

Perlis, R. H. (2013). "A clinical risk stratification tool for predicting treatment resistance in major depressive disorder." Biological psychiatry **74**(1): 7-14.

Pernet, C., M. Latinus, T. Nichols and G. Rousselet (2015). "Cluster-based computational methods for mass univariate analyses of event-related brain potentials/fields: A simulation study." Journal of Neuroscience Methods **250**: 85-93.

Perrin, F., J. Pernier, O. Bertrand and J. Echallier (1989). "Spherical splines for scalp potential and current density mapping." Electroencephalography and clinical neurophysiology **72**(2): 184-187.

Perrin, J. S., S. Merz, D. M. Bennett, J. Currie, D. J. Steele, I. C. Reid and C. Schwarzbauer (2012). "Electroconvulsive therapy reduces frontal cortical connectivity in severe depressive disorder." Proceedings of the National Academy of Sciences **109**(14): 5464-5468.

Peterchev, A. V., M. A. Rosa, Z. D. Deng, J. Prudic and S. H. Lisanby (2010). "Electroconvulsive therapy stimulus parameters: rethinking dosage." J ECT **26**(3): 159-174.

Pittenger, C. and R. S. Duman (2008). "Stress, depression, and neuroplasticity: a convergence of mechanisms." Neuropsychopharmacology **33**(1): 88.

Pizzagalli, D., R. D. Pascual-Marqui, J. B. Nitschke, T. R. Oakes, C. L. Larson, H. C. Abercrombie, S. M. Schaefer, J. V. Koger, R. M. Benca and R. J. Davidson (2001). "Anterior cingulate activity as a predictor of degree of treatment response in major depression: evidence from brain electrical tomography analysis." American Journal of Psychiatry **158**(3): 405-415.

Pizzagalli, D. A. (2011). "Frontocingulate dysfunction in depression: toward biomarkers of treatment response." Neuropsychopharmacology **36**(1): 183.

Pizzagalli, D. A., J. B. Nitschke, T. R. Oakes, A. M. Hendrick, K. A. Horras, C. L. Larson, H. C. Abercrombie, S. M. Schaefer, J. V. Koger and R. M. Benca (2002). "Brain electrical tomography in depression: the importance of symptom severity, anxiety, and melancholic features." Biological psychiatry **52**(2): 73-85.

Pizzagalli, D. A., T. R. Oakes and R. J. Davidson (2003). "Coupling of theta activity and glucose metabolism in the human rostral anterior cingulate cortex: an EEG/PET study of normal and depressed subjects." Psychophysiology **40**(6): 939-949.

Poline, J.-B. and B. Mazoyer (1994). "Analysis of individual brain activation maps using hierarchical description and multiscale detection." IEEE transactions on medical imaging **13**(4): 702-710.

Pollock, V. E. and L. S. Schneider (1990). "Quantitative, waking EEG research on depression." Biological psychiatry **27**(7): 757-780.

Price, G., J. Lee, C. Garvey and N. Gibson (2008). "Appraisal of sessional EEG features as a correlate of clinical changes in an rTMS treatment of depression." Clinical EEG and neuroscience **39**(3): 131-138.

Price, J. L. and W. C. Drevets (2010). "Neurocircuitry of mood disorders." Neuropsychopharmacology **35**(1): 192.

Prichep, L. S. and E. R. John (1992). "QEEG profiles of psychiatric disorders." Brain topography **4**(4): 249-257.

Quitkin, F. M., J. G. Rabkin, D. Ross and J. W. Stewart (1984). "Identification of true drug response to antidepressants: use of pattern analysis." Archives of general psychiatry **41**(8): 782-786.

Rabinoff, M., C. Kitchen, I. Cook and A. Leuchter (2011). "Evaluation of quantitative EEG by classification and regression trees to characterize responders to antidepressant and placebo treatment." The open medical informatics journal **5**: 1.

Raichle, M. E. and A. Z. Snyder (2007). "A default mode of brain function: a brief history of an evolving idea." Neuroimage **37**(4): 1083-1090.

Reid, S. A., L. M. Duke and J. J. Allen (1998). "Resting frontal electroencephalographic asymmetry in depression: Inconsistencies suggest the need to identify mediating factors." Psychophysiology **35**(4): 389-404.

Remijnse, P., M. Nielen, A. Van Balkom, G.-J. Hendriks, W. Hoogendijk, H. Uylings and D. Veltman (2009). "Differential frontal–striatal and paralimbic activity during reversal learning in major depressive disorder and obsessive–compulsive disorder." Psychological medicine **39**(9): 1503-1518.

Ressler, K. J. and C. B. Nemeroff (1999). "Role of norepinephrine in the pathophysiology and treatment of mood disorders." Biological psychiatry **46**(9): 1219-1233.

Richards, D. (2011). "Prevalence and clinical course of depression: a review." Clinical psychology review **31**(7): 1117-1125.

Rieger, K., L. D. Hernandez, A. Baenninger and T. Koenig (2016). "15 Years of Microstate research in schizophrenia–Where are We? a Meta-analysis." Frontiers in psychiatry **7**.

Rigucci, S., G. Serafini, M. Pompili, G. D. Kotzalidis and R. Tatarelli (2010). "Anatomical and functional correlates in major depressive disorder: the contribution of neuroimaging studies." The World Journal of Biological Psychiatry **11**(2-2): 165-180.

Rizvi, S. J., E. Grima, M. Tan, S. Rotzinger, P. Lin, R. S. McIntyre and S. H. Kennedy (2014). "Treatment-resistant depression in primary care across Canada." The Canadian Journal of Psychiatry **59**(7): 349-357.

Robbins, T. W. (2003). "Dopamine and cognition." Current opinion in neurology **16**: S1-S2.

Rodriguez, G., P. Vitali, C. De Leo, F. De Carli, N. Girtler and F. Nobili (2002). "Quantitative EEG changes in Alzheimer patients during long-term donepezil therapy." Neuropsychobiology **46**(1): 49-56.

Roemer, R. A., C. Shagass, W. Dubin, R. Jaffe and L. Siegal (1992). "Quantitative EEG in elderly depressives." Brain topography **4**(4): 285-290.

Rojas, R. (2013). Neural networks: a systematic introduction, Springer Science & Business Media.

Rolls, E. T., A. Treves and E. T. Rolls (1998). Neural networks and brain function, Oxford university press Oxford.

Rozycki, M., T. D. Satterthwaite, N. Koutsouleris, G. Erus, J. Doshi, D. H. Wolf, Y. Fan, R. E. Gur, R. C. Gur and E. M. Meisenzahl (2017). "Multisite machine learning analysis provides a robust structural imaging signature of schizophrenia detectable across diverse patient populations and within individuals." Schizophr Bull **10**.

Rubinov, M. and O. Sporns (2010). "Complex network measures of brain connectivity: uses and interpretations." Neuroimage **52**(3): 1059-1069.

Rush, A. J., M. H. Trivedi, S. R. Wisniewski, A. A. Nierenberg, J. W. Stewart, D. Warden, G. Niederehe, M. E. Thase, P. W. Lavori and B. D. Lebowitz (2006). "Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR* D report." American Journal of Psychiatry **163**(11): 1905-1917.

Rush, A. J., S. R. Wisniewski, D. Warden, J. F. Luther, L. L. Davis, M. Fava, A. A. Nierenberg and M. H. Trivedi (2008). "Selecting among second-step antidepressant medication monotherapies: predictive

value of clinical, demographic, or first-step treatment features." Archives of General Psychiatry **65**(8): 870-880.

Russo-Neustadt, A., R. Beard, Y. Huang and C. Cotman (2000). "Physical activity and antidepressant treatment potentiate the expression of specific brain-derived neurotrophic factor transcripts in the rat hippocampus." Neuroscience **101**(2): 305-312.

Sackeim, H. A. (1999). "The anticonvulsant hypothesis of the mechanisms of action of ECT: current status." The journal of ECT **15**(1): 5-26.

Sackeim, H. A., P. Decina, S. Portnoy, P. Neeley and S. Malitz (1987). "Studies of dosage, seizure threshold, and seizure duration in ECT." Biological psychiatry **22**(3): 249-268.

Sackeim, H. A., B. Luber, G. P. Katzman, J. R. Moeller, J. Prudic, D. Devanand and M. S. Nobler (1996). "The effects of electroconvulsive therapy on quantitative electroencephalograms: relationship to clinical outcome." Archives of general psychiatry **53**(9): 814-824.

Sackeim, H. A., B. Luber, J. R. Moeller, J. Prudic, D. P. Devanand and M. S. Nobler (2000). "Electrophysiological correlates of the adverse cognitive effects of electroconvulsive therapy." J ECT **16**(2): 110-120.

Sackeim, H. A., J. Prudic, M. S. Nobler, L. Fitzsimons, S. H. Lisanby, N. Payne, R. M. Berman, E.-L. Brakemeier, T. Perera and D. Devanand (2008). "Effects of pulse width and electrode placement on the efficacy and cognitive effects of electroconvulsive therapy." Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation **1**(2): 71-83.

Saeb, S., L. Lonini, A. Jayaraman, D. C. Mohr and K. P. Kording (2016). "Voodoo machine learning for clinical predictions." Biorxiv: 059774.

Saeys, Y., I. Inza and P. Larrañaga (2007). "A review of feature selection techniques in bioinformatics." bioinformatics **23**(19): 2507-2517.

Sale, M. V., J. B. Mattingley, A. Zalesky and L. Cocchi (2015). "Imaging human brain networks to improve the clinical efficacy of non-invasive brain stimulation." Neurosci Biobehav Rev **57**: 187-198.

Saletu, B., J. Grünberger and L. Linzmayer (1986). "On central effects of serotonin re-uptake inhibitors: quantitative EEG and psychometric studies with sertraline and zimelidine." Journal of neural transmission **67**(3-4): 241-266.

Salomon, R. M. and R. L. Cowan (2013). "Oscillatory serotonin function in depression." Synapse **67**(11): 801-820.

Sanacora, G., R. Gueorguieva, C. N. Epperson, Y.-T. Wu, M. Appel, D. L. Rothman, J. H. Krystal and G. F. Mason (2004). "Subtype-specific alterations of γ-aminobutyric acid and glutamatein patients with major depression." Archives of general psychiatry **61**(7): 705-713.

Sanacora, G., G. F. Mason, D. L. Rothman, K. L. Behar, F. Hyder, O. A. Petroff, R. M. Berman, D. S. Charney and J. H. Krystal (1999). "Reduced cortical γ-aminobutyric acid levels in depressed patients determined by proton magnetic resonance spectroscopy." Archives of general psychiatry **56**(11): 1043-1047.

Sanacora, G., G. F. Mason, D. L. Rothman, F. Hyder, J. J. Ciarcia, R. B. Ostroff, R. M. Berman and J. H. Krystal (2003). "Increased cortical GABA concentrations in depressed patients receiving ECT." American Journal of Psychiatry **160**(3): 577-579.

Santarnecchi, E., A. R. Khanna, C. S. Musaeus, C. S. Benwell, P. Davila, F. Farzan, S. Matham, A. Pascual-Leone and M. M. Shafi (2017). "EEG microstate correlates of fluid intelligence and response to cognitive training." Brain topography **30**(4): 502-520.

Sarasso, S., M. Rosanova, A. G. Casali, S. Casarotto, M. Fecchio, M. Boly, O. Gosseries, G. Tononi, S. Laureys and M. Massimini (2014). "Quantifying cortical EEG responses to TMS in (un) consciousness." Clinical EEG and neuroscience **45**(1): 40-49.

Saxena, S., A. L. Brody, M. L. Ho, N. Zohrabi, K. M. Maidment and L. R. Baxter Jr (2003). "Differential brain metabolic predictors of response to paroxetine in obsessive-compulsive disorder versus major depression." American Journal of Psychiatry **160**(3): 522-532.

Schaffer, C. E., R. J. Davidson and C. Saron (1983). "Frontal and parietal electroencephalogram asymmetry in depressed and nondepressed subjects." Biological psychiatry.

Schnitzler, A. and J. Gross (2005). "Normal and pathological oscillatory communication in the brain." Nature reviews neuroscience **6**(4): 285.

Schwab, S., T. Koenig, Y. Morishima, T. Dierks, A. Federspiel and K. Jann (2015). "Discovering frequency sensitive thalamic nuclei from EEG microstate informed resting state fMRI." NeuroImage **118**: 368-375.

Schweitzer, I., V. Tuckwell, D. Ames and J. O'brien (2001). "Structural neuroimaging studies in late-life depression: a review." The World Journal of Biological Psychiatry **2**(2): 83-88.

Seeley, W. W., V. Menon, A. F. Schatzberg, J. Keller, G. H. Glover, H. Kenna, A. L. Reiss and M. D. Greicius (2007). "Dissociable intrinsic connectivity networks for salience processing and executive control." Journal of Neuroscience **27**(9): 2349-2356.

Segrave, R. A., N. Cooper, R. Thomson, R. J. Croft, D. Sheppard and P. Fitzgerald (2011). "Individualized alpha activity and frontal asymmetry in major depression." Clinical EEG and Neuroscience **42**(1): 45-52.

Sen, S., R. Duman and G. Sanacora (2008). "Serum brain-derived neurotrophic factor, depression, and antidepressant medications: meta-analyses and implications." Biological psychiatry **64**(6): 527-532.

Serretti, A., P. Olgiati, M. N. Liebman, H. Hu, Y. Zhang, R. Zanardi, C. Colombo and E. Smeraldi (2007). "Clinical prediction of antidepressant response in mood disorders: linear multivariate vs. neural network models." Psychiatry research **152**(2-3): 223-231.

Shaffer, J. P. (1995). "Multiple hypothesis testing." Annual review of psychology **46**(1): 561-584.

Sheehan, D., Y. Lecrubier, K. H. Sheehan, J. Janavs, E. Weiller, A. Keskiner, J. Schinka, E. Knapp, M. Sheehan and G. Dunbar (1997). "The validity of the Mini International Neuropsychiatric Interview (MINI) according to the SCID-P and its reliability." European Psychiatry **12**(5): 232-241.

Sheline, Y. I., D. M. Barch, J. L. Price, M. M. Rundle, S. N. Vaishnavi, A. Z. Snyder, M. A. Mintun, S. Wang, R. S. Coalson and M. E. Raichle (2009). "The default mode network and self-referential processes in depression." Proceedings of the National Academy of Sciences **106**(6): 1942-1947.

Sheline, Y. I., M. H. Gado and J. L. Price (1998). "Amygdala core nuclei volumes are decreased in recurrent major depression." Neuroreport **9**(9): 2023-2028.

Sheline, Y. I., J. L. Price, Z. Yan and M. A. Mintun (2010). "Resting-state functional MRI in depression unmasks increased connectivity between networks via the dorsal nexus." Proceedings of the National Academy of Sciences **107**(24): 11020-11025.

Shelp, B. J., A. W. Bown and M. D. McLean (1999). "Metabolism and functions of gamma-aminobutyric acid." Trends in plant science **4**(11): 446-452.

Shimoda, K. and R. G. Robinson (1999). "The relationship between poststroke depression and lesion location in long-term follow-up." Biol Psychiatry **45**(2): 187-192.

Small, G. W., I. F. Small and V. Milstein (1978). Electrophysiology of ECT. New York, Raven Press.

Smit, D. J., C. J. Stam, D. Posthuma, D. I. Boomsma and E. J. De Geus (2008). "Heritability of "small-world" networks in the brain: A graph theoretical analysis of resting-state EEG functional connectivity." Human brain mapping **29**(12): 1368-1378.

Smith, S. M. and T. E. Nichols (2009). "Threshold-free cluster enhancement: addressing problems of smoothing, threshold dependence and localisation in cluster inference." Neuroimage **44**(1): 83-98.

Solomon, D. A., M. B. Keller, A. C. Leon, T. I. Mueller, M. T. Shea, M. Warshaw, J. D. Maser, W. Coryell and J. Endicott (1997). "Recovery from major depression: a 10-year prospective follow-up across multiple episodes." Archives of general psychiatry **54**(11): 1001-1006.

Souery, D., P. Oswald, I. Massat, U. Bailer, J. Bollen, K. Demyttenaere, S. Kasper, Y. Lecrubier, S. Montgomery and A. Serretti (2007). "Clinical factors associated with treatment resistance in major depressive disorder: results from a European multicenter study." Journal of Clinical Psychiatry **68**(7): 1062-1070.

Souery, D., G. I. Papakostas and M. H. Trivedi (2006). "Treatment-resistant depression." The Journal of clinical psychiatry **67**: 16-22.

Spellman, T., S. M. McClintock, H. Terrace, B. Luber, M. M. Husain and S. H. Lisanby (2008). "Differential effects of high-dose magnetic seizure therapy and electroconvulsive shock on cognitive function." Biological psychiatry **63**(12): 1163-1170.

Sporns, O., G. Tononi and G. M. Edelman (2000). "Connectivity and complexity: the relationship between neuroanatomy and brain dynamics." Neural networks **13**(8-9): 909-922.

Sprengelmeyer, R., J. D. Steele, B. Mwangi, P. Kumar, D. Christmas, M. Milders and K. Matthews (2011). "The insular cortex and the neuroanatomy of major depression." Journal of affective disorders **133**(1-2): 120-127.

Spronk, D., M. Arns, K. Barnett, N. Cooper and E. Gordon (2011). "An investigation of EEG, genetic and cognitive markers of treatment response to antidepressant medication in patients with major depressive disorder: a pilot study." Journal of affective disorders **128**(1-2): 41-48.

Stahl, S. M. (1998). "Mechanism of action of serotonin selective reuptake inhibitors: serotonin receptors and pathways mediate therapeutic effects and side effects." Journal of affective disorders **51**(3): 215-235.

Stahl, S. M., M. M. Grady, C. Moret and M. Briley (2005). "SNRIs: the pharmacology, clinical efficacy, and tolerability in comparison with other classes of antidepressants." CNS spectrums **10**(9): 732-747.

Stam, C., B. Jones, G. Nolte, M. Breakspear and P. Scheltens (2006). "Small-world networks and functional connectivity in Alzheimer's disease." Cerebral cortex **17**(1): 92-99.

Stam, C. J. (2005). "Nonlinear dynamical analysis of EEG and MEG: review of an emerging field." Clinical neurophysiology **116**(10): 2266-2301.

Stassen, H., J. Angst and A. Delini-Stula (1999). "Fluoxetine versus moclobemide: cross-comparison between the time courses of improvement." Pharmacopsychiatry **32**(02): 56-60.

Steele, J. D., J. Currie, S. M. Lawrie and I. Reid (2007). "Prefrontal cortical functional abnormality in major depressive disorder: a stereotactic meta-analysis." Journal of affective disorders **101**(1-3): 1-11.

Steriade, M. (2001). "Impact of network activities on neuronal properties in corticothalamic systems." Journal of neurophysiology **86**(1): 1-39.

Stimmel, G. L., J. A. Dopheide and S. M. Stahl (1997). "Mirtazapine: an antidepressant with noradrenergic and specific serotonergic effects." Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy **17**(1): 10-21.

Strik, W., T. Dierks, T. Becker and D. Lehmann (1995). "Larger topographical variance and decreased duration of brain electric microstates in depression." Journal of neural transmission **99**(1): 213-222.

Suffin, S. C. and W. H. Emory (1995). "Neurometric subgroups in attentional and affective disorders and their association with pharmacotherapeutic outcome." Clinical Electroencephalography **26**(2): 76-83.

Suhhova, A., M. Bachmann, K. Aadamsoo, Ü. Võhma, J. Lass and H. Hinrikus (2009). EEG Coherence as measure of depressive disorder. 4th European Conference of the International Federation for Medical and Biological Engineering, Springer.

Sullivan, E. A. and K. I. Shulman (1984). "Diet and monoamine oxidase inhibitors: a re-examination." The Canadian Journal of Psychiatry **29**(8): 707-711.

Sun, Y., F. Farzan, B. H. Mulsant, T. K. Rajji, P. B. Fitzgerald, M. S. Barr, J. Downar, W. Wong, D. M. Blumberger and Z. J. Daskalakis (2016). "Indicators for remission of suicidal ideation following magnetic seizure therapy in patients with treatment-resistant depression." JAMA psychiatry **73**(4): 337-345.

Surguladze, S., M. J. Brammer, P. Keedwell, V. Giampietro, A. W. Young, M. J. Travis, S. C. Williams and M. L. Phillips (2005). "A differential pattern of neural response toward sad versus happy facial expressions in major depressive disorder." Biological psychiatry **57**(3): 201-209.

Szegedi, A., W. T. Jansen, A. P. van Willigenburg, E. van der Meulen, H. H. Stassen and M. E. Thase (2009). "Early improvement in the first 2 weeks as a predictor of treatment outcome in patients with major depressive disorder: a meta-analysis including 6562 patients." The Journal of clinical psychiatry.

Tadel, F., S. Baillet, J. C. Mosher, D. Pantazis and R. M. Leahy (2011). "Brainstorm: a user-friendly application for MEG/EEG analysis." Computational intelligence and neuroscience **2011**: 8.

Takahashi, T., R. Y. Cho, T. Mizuno, M. Kikuchi, T. Murata, K. Takahashi and Y. Wada (2010). "Antipsychotics reverse abnormal EEG complexity in drug-naive schizophrenia: a multiscale entropy analysis." Neuroimage **51**(1): 173-182.

Tardito, D., J. Perez, E. Tiraboschi, L. Musazzi, G. Racagni and M. Popoli (2006). "Signaling pathways regulating gene expression, neuroplasticity, and neurotrophic mechanisms in the action of antidepressants: a critical overview." Pharmacological reviews **58**(1): 115-134.

Tarn, M., J. G. Edwards and E. Sedgwick (1993). "Fluoxetine, amitriptyline and the electroencephalogram." Journal of affective disorders **29**(1): 7-10.

Tenke, C. E., J. Kayser, C. G. Manna, S. Fekri, C. J. Kroppmann, J. D. Schaller, D. M. Alschuler, J. W. Stewart, P. J. McGrath and G. E. Bruder (2011). "Current source density measures of electroencephalographic alpha predict antidepressant treatment response." Biological psychiatry **70**(4): 388-394.

Thase, M. E., A. R. Entsuah and R. L. Rudolph (2001). "Remission rates during treatment with venlafaxine or selective serotonin reuptake inhibitors." The British journal of psychiatry **178**(3): 234-241.

Thase, M. E., J. B. Greenhouse, E. Frank, C. F. Reynolds, P. A. Pilkonis, K. Hurley, V. Grochocinski and D. J. Kupfer (1997). "Treatment of major depression with psychotherapy or psychotherapy-pharmacotherapy combinations." Archives of general psychiatry **54**(11): 1009-1015.

Thase, M. E. and A. J. Rush (1995). Treatment-resistant depression. Psychopharmacology: The Fourth Generation of Progress. F. E. Bloom and K. D. J. New York, Raven Press Ltd**:** 1081-1097.

Thomason, M. E., D. J. Yoo, G. H. Glover and I. H. Gotlib (2009). "BDNF genotype modulates resting functional connectivity in children." Frontiers in human neuroscience **3**: 55.

Thomasson, N., L. Pezard, J.-F. Allilaire, B. Renault and J. Martinerie (2000). "Nonlinear EEG changes associated with clinical improvement in depressed patients." Nonlinear Dynamics, Psychology, and Life Sciences **4**(3): 203-218.

Thut, G. and C. Miniussi (2009). "New insights into rhythmic brain activity from TMS–EEG studies." Trends in cognitive sciences **13**(4): 182-189.

Tibshirani, R. and G. Walther (2005). "Cluster validation by prediction strength." Journal of Computational and Graphical Statistics **14**(3): 511-528.

Tomescu, M. I., T. A. Rihs, R. Becker, J. Britz, A. Custo, F. Grouiller, M. Schneider, M. Debbané, S. Eliez and C. M. Michel (2014). "Deviant dynamics of EEG resting state pattern in 22q11. 2 deletion syndrome adolescents: A vulnerability marker of schizophrenia?" Schizophrenia research **157**(1): 175-181.

Tononi, G. and G. M. Edelman (1998). "Consciousness and complexity." Science **282**(5395): 1846-1851.

Tononi, G., O. Sporns and G. M. Edelman (1994). "A measure for brain complexity: relating functional segregation and integration in the nervous system." Proceedings of the National Academy of Sciences **91**(11): 5033-5037.

Tran, P. V., F. P. Bymaster, R. K. McNamara and W. Z. Potter (2003). "Dual monoamine modulation for improved treatment of major depressive disorder." Journal of clinical psychopharmacology **23**(1): 78-86.

Trivedi, M. H., A. J. Rush, S. R. Wisniewski, A. A. Nierenberg, D. Warden, L. Ritz, G. Norquist, R. H. Howland, B. Lebowitz and P. J. McGrath (2006). "Evaluation of outcomes with citalopram for depression using measurement-based care in STAR* D: implications for clinical practice." American journal of Psychiatry **163**(1): 28-40.

Uher, R., R. H. Perlis, A. Placentino, M. Z. Dernovšek, N. Henigsberg, O. Mors, W. Maier, P. McGuffin and A. Farmer (2012). "Self-report and clinician-rated measures of depression severity: can one replace the other?" Depression and anxiety **29**(12): 1043-1049.

Ulrich, G., H.-J. Haug and E. Fähndrich (1994). "Acute vs. chronic EEG effects in maprotiline-and in clomipramine-treated depressive inpatients and the prediction of therapeutic outcome." Journal of affective disorders **32**(3): 213-217.

Ulrich, G., H.-J. Haug, R.-D. Stieglitz and E. Fähndrich (1988). "Are there distinct biochemical subtypes of depression? EEG characteristics of clinically defined on-drug responders and non-responders." Journal of affective disorders **15**(2): 181-185.

Ulrich, G., H.-J. Haug, R.-D. Stieglitz and E. Fähndrich (1988). "EEG characteristics of clinically defined on-drug-responders and non-responders-a comparison clomipramine vs. maprotiline." Pharmacopsychiatry **21**(06): 367-368.

Ulrich, G., E. Renfordt and K. Frick (1986). "Electroencephalographic dynamics of vigilance and drug-response in endogenous depression. A comparison of patients treated with amitriptyline or pirlindole." Pharmacopsychiatry **19**(04): 270-271.

Ulrich, G., E. Renfordt and K. Frick (1986). "The topographical distribution of alpha-activity in the resting EEG of endogenous-depressive in-patients with and without clinical response to pharmacotherapy." Pharmacopsychiatry **19**(04): 272-273.

Ulrich, G., E. Renfordt, G. Zeller and K. Frick (1984). "Interrelation between changes in the EEG and psychopathology under pharmacotherapy for endogenous depression." Pharmacopsychiatry **17**(06): 178-183.

Üstün, T. B., J. L. Ayuso-Mateos, S. Chatterji, C. Mathers and C. J. Murray (2004). "Global burden of depressive disorders in the year 2000." The British journal of psychiatry **184**(5): 386-392.

Vakorin, V. A., S. Lippé and A. R. McIntosh (2011). "Variability of brain signals processed locally transforms into higher connectivity with brain development." Journal of neuroscience **31**(17): 6405-6413.

Vakorin, V. A., A. R. McIntosh, B. Misic, O. Krakovska, C. Poulsen, K. Martinu and T. Paus (2013). "Exploring age-related changes in dynamical non-stationarity in electroencephalographic signals during early adolescence." PLoS One **8**(3): e57217.

Van Diessen, E., T. Numan, E. Van Dellen, A. Van Der Kooi, M. Boersma, D. Hofman, R. Van Lutterveld, B. Van Dijk, E. van Straaten and A. Hillebrand (2015). "Opportunities and methodological challenges in EEG and MEG resting state functional brain network research." Clinical Neurophysiology **126**(8): 1468-1481.

Van Walraven, C. and R. G. Hart (2008). "Leave 'em alone–why continuous variables should be analyzed as such." Neuroepidemiology **30**(3): 138-139.

Varela, F., J.-P. Lachaux, E. Rodriguez and J. Martinerie (2001). "The brainweb: phase synchronization and large-scale integration." Nature reviews neuroscience **2**(4): 229.

Veer, I. M., C. Beckmann, M.-J. Van Tol, L. Ferrarini, J. Milles, D. Veltman, A. Aleman, M. A. Van Buchem, N. J. Van Der Wee and S. A. Rombouts (2010). "Whole brain resting-state analysis reveals decreased functional connectivity in major depression." Frontiers in systems neuroscience **4**: 41.

Verhagen, M., A. Van Der Meij, P. Van Deurzen, J. Janzing, A. Arias-Vasquez, J. Buitelaar and B. Franke (2010). "Meta-analysis of the BDNF Val66Met polymorphism in major depressive disorder: effects of gender and ethnicity." Molecular psychiatry **15**(3): 260.

von Knorring, L., C. Perris, L. Goldstein, D. Kemali, K. Monakhov and L. Vacca (1983). Intercorrelations between different computer-based measures of the EEG alpha amplitude and its variability over time

and their validity in differentiating healthy volunteers from depressed patients1. Neurophysiological Correlates of Normal Cognition and Psychopathology, Karger Publishers. **13:** 172-181.

Vos, T., C. Allen, M. Arora, R. M. Barber, Z. A. Bhutta, A. Brown, A. Carter, D. C. Casey, F. J. Charlson and A. Z. Chen (2016). "Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015." The Lancet **388**(10053): 1545-1602.

Vuilleumier, P. (2005). "How brains beware: neural mechanisms of emotional attention." Trends in cognitive sciences **9**(12): 585-594.

Wager, T. D. and C.-W. Woo (2017). "Imaging biomarkers and biotypes for depression." Nature medicine **23**(1): 16.

Wang, L., C. Zhu, Y. He, Y. Zang, Q. Cao, H. Zhang, Q. Zhong and Y. Wang (2009). "Altered small-world brain functional networks in children with attention-deficit/hyperactivity disorder." Human brain mapping **30**(2): 638-649.

Wang, X.-J. (2010). "Neurophysiological and computational principles of cortical rhythms in cognition." Physiological reviews **90**(3): 1195-1268.

Warden, D., M. H. Trivedi, S. R. Wisniewski, L. Davis, A. A. Nierenberg, B. N. Gaynes, S. Zisook, S. D. Hollon, G. Balasubramani and R. Howland (2007). "Predictors of attrition during initial (citalopram) treatment for depression: a STAR* D report." American Journal of Psychiatry **164**(8): 1189-1197.

West, B. J. and P. Grigolini (2010). "The living matter way to exchange information." Medical hypotheses **75**(6): 475-478.

Whitfield-Gabrieli, S. and J. M. Ford (2012). "Default mode network activity and connectivity in psychopathology." Annual review of clinical psychology **8**: 49-76.

Winokur, G. (1997). "All roads lead to depression: clinically homogeneous, etiologically heterogeneous." Journal of Affective Disorders **45**(1-2): 97-108.

World Health Organization (2009). "Global Health Risks: Mortality and Burden of Disease Attributable to Selected Major Risks." from http://www.who.int/healthinfo/global_burden_disease/GlobalHealthRisks_report_full.pdf.

Wu, J., M. S. Buchsbaum, J. C. Gillin, C. Tang, S. Cadwell, M. Wiegand, A. Najafi, E. Klein, K. Hazen and W. E. Bunney Jr (1999). "Prediction of antidepressant effects of sleep deprivation by metabolic rates in the ventral anterior cingulate and medial prefrontal cortex." American Journal of Psychiatry **156**(8): 1149-1158.

Wu, M., C. Andreescu, M. A. Butters, R. Tamburo, C. F. Reynolds III and H. Aizenstein (2011). "Default-mode network connectivity and white matter burden in late-life depression." Psychiatry Research: Neuroimaging **194**(1): 39-46.

Wu, Q. Z., D. M. Li, W. H. Kuang, T. J. Zhang, S. Lui, X. Q. Huang, R. C. Chan, G. J. Kemp and Q. Y. Gong (2011). "Abnormal regional spontaneous neural activity in treatment-refractory depression revealed by resting-state fMRI." Human brain mapping **32**(8): 1290-1299.

Xu, X., L. An, X. Mi and T. Zhang (2013). "Impairment of cognitive function and synaptic plasticity associated with alteration of information flow in theta and gamma oscillations in melamine-treated rats." PloS one **8**(10): e77796.

Yamamura, T., Y. Okamoto, G. Okada, Y. Takaishi, M. Takamura, A. Mantani, A. Kurata, Y. Otagaki, H. Yamashita and S. Yamawaki (2016). "Association of thalamic hyperactivity with treatment-resistant depression and poor response in early treatment for major depression: a resting-state fMRI study using fractional amplitude of low-frequency fluctuations." Translational psychiatry **6**(3): e754.

Yang, A. C. and S.-J. Tsai (2013). "Is mental illness complex? From behavior to brain." Progress in Neuro-Psychopharmacology and Biological Psychiatry **45**: 253-257.

Yatham, L. N., P. F. Liddle, R. W. Lam, A. P. Zis, A. J. Stoessl, V. Sossi, M. J. Adam and T. J. Ruth (2010). "Effect of electroconvulsive therapy on brain 5-HT 2 receptors in major depression." The British Journal of Psychiatry **196**(6): 474-479.

Youdim, M. B., D. Edmondson and K. F. Tipton (2006). "The therapeutic potential of monoamine oxidase inhibitors." Nature Reviews Neuroscience **7**(4): 295.

Yuan, H., V. Zotev, R. Phillips, W. C. Drevets and J. Bodurka (2012). "Spatiotemporal dynamics of the brain at rest—exploring EEG microstates as electrophysiological signatures of BOLD resting state networks." Neuroimage **60**(4): 2062-2072.

Zanto, T. P. and A. Gazzaley (2013). "Fronto-parietal network: flexible hub of cognitive control." Trends in cognitive sciences **17**(12): 602-603.

Zetterström, T., Q. Pei and D. Grahame-Smith (1998). "Repeated electroconvulsive shock extends the duration of enhanced gene expression for BDNF in rat brain compared with a single administration." Molecular brain research **57**(1): 106-110.

Zhang, J., J. Wang, Q. Wu, W. Kuang, X. Huang, Y. He and Q. Gong (2011). "Disrupted brain connectivity networks in drug-naive, first-episode major depressive disorder." Biological psychiatry **70**(4): 334-342.

Zhong, H., N. Haddjeri and C. Sánchez (2012). "Escitalopram, an antidepressant with an allosteric effect at the serotonin transporter—a review of current understanding of its mechanism of action." Psychopharmacology **219**(1): 1-13.

Zoon, H. F., C. Veth, M. Arns, W. Drinkenburg, W. Talloen, P. J. Peeters and J. Kenemans (2013). "EEG alpha power as an intermediate measure between brain-derived neurotrophic factor Val66Met and depression severity in patients with major depressive disorder." Journal of Clinical Neurophysiology **30**(3): 261-267.

Zou, Y.-F., D.-Q. Ye, X.-L. Feng, H. Su, F.-M. Pan and F.-F. Liao (2010). "Meta-analysis of BDNF Val66Met polymorphism association with treatment response in patients with major depressive disorder." European Neuropsychopharmacology **20**(8): 535-544.