EEG BIOMARKERS IN DEPRESSION

PAVING THE WAY FOR STRATIFIED PSYCHIATRY

NIKITA VAN DER VINNE

To my parents

The work described in this doctoral dissertation was carried out at Synaeda Psycho Medisch Centrum and Research Institute Brainclinics, in collaboration with University of Twente.

Vinne, Nikita van der
EEG biomarkers in depression
paving the way for stratified psychiatry

Published by Brainclinics Insights,
Nijmegen, The Netherlands

An electronic version of this dissertation is available at:
brainclinics.com

EEG BIOMARKERS IN DEPRESSION

PAVING THE WAY FOR STRATIFIED PSYCHIATRY

Dissertation

to obtain
the degree of doctor at the University of Twente,
on the authority of the rector magnificus,
prof. dr. T. T. M. Palstra,
on account of the decision of the graduation committee,
to be publicly defended
on Wednesday, the 1st of July 2020, at 16:45 hrs.

by

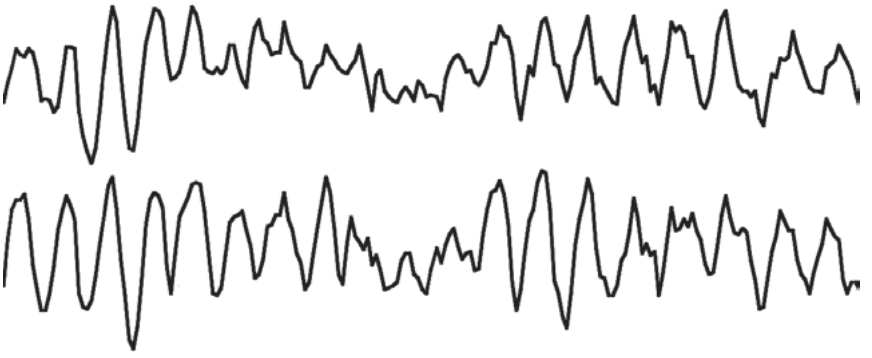
Nikita van der Vinne
born on the 20th of January 1987
in Deventer, The Netherlands

This dissertation has been approved by:

Supervisor
Prof. dr. ir. M. J. A. M. van Putten

Co-supervisors
Dr. M. Arns
Dr. M. A. Vollebregt

Cover and lay out: Brainclinics Insights, Nijmegen
Printed by GVO, Ede, The Netherlands
ISBN-978-90-830-0133-3



The Graduation Committee

Chairman and secretary
Prof. dr. J. L. Herek
University of Twente

Supervisor
Prof. dr. ir. M. J. A. M. van Putten
University of Twente

Co-supervisors
Dr. M. Arns
Research Institute Brainclinics
Dr. M. A. Vollebregt
Research Institute Brainclinics

Committee Members:
Prof. N. N. Boutros
Saint Luke's Marion Bloch Neuroscience Institute
Prof. G. A. van Wingen
University of Amsterdam
Prof. dr. ir. H. J. Hermens
University of Twente
Prof. dr. J. A. M. van der Palen
University of Twente
Dr. N. Bailey
Monash University

TABLE OF CONTENTS

1

GENERAL INTRODUCTION

1.1 MAJOR DEPRESSIVE DISORDER & TREATMENT

Throughout history, major depressive disorder has been thoroughly described, often known as melancholia (i.e. Hippocrates, 1923). Despite different terminology and the various explanations that were given by e.g. shamans, philosophers, and clergymen, old descriptions of the disease (Telles-Correia & Marques, 2015) show remarkable similarities to our 20th and 21st century classification for major depressive disorder (MDD; Diagnostic and Statistical Manual of Mental Disorders: DSM-5, 2013). Symptoms include a depressed mood most of the day (nearly every day), loss of interest in daily activities, diet independent weight loss or weight gain, fatigue, and feelings of worthlessness or guilt. This mood disorder is characterized by a long-lasting course, which in many cases turns into a chronic problem.

In the course of history, various treatments have been proposed, ranging from rituals in ancient Mesopotamia in the second millennium B.C. (Reynolds & Wilson, 2013), dietary restrictions, bloodletting and compound medicines in the seventeenth century (Burton, 1883), to cocaine and psychoanalysis around the turn of century in 1900 (Grinspoon & Bakalar, 1981). Yet, even in the current era of modern medicine, patients and doctors are still facing the challenges of combating this disease. The most common current treatments are antidepressant medication or psychotherapies like cognitive behavioral therapy (CBT).

Serotonin reuptake inhibitors (SSRIs, such as escitalopram and sertraline) are nowadays within the most commonly prescribed antidepressants (ADs). Although growing evidence indicates improvement in *mild* depression with ADs (Hieronymus et al., 2019; Stewart et al., 2012), ADs are usually seen as most effective only for people suffering from moderate to severe depression (Fournier et al., 2010). The presumed mechanism of action is the restriction of serotonin reuptake into the presynaptic cell.

Reuptake inhibitors of both serotonin and norepinephrine (SNRIs, such as venlafaxine) are also prescribed in large numbers and are presumed to limit the reuptake of the neurotransmitter norepinephrine in addition to serotonin. Theories on the working mechanisms are still not unequivocally proven, but these drugs serve many patients suffering from this disorder.

Psychotherapy (including CBT) is also a commonly applied treatment of MDD. The effectiveness of AD and psychotherapy treatment is similar (Cuijpers et al., 2008), but especially a combination of the two is considered to be superior to medication or CBT alone (Cuijpers et al., 2014; Cuijpers et al., 2009). For ADs alone, clinical efficacy ranges from 37% remission after a first antidepressant (AD) prescription to declining remission rates of respectively 31%, 14%, and 13% only, after each consecutive AD trial including augmentation strategies (Rush et al., 2006).

Disconcerting is the large (and continuously growing) prevalence of MDD in the world: while in 2005 183 million people were reported to be affected, this number increased to 216 million in 2015 (Vos et al., 2016), a large increase compared to the world population growth in this period (The World Bank, Population total). Publication bias and other biases plague the formation of scientific foundations for psychotherapy (Cuijpers et al., 2019) and the development of new antidepressant medication is subjected to suspended research and development budgets for central nervous system drugs, including ADs (Miller, 2010). New approaches to the treatment of MDD are needed to serve a growing group of people affected by this disorder.

1.2 BIOMARKERS

In a wide range of clinical applications, biomarkers become more and more important for providing health care that is adjusted to the level of the individual patient, otherwise known as personalized medicine. Biomarkers give us information on the presence of an illness or determining the right treatment. They give us the opportunity of doing measurements on the biological level, which opened up a new world of possibilities in hospital settings and the like (i.e. in digestive diseases, Carethers, et al., 2015, or in lung cancer, Rosell et al., 2013).

In the field of mental health, it is hard to find clinically accepted biomarkers for disorder detection and treatment. Decades of research has provided us with little solid evidence and replication. For MDD, several biomarkers derived from electrophysiological characteristics of the brain have been proposed, but their true effectiveness was recently questioned by Widge and colleagues (2019). Possibly contributing to the reservations about biomarkers' effectiveness, is that the predictive value of the most promising electrophysiological biomarkers has rarely (or possibly never) been tested prospectively.

1.3 ELECTROENCEPHALOGRAPHY

To establish electrophysiological characteristics of the brain that could be used as biomarkers, measurement of the electroencephalogram (EEG) is being used. In 1929, Hans Berger presented his observations of brain activity in the EEG to the world (Berger, 1929), being one of the first to study different patterns of human brain activity through this method. His son would become one of his most tested subjects, in whom Berger discovered the signifying alpha EEG rhythm. It appears when the eyes are being closed, and is attenuated by opening the eyes (see figure 1.1).

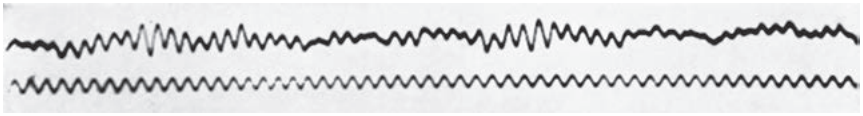


Figure 1.1: *A segment from the first reports of the human EEG by Hans Berger (1929). The figure represents a segment with what he called the 'alpha rhythm'. The tracing underneath is a generated 10 Hz sine wave. From: Berger (1929).*

The alpha rhythm studied by Berger consists of neural oscillations, i.e. rhythmic or repetitive neural activity in the central nervous system, between 8 and 12 Hz (voltage increases and declines 8 to 12 times per second). It reflects a state in which millions of cortical neurons oscillate synchronously with the same phase and within a comparatively narrow frequency range (Klimesch, 1999).

For thirty years, the EEG was mainly visually analyzed. This changed in the late 1960's, when digital equipment became available that allowed researchers to apply Fourier analysis to EEG data, to extract

frequency properties of a signal. Neural oscillations are traditionally being grouped into different frequency bands. This is useful when oscillations in a particular range of frequencies is the topic of interest – such as for example the alpha rhythm.

The synchronization of brain activity as measured by alpha waves, is foremost visible in the posterior part of the brain, in the occipital cortex. However, activity within this frequency range can be measured in other brain areas as well, and are presumably linked to several cognitive and mental processes. A vast amount of literature exists on alpha oscillations and what they represent. Alpha activity is traditionally seen as an idling rhythm (e.g. Pfurtscheller et al., 1996). A more active role of alpha oscillations, i.e. involvement in inhibition processes, dominates the current view on its role (Jensen et al., 2002; Klimesch et al., 2006; van Dijk et al., 2008). In this dissertation, the focus will be directed towards the role of alpha activity in MDD.

1.4 FRONTAL ALPHA ASYMMETRY AND DEPRESSION

In 1936, Lemere was the first to link low alpha activity to depression (1936). He inspected EEGs of healthy people and several psychiatric patients and concluded: *"...The ability to produce "good" alpha waves seems to be a neurophysiological characteristic which is related in some way to the affective capacity of the individual"* (Lemere, 1936, p. 374).

The presumed connection between alpha power and affect became more pronounced when asymmetry between hemispheres in the alpha frequency range was investigated in relation to depression (d'Elia & Perris, 1973). The objective in subsequent studies was to measure either more relatively right- or left-sidedness in alpha power, and relate this to depressive symptoms. By 1983, a group led by Richard Davidson started publishing on alpha asymmetry, focusing on frontal areas after they found no relation between symptoms and parietal alpha (Schaffer et al., 1983). Frontal alpha asymmetry (FAA) was calculated by subtracting alpha power of electrode F3 (which is located over the left dorsolateral prefrontal cortex, or DLPFC) from alpha power of electrode F4 (over the right DLPFC). The research group investigated affective capacities in terms of this FAA, and formulated

the diathesis model, a framework in which approach and avoidance form the basis (Davidson and Tomarken in 1989, described in Henriques & Davidson, 1991).

The diathesis model of approach and avoidance states that people showing a relative excess of alpha on the *left* side (compared to right) are characterized by more negative affect and withdrawal related emotion (see also Davidson, 1998). Conversely, people showing the opposite with a relative alpha excess on the *right* side, show more positive affect and approach behavior. When Henriques and Davidson (1991) found relatively more left-sided FAA in depressed patients and more right-sided FAA in healthy controls, they interpreted these findings as a deficit in the approach system in humans with left-sided FAA. These patients were said to be more prone to certain negative affect states and depressive disorders, given a certain level of environmental stress. This was endorsed by Harmon-Jones and Allen (Harmon-Jones & Allen, 1997), who concluded that FAA *"...may hold prognostic value for identifying those at risk for psychopathology characterized by a deficiency in approach motivation (e.g. depression)".*

During the late 1990's and 2000's, more studies on FAA and depression were published. Results appeared more ambiguous than expected, and some doubt arose whether FAA could actually reliably differentiate those at risk for depression from healthy humans. Instead of assuming homogeneity in this patient group, it seemed more heterogenic in terms of FAA. Rather embracing such heterogeneity, a *prognostic* effect of FAA was found (Arns et al., 2016): for females, right-sided FAA was associated with response and remission to the SSRIs escitalopram and sertraline, and left-sided FAA was associated with non-response and non-remission. No such effects were found for males.

1.5 EEG PAROXYSMS AND TREATMENT EFFECTS

In the clinic, EEGs are used for diagnostics in a large variety of central nervous system disorders, including epilepsy, coma, or assessment of sleep disorders. In these applications, brain activity is typically directly linked to observable human behavior, functioning or phenomena. In these circumstances, EEGs may show spike-wave

discharges, diffuse slowing or fragmentation of sleep cycles. Slowing of the EEG in the form of a slow background pattern can be captured by determining the alpha peak frequency (APF) or the dominant frequency (which has a broader frequency domain: 5-15 Hz compared to 7.5-13 Hz in the APF).

Only few studies explored how the EEG may relate to treatment effects in affective disorders. On this topic, Boutros and colleagues theorized *"...that milder degrees of increased neural excitability (i.e., a subthreshold excitation insufficient to cause seizures) may nonetheless be capable of causing observable phenotypic changes"* (Boutros et al., 2015). Examples are presented in figure 1.2. See chapter 4 for a more detailed description of different types of abnormalities and their associations to MDD and ADs. These subclinical abnormalities are reported in 3–5% of patients with MDD, a rate similar to controls (1–6%: Arns et al., 2017; Arns et al., 2008; Goodwin 1947; Lennox-Buchtal et al., 1960; Monin et al., 2018; Oh et al., 2018; Richter et al., 1971; Shelley et al., 2008). Arns and colleagues (2017) showed that a subgroup of MDD patients with abnormal EEG patterns was more likely to not respond to the ADs escitalopram and venlafaxine, whereas response to sertraline was not different for patients with or without EEG abnormalities. These findings suggest that patients with both MDD and abnormal EEGs may differentially respond to AD treatment.

1.6 COMPUTER ASSISTED EEG INTERPRETATION

Conventionally, visual inspection of the human EEG is used in several situations. A neurologist inspects the EEG by examining the background pattern, or searching for deviant, epileptiform activity. Various quantitative EEG analyses exist, ranging from elementary assessment of the posterior dominant rhythm with a Fourier transform transform (Lodder & van Putten, 2011), to the use of machine learning, including deep learning (Tjepkema-Cloostermans et al., 2018; Van Leeuwen et al., 2019; da Silva Lourenço et al., 2020). Algorithms based on machine learning could autonomously detect patterns that are invisible to the human eye, which would be suitable for the digitally recorded EEG. When providing sufficient data, deep learning is capable of learning a hierarchical feature representation automati-

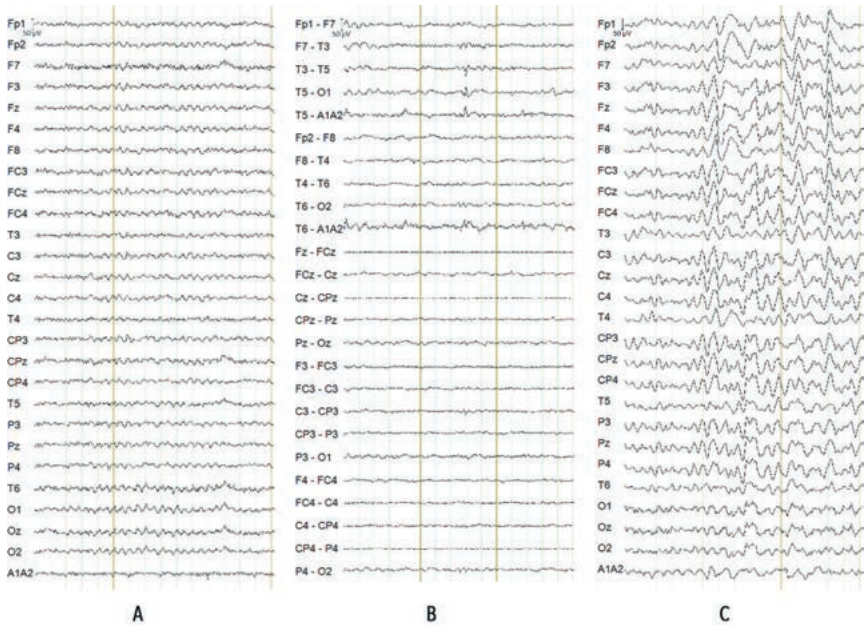


Figure 1.2: *Examples of EEG segments of 3 seconds, showing normal activity (A, montage: linked ears) and subclinically abnormal activity: sharp activity (B, montage: bipolar) diffuse slowing (C, montage: linked ears). Filter settings: 0.5-35 Hz.*

cally, it effectively learns by example (Tjepkema-Cloostermans et al., 2018). If a good deep learning paradigm is being developed, this would provide a more reliable alternative to the visual inspection of EEGs, because it does not suffer from the low interrater reliability of visual inspection: when classifying 300 EEGs as normal or containing seizures or epileptiform discharges, from a general clinically heterogeneous population, six board-certified neurophysiologists achieved an agreement (Fleiss's kappa) of 55% (Grant et al., 2014). Recent studies employing deep convolutional networks revealed remarkable distinctions between normal and EEGs with deviations. Studies focused on detecting interictal epileptiform discharges (IEDs) demonstrated substantial areas under the curve (AUC in ROC analyses) of 0.94 and 0.96 (respectively Tjepkema-Cloostermans et al. (2018) and da Silva Lourenço et al. (2020)). An AUC of 0.917 was achieved in a study focused on the distinction between normal and abnormal EEGs (Van Leeuwen et al., 2019).

1.7 EEG IN PSYCHIATRY

At present, no evidence-based clinical applications of EEG in psychiatry exist. Historically, this is interesting, as neurology and psychiatry do originate from one medical field in the 19th century (Baker et al., 2002). With discoveries in neuroanatomy and neuropathology at that time, the possibility of localizing diseases in the central nervous system shaped the discipline of neurology (Kazamel, 2018). However, the lack of a biological basis for mental disorders contributed to the expansion of psychoanalytical psychiatry, ultimately ending in the definitive split between neurology and psychiatry in the 20th century (Baker et al., 2002). As a result, EEGs are rarely investigated or deployed in psychiatry (Boutros, 2018). This, to the dissatisfaction of professionals who regard the division between neurology and psychiatry not a realistic one, and advocate a more integrated approach to diseases and disorders (Boutros, 2018; Reilly, 2015).

1.8 AIMS AND OUTLINE OF DISSERTATION

The primary aim of this doctoral dissertation was to investigate the value of neurophysiological biomarkers measured by the EEG, in the prognosis of treatment outcome in depression. Through this aim, we wished to add to integrating scientific approaches of psychiatry and neurology, transcending the currently conventional boundaries between medical fields, by utilizing neurophysiological methods in the treatment of mental disorders, in this specific case major depressive disorder, or MDD.

To investigate the value of EEG biomarkers for improvement of treatment, we asked ourselves the following questions:

1. Is the depressed population best characterized by homogeneity or heterogeneity?
2. What can we learn from state, trait, and drug effects on the biomarkers FAA and abnormalities in the EEG?
3. What can we learn about the underlying pathophysiology when studying the time course of biomarkers?
4. Can EEG biomarkers be reliably implemented for the treatment of depression?
5. Which circumstances are needed to provide patients and professionals with trustworthy advice?
6. Can we automate the detection of EEG abnormalities presumed relevant for psychiatry?

Chapter 2 focuses on frontal alpha asymmetry (FAA), and its presumed diagnostic properties. Contradictory study outcomes in the past motivated us to review these results and perform a meta-analysis. To explain discrepancies between studies, we employed post-hoc analyses for uncovering factors that possibly underlie these discrepancies. In **chapters 3 and 4**, we studied the time-course of two prognostic biomarkers (FAA and subclinical abnormalities in the EEG) over a period of eight weeks of antidepressant treatment. State, trait and drug effects were investigated, along with whether the time-course provides information about the etiology in earlier described subgroups. In an effort to automate the visual assessment of the proposed prognostic biomarkers (i.e., subclinical abnormal EEG activity) we investigated and combined several methods which are described in **chapter 5**. We employed frequency and spectral analysis, deep learning and random forest models. Eventually, in **chapter 6**, we report how we tested the feasibility and clinical effects of an EEG informed treatment allocation algorithm in a prospective feasibility study.
Findings are integrated in the general discussion, **chapter 7**.

2

FRONTAL ALPHA ASYMMETRY AS A DIAGNOSTIC MARKER IN DEPRESSION: FACT OR FICTION?

A META-ANALYSIS

N. van der Vinne, M. A. Vollebregt, M. J. A. M. van Putten and M. Arns
NeuroImage: Clinical, 2017; 16: 79–87

ABSTRACT

Introduction

Frontal alpha asymmetry (FAA) has frequently been reported as potential discriminator between depressed and healthy individuals, although contradicting results have been published. The aim of the current study was to provide an up-to-date meta-analysis on the diagnostic value of FAA in major depressive disorder (MDD) and to further investigate discrepancies in a large cross-sectional dataset.

Methods and materials

SCOPUS database was searched through February 2017. Studies were included if the article reported on both MDD and controls, provided an FAA measure involving EEG electrodes F3/F4, and provided data regarding potential covariates. Hedges' d was calculated from FAA means and standard deviations (SDs). Potential covariates, such as age and gender, were explored. Post hoc analysis was performed to elucidate interindividual differences that could explain interstudy discrepancies.

Results

Sixteen studies were included (MDD: n = 1883, controls: n = 2161). After resolving significant heterogeneity by excluding studies, a non-significant Grand Mean effect size (ES) was obtained (d = -0.007; CI = [-0.090] – [0.075]). Cross-sectional analyses showed a significant three-way interaction for Gender × Age × Depression severity in the depressed group, which was prospectively replicated in an independent sample.

Conclusions

The main result was a non-significant, negligible ES, demonstrating limited diagnostic value of FAA in MDD. The high degree of heterogeneity across studies indicates covariate influence, as was confirmed by cross-sectional analyses, suggesting future studies should address this Gender × Age × Depression severity interaction. Upcoming studies should focus more on prognostic and research domain usages of FAA rather than a pure diagnostic tool.

2.1 INTRODUCTION

With a lifetime prevalence of 16.2% in the United States, major depressive disorder (MDD) is a common disorder affecting many people (Kessler et al., 2003). Projections for 2030, reported by the WHO, show that MDD will become the second most debilitating disease worldwide (Mathers & Loncar, 2006). However, despite many pursuits of research groups into improving diagnostics and prognostics, MDD prevalence is still high (Patten et al., 2016). Improving differential diagnostic procedures should lead to a more reliable distinction between MDD and other mental disorders with overlapping symptoms, ultimately enabling better prognosis with more effective treatment.

Changes in affect, in particular a depressed mood, are one of the diagnostic criteria of MDD (*Diagnostic and Statistical Manual of Mental Disorders: DSM-5*, 2013). A model that focuses on affect, also known as the approach-withdrawal hypothesis, was developed to describe basic features of emotional affect (later described as the diathesis model by Davidson and Tomarken in 1989 [Henriques & Davidson, 1991]). According to this model, two major motivational systems in response to stimuli exist: one is appetitive whereas the other is aversive. This corresponds to positive and negative affect respectively, inducing approach or withdrawal behavior. The balance in the activation of these systems is also assumed to be reflected in differential activity in the EEG. In particular, anterior left activation (reflected by relatively diminished anterior left alpha activity, compared to right) was hypothesized to correspond with appetitive behavior (approach), and anterior right activation (reflected by relatively diminished anterior right alpha activity, compared to left) was hypothesized to correspond to aversive behavior (withdrawal; Davidson, 1984; Kelley et al., 2017). This asymmetry between left and right frontal alpha is referred to as frontal alpha asymmetry (FAA).
Initial EEG studies comparing depressed people with controls indeed

provided evidence for left-sided FAA (higher left than right frontal alpha activity) in depressed patients (Bell et al., 1998; Debener et al., 2000; Gotlib et al., 1998; Henriques & Davidson, 1991; Pizzagalli et al., 2002), compared to a dominant right-sided FAA in controls (Fingelkurts et al., 2006; Schaffer et al., 1983). Note that left-sided FAA is inversely related to relatively greater right than left cortical activity, as cortical processing typically results in a reduction of synchronous rhythmic activity (e.g. a reduction in alpha power). A significant correlation between FAA and Behavioral Activation System (BAS) sensitivity (of which low scores indicate a predisposition toward certain types of MDD), suggested that this pattern of FAA *"...may hold prognostic value for identifying those at risk for psychopathology characterized by a deficiency in approach motivation* (*e.g. depression)"* (Harmon-Jones & Allen, 1997). Furthermore, left-sided FAA is hypothesized to specifically expose subgroups reporting anhedonia (a common MDD symptom described as diminished interest or experience of pleasure), while anxious apprehension, related to an opposite pattern of right-sided FAA, might possibly mark another subgroup (Nusslock et al., 2015). Defining such subgroups needs further investigation.

Although recent studies have confirmed an association between MDD and FAA (Beeney et al., 2014; Gollan et al., 2014; Jaworska et al., 2012; Kemp et al., 2010), which was also reflected by two reviews (Baskaran et al., 2012; Fingelkurts & Fingelkurts, 2015), multiple methodologically sound studies have failed to confirm the diagnostic value of FAA regarding MDD and other mental illnesses (Allen et al., 2004; Carvalho et al., 2011; Deldin & Chiu, 2005; Gold et al, 2013; Kaiser et al., 2016; Kentgen et al., 2000; Knott et al., 2001; Mathersul et al., 2008; Price et al., 2008; Quraan et al., 2014; Reid et al., 1998), including the largest EEG study to date in MDD in a sample of 1008 MDD patients compared to 336 controls (Arns et al., 2016) from our research group. Questions should be raised on the uniformity and generalizability of all studies regarding FAA. This concerns technical properties of the EEG recordings and further processing of the data, as well as sample characteristics. This makes updating previous reviews and a meta-analysis relevant, with adding results of more recent studies.

A decade ago, Thibodeau, Jorgensen and Kim (2006) addressed the use of FAA in a meta-analytic review, including a maximum of 1614 adults (depressed and healthy, exact sample size is unknown), and concluded that depression is meaningfully related to relatively greater right than left frontal cortical activity at rest (left-sided FAA), with moderate weighted mean effect sizes (ES) for the depressed adults with Pearson r = 0.26 and Cohen's d = 0.54. Their meta-analysis did not include recent large methodologically sound studies and had several limitations, e.g. it included a wide range of groups defined by other characteristics than an MDD diagnosis or defined as sub-clinical MDD characteristics, FAA measures based on different scalp sites, and different types of ESs as reported in original articles. When controlling for sub-clinical MDD, the authors found an equally moderate ES for FAA with r =0.27, indicating a limited influence of operationalization of depression. Several studies and reviews (Davidson, 1998; Hagemann et al., 1998; Jaworska et al., 2012; Segrave et al., 2011; Smith et al., 2017; Stewart et al., 2010; Thibodeau et al., 2006) have indicated that methodological aspects could explain discrepant findings, such as the EEG reference montage and frequency range considered. Further, the FAA calculation is not often discussed in studies and reviews, but varies in normalization application. Normalizing by dividing $F4$ - $F3$ by its sum ($F4$ + $F3$) enables researchers to rule out interindividual EEG differences like individual EEG power (as a result of skull thickness for instance).

The purpose of the current study was to provide an up-to-date meta-analysis, further clarifying the role of FAA in MDD using a standardized approach. This is achieved by calculating a weighted mean effect size (ES) only based on original means and standard deviations (SDs), obtained from EEG electrode F3 and F4 only and using a more homogenous sample with clear inclusion criteria (MDD vs non-MDD only, excluding subclinical samples). Furthermore, we also used data from a large cross-sectional dataset (MDD: n = 938, Controls: n = 306) to investigate interindividual differences, and the impact of methodological aspects such as EEG montaging and use of normalization.

2.2 METHODS AND MATERIALS

A literature search was carried out in SCOPUS for the period up until February 2017, using the query "depression AND EEG OR electroencephalogram AND alpha asymmetry", which yielded 172 hits. The database search outlined above was supplemented by manual searches. To identify additional publications, we further inspected reference lists from prior meta-analyses (Thibodeau et al., 2006) and reviews (Fingelkurts & Fingelkurts, 2015; Jesulola et al., 2015). PRISMA guidelines for conducting and reporting systematic reviews were followed during this analysis (Moher et al., 2009).

Studies had to meet the following inclusion criteria: (a) DSM-IV diagnosis of MDD or MDD classification after a structured clinical interview using the SCID or MINI; (b) availability of mean, standard deviation (SD), and sample size of resting FAA (electrode F4 minus F3); (c) availability of a healthy control group; (d) reporting of EEG reference montage; (e) published in English. When means, SDs and/or sample sizes were not provided in the article, authors were e-mailed to request the relevant data. Additional subject information on the following variables was gathered: mean age and SD, comorbid classifications (% and type of comorbidity), comorbid anxiety, medication status (% receiving an antidepressant), gender (% female), depression severity mean and SD. For each study, we also recorded the year of publication, reference montage, resting EEG condition (eyes open (EO), eyes closed (EC), or both), recording length, alpha bandwidth and continent where the study is carried out.

Statistical analyses were performed using SPSS 17.0. MetaWin 2.1 (Rosenberg et al., 2000) was used to conduct the meta-analysis and generate all variables of interest. ESs (the standardized mean difference Hedges' *d*) were calculated based on the FAA statistic from the MDD group and control group means and SDs. This ES is a scale-free statistic, thereby allowing comparison of scores from various studies. A grand mean ES was calculated with a 95% confidence interval (CI) providing the weighted mean ES for all studies. Larger ES values indicate stronger clinical relevance. Furthermore, Qt (heterogeneity of ESs), and the failsafe number (Rosenthal's method: $\alpha < 0.05$, and Orwin's method) were calculated. The fail-safe number is the number

of studies indicating how many unpublished null findings are needed to render an effect non-significant. When the total heterogeneity of a sample (Qt) was significant – indicating that the variance among ESs is greater than expected by sampling error – the study contributing most to the significance of the Qt value was excluded from further analysis for that variable until the Qt value was no longer significant. This was done for a maximum of three iterations. If more than three studies needed to be excluded to obtain a non-significant Qt value, then other explanatory variables for the effects had to be assumed (Rosenberg et al., 2000) and were investigated in post hoc tests.

To investigate specific interstudy differences (or a lack thereof), the cross-sectional dataset of Arns et al. (2016) was used to elucidate interindividual differences that could drive differences between studies. To this end, main and interactional effects of group, gender, age, depression severity (HRSD-17), and anxiety severity (HAM-A), were investigated through univariate ANCOVAs. To test the stability of the significant results in this paper across EEG reference montages and different FAA definitions, FAA was also analyzed after re-referencing to Cz and the linked ears from the original average reference montage.

2.3 RESULTS

2.3.1 META-ANALYSIS

A total of 214 studies were identified between January 1998 and July 2016. One additional relevant study was identified out of studies covered by an earlier meta-analysis (Thibodeau et al., 2006) and reviews (Baskaran et al., 2012; Fingelkurts & Fingelkurts, 2015; Jesulola et al., 2015). A final search conducted in February 2017 yielded eight new hits, resulting in one extra study in the meta-analysis. See figure 2.1 on page 28 for a flow diagram of the inclusion process.

Most excluded studies were not selected due to the absence of a control group (n = 58) or the absence of a clinical MDD group (n = 47). Sixteen studies (Arns et al., 2016; Baehr et al., 1998; Beeney et al., 2014; Brzezicka et al., 2016; Cantisani et al., 2015; Carvalho et al., 2011; Deslandes et al., 2008; Gollan et al., 2014; Gordon et al., 2010;

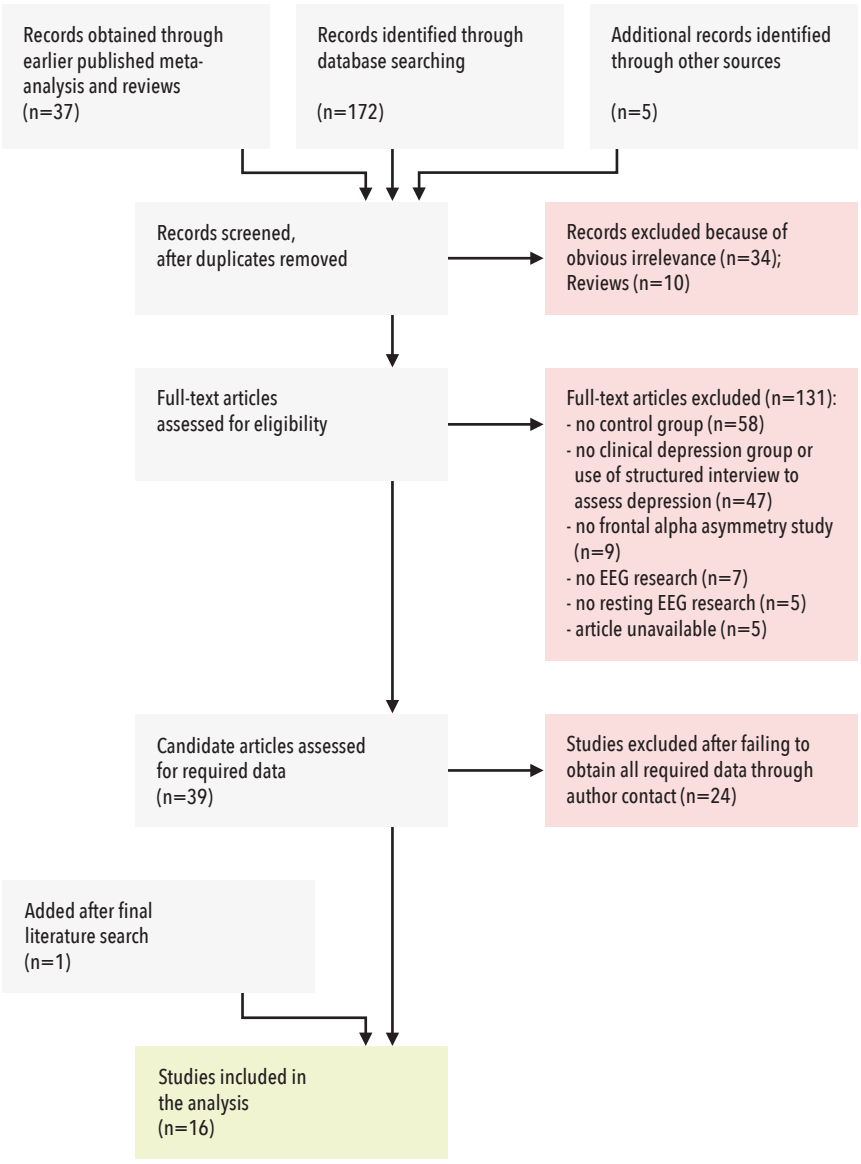


Figure 2.1: *Flow diagram of the inclusion process.*

Jaworska et al., 2012; Kaiser et al., 2016; Keeser et al., 2013; Liu et al., 2016; Quinn et al., 2014; Saletu et al., 1996; Segrave et al., 2011; Stewart et al., 2010) met all inclusion criteria and were included in this meta-analysis, see table 2.1 on pages 30 and 31 for an overview. Note

that only one study was included in both the previous meta-analysis by Thibodeau et al. (2006) and the current meta-analysis, because most of the other previous studies were either based on a depression group defined by solely a severity measure (no official diagnosis), on continuous depression severity measures (no control group), or the requested means were not available.

Due to overlapping samples of Gordon et al. (2010) and Quinn et al. (2014), original data of these studies were requested and combined to prevent overlapping samples, now referred to as Gordon/Quinn. Note that Gordon et al. originally reported on electrode FC4 and FC3 and Quinn et al. on the more frequently used F4 and F3. Considering the inclusion criteria of this meta-analysis, only *F*4 - *F*3 data were merged (this data was provided by BRAINnet). FAA Means, SDs and n were recalculated for the studies of Arns et al. (2016; analysis of original data for EC only as well as age means and SDs), Stewart et al. (2010; merging of subgroup data), and Brzezicka et al. (2016; merging of individual data). Additional statistics (means of (*F*4 - *F*3)/(*F*4 + *F*3) and SDs) and subject data were calculated for the data provided by Daniel Keeser of the neurophysiological research group of the Ludwig-Maximilians-University of Munich, Germany (Keeser et al., 2013). These data were updated with data from newly included subjects since the publishing of the cited conference abstract.
A total of 1883 MDD subjects and 2161 control subjects was included in the meta-analysis. A fixed-effects model meta-analysis yielded a significant heterogeneity test ($Qt = 37.65$, $p = 0.001$), a non-significant grand mean ES of –0.041 (*CI* = [-0.1204-0.0375]), and a fail-safe number of 11.4 (Rosenthal's method) and 0 (Orwin's method). The forest plot in figure 2.2, shown on page 32, and funnel plot in figure 2.3 (page 33) show a graphical overview of the ESs and grand mean.

Table 2.1 Overview of all included studies in the meta-analysis, covering the period 1996-2017

			FAA MDD Group			FAA Controls Group			EEG details				
N°	Study	ES	Mean[1]	SD	n	Mean[1]	SD	n	Reference montage	EO/EC[2]	Recording length (s)	Alpha band (Hz)	FAA measure
1	Arns, 2015	0.059	0.002	0.13	938	-0.005	0.10	306	CA[4]	EC	120	8-13	(F4-F3)/(F4+F3)
2	Baehr, 1998	-1.345	-1.160	7.82	13	9.760	7.86	11	Cz	EC	300	8-13	(F4-F3)/(F4+F3)
3	Beeney, 2014	-0.261	0.020	0.36	13	0.090	0.18	21	Cz	EO/EC	480	8-13	F4-F3
4	Brzezicka, 2016	-0.254	0.799	1.36	26	1.109	1.01	26	CSD[4]	EC	300	8-13	F4-F3
5	Cantisani, 2015	-0.834	-0.127	0.34	20	0.102	0.17	19	CA	EC	300	8-12.5	F4-F3
6	Carvalho, 2011	0.533	0.059	0.11	12	0.003	0.08	7	LEa	EC	480	8-12.9	F4-F3
7	Deslandes, 2008	-0.283	-0.179	0.28	22	-0.105	0.21	14	LE[4]	EC	480	8-13	F4-F3
8	Gollan, 2014	0.480	0.430	0.73	37	0.160	0.27	35	LE	EO/EC	480	8-13	F4-F3
9	Gordon/Quinn, 2010/2014	-0.072	-0.010	0.15	93	-0.002	0.11	1037	CA	EC	120	8-13	F4-F3
10	Jaworska, 2012	-0.166	-0.005	0.16	53	0.034	0.30	43	CA,Cz,LE	EC	360	8-13	F4-F3
11	Kaiser, 2016	-0.006	0.212	0.20	14	0.213	0.12	14	LE	EC	180	8.9-10.9	F4-F3
12	Liu, 2016	-0.307	0.000	0.06	141	0.018	0.05	113	LE	EO/EC	360	7.8-12.7	F4-F3
13	Keeser, 2013	0.005	-0.016	0.09	233	-0.016	0.07	291	CA	EC	600	8-12	(F4-F3)/(F4+F3)
14	Saletu, 1996	-0.600	-0.014	0.07	60	0.031	0.08	29	CA	EC[3]	180	7.5-13	(F4-F3)/(F4+F3)
15	Segrave, 2011	0.443	0.065	0.10	16	0.010	0.14	18	CA,Cz	EO/EC	360	8-13	F4-F3
16	Stewart, 2010	0.066	0.013	0.08	143	0.007	0.09	163	CA,Cz,LE,CSD	EO/EC	480	8-13	F4-F3

[1]In the occurrence of multiple reference montages, one is selected for calculation of the grand mean in the following order of priority: CA, Cz, Mas, CSD; [2]When both EO and EC data was available, EC was used for further analysis; [3]An auditory stimulus was presented when a drowsiness pattern was visible in the EEG; [4]Abbreviations used: CA=common average reference, CSD=current source density, LE=Linked ears

Table 2.1, *continued*

N°	Age mean (SD) MDD	Age mean (SD) Controls	Comorbidity	% Female	% Medicated (MDD)	MDD severity measure	Severity mean (SD) MDD	Severity mean (SD) Controls	Continent
1	37.6 (12.6)	36.9 (13.1)	Anx[5] (n=62), Soc phob[5] (n=105)	57%	0%	HRSD	22 (4.1)	1.2 (1.7)	International
2	43.5 (7)	44.2 (13.3)	Unknown	n/a	17%	BDI	21.9 (9.2)	3.4 (2.9)	North-America
3	32.1 (8.8)	27.8 (11.7)	Anx (n=4), PTSD (n=1)	100%	n/a	BDI-II	18.5 (9.4)	2.9 (2.8)	North-America
4	28 (8.3)	24.9 (5.2)	Unknown	60%	n/a	BDI	20.5 (8.3)	2.9 (2.5)	Europe
5	43.4 (14)	41.1 (13.8)	Unknown	54%	95%	HRSD	25.5 (5)	n/a	Europe
6	71 (7.8)	72 (9.2)	None	63%	100%	BDI-II	16.4 (4.4)	2 (2.3)	South-America
7	71.6 (1.2)	72.4 (1.7)	None	94%	100%	HRSD[6]	9.4 (1.5)	1.1 (2.6)	South-America
8	36.2 (12.4)	35.1 (13.7)	Unknown	63%	0%	IDS-C[6]	33.7 (7.7)	2.3 (2.6)	North-America
9	40.7 (14.8)	40.2 (17.1)	None	50%	0%	DASS	13.8 (5.1)	1.5 (2.2)	International
10	40.7 (11.9)	36.6 (9.9)	Anx (n=8)	54%	0%	MDD: HRSD[6] Control: BDI-II	22.4 (5.1)	4.4 (5)	North-America
11	80.5 (5.7)	80.9 (7.0)	Unknown (no Anx)	100%	Unknown	HADS-D	7.7 (2.5)	2.0 (1.4)	Europe
12	33.1 (12.1)	32.6 (12.5)	Anx (n=74)	82%	37%	HRSD	26.3 (7.8)	3.2 (4.9)	North-America
13	22.3 (14.3)	46.3 (14.2)	Adj[5] (n=3), Sub[5] (n=3), Anx (n=1), Depe[5] (n=1), OCD (n=1)	56%	63%	n/a	n/a	n/a	Europe
14	51.1 (3.1)	53.4 (2.9)	Unknown (n=2)	100%	0%	HRSD	18.3 (5.7)	2.9 (2.4)	Europe
15	40.8 (11.4)	42.1 (13)	None	100%	44%	BDI-II[6]	39.3 (10.6)	2.1 (2.5)	Australia
16	19.1 (0.1)	Unknown	None	69%	0%	HRSD[6]	11.1 (1.1)	4 (0.6)	North-America

[5] Abbreviations used: Anx = anxiety, Soc phob=social phobia, Adj = adjustment disorder, Sub=substance abuse, Depe = dependent personality disorder
[6]Multiple depression severity measures available, one is selected in the following order of priority: HRSD/HAM-D, BDI-II, MADRS, IDS-C

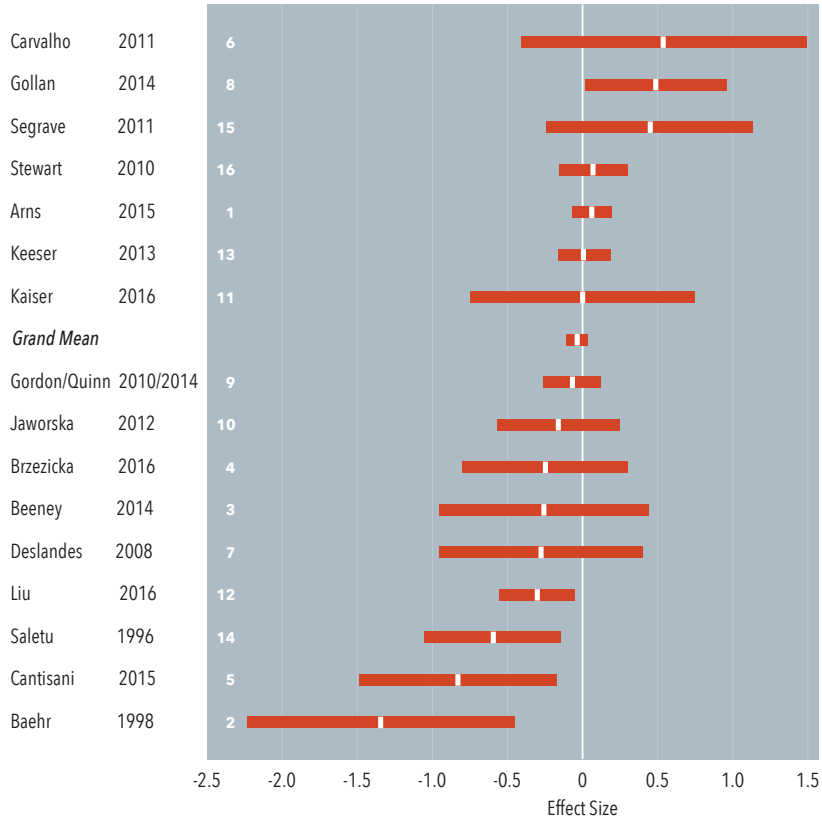


Figure 2.2: *Forest plot of the effect sizes (ES) of all included studies and the grand mean ES for all studies. The grand mean ES after resolving heterogeneity was –0.007 (not significant). Numbers correspond to study numbers in table 2.1.*

Exclusion of three studies (Baehr et al., 1998; Cantisani et al., 2015; Saletu et al., 1996) abolished the significant heterogeneity, resulting in a non-significant grand mean ES of –0.007 (*CI* = [-0.090-0.075]) and fail-safe numbers of 0 (Rosenthal's method) and 0 (Orwin's method). In subsequent post hoc analysis, we attempted to identify the source of heterogeneity (outlined below).

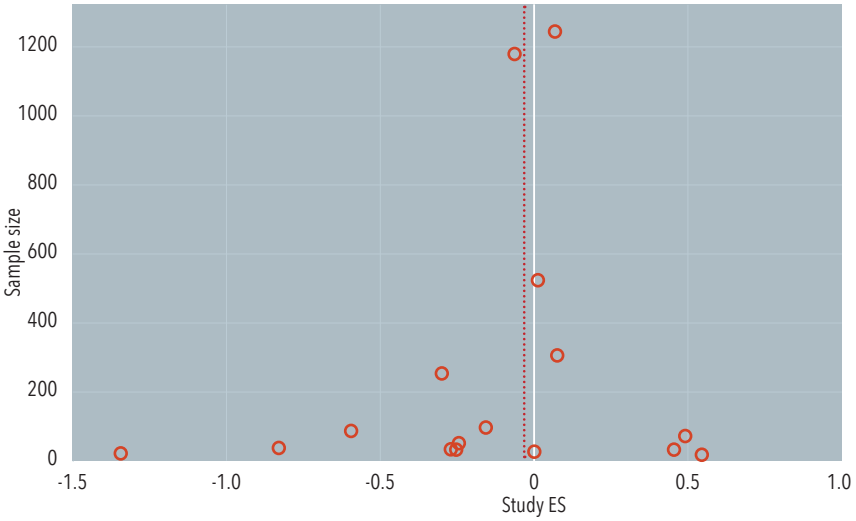


Figure 2.3: *Funnel plot of the study effect sizes (ES) and corresponding sample sizes, with the white line indicating x= 0 and the dotted line indicating the grand mean ES = -0.041. Note that the largest studies with sample size N>200 all approach the same ES close to 0, suggesting that a sample size of 300 and larger is required to obtain stable and biologically plausible effects for FAA in MDD.*

2.3.2 POST HOC TESTS

Post hoc, the influence of several potential moderators was investigated. Detailed results can be found in supplement 2.5. One potentially important moderator is the choice of the reference montage, which differs across the included studies. We performed post hoc tests where the relationship between study ES and reference montage was investigated. This did not result in significant ESs, or left the analyses with an insufficient number of studies, and therefore insufficient power, to achieve reliable results. Additional analyses (with combined montages as well as separated analyses per type of montage) between study ES and most potential moderators demonstrated no significant correlations, including anxiety. This was investigated further in one of the included studies by Arns et al. (2016), who found no changes in results after excluding subjects diagnosed with comorbid anxiety (female responders showed greater alpha (less cortical activity) over the right frontal site, whereas non-remitters showed the opposite asymmetry).

2.3.3 CROSS-SECTIONAL ANALYSIS

To explain the different study outcomes, we used 1244 participants (out of 1344 subjects, these had successful EEGs combined with BDI scores) from the cross-sectional dataset iSPOT-D (Arns et al., 2016) to extract candidate factors that could influence FAA or explain differences between studies. No significant contribution of singular variables to FAA was found through univariate ANCOVA (variables included group and gender as fixed factors, and age, depression severity, and anxiety as covariates). Although no significant interaction effect was found for Group × Gender × Age × Severity, within the depressed group, a significant three-way interaction of Gender × Age × Severity was found ($F(1,930) = 6.096$, $p = 0.014$) when these variables were exclusively part of the model, but this was not the case within the control group. Replacing depression severity with anxiety severity in this model did not yield any significant effects. To study the stability of Gender × Age × Severity across datasets, the same analysis was prospectively conducted in the Gordon/Quinn dataset (MDD: $n = 93$, controls: $n = 1037$). Note that one-sided testing of the replication of an interaction effect was not possible through ANOVA. Nevertheless, considering the a priori hypothesis and the smaller replication sample at hand, a more liberal criterion of $p < 0.10$ was employed. This replicated the significant three-way interaction effect as well ($F(1,85) = 3.400$, $p = 0.069$).

To visualize this three-way interaction, the Curve Fitting Toolbox in MATLAB 2016b (The Mathworks, Inc., Natick, MA) was used. In figure 2.4, the linear fitting of the surface is illustrated, comparing females and males based on their FAA, age and depression severity. A pattern becomes visible where differences between females and males seem to exist for older and severely depressed subjects, especially from an age of approximately 53 years and older, with opposing effects for males compared to females. Based on these results, four groups were formed dividing young and old (<53 and ≥53 years old), and moderately and severely depressed subjects (HDRS score <24 and ≥24, based on recent labelling of HDRS depression scores (Zimmerman et al. , 2013)). In these groups, univariate ANOVAs with gender, age, and severity as dependent variables were performed separately. No significant gender effects in FAA were found in both the young and moderately depressed groups. In the old, severely depressed group however, females had significantly

higher, i.e. right-sided, FAA than males ($F(1,46) = 8.094$, $p = 0.007$). This seems to drive the three-way interaction effect found earlier.

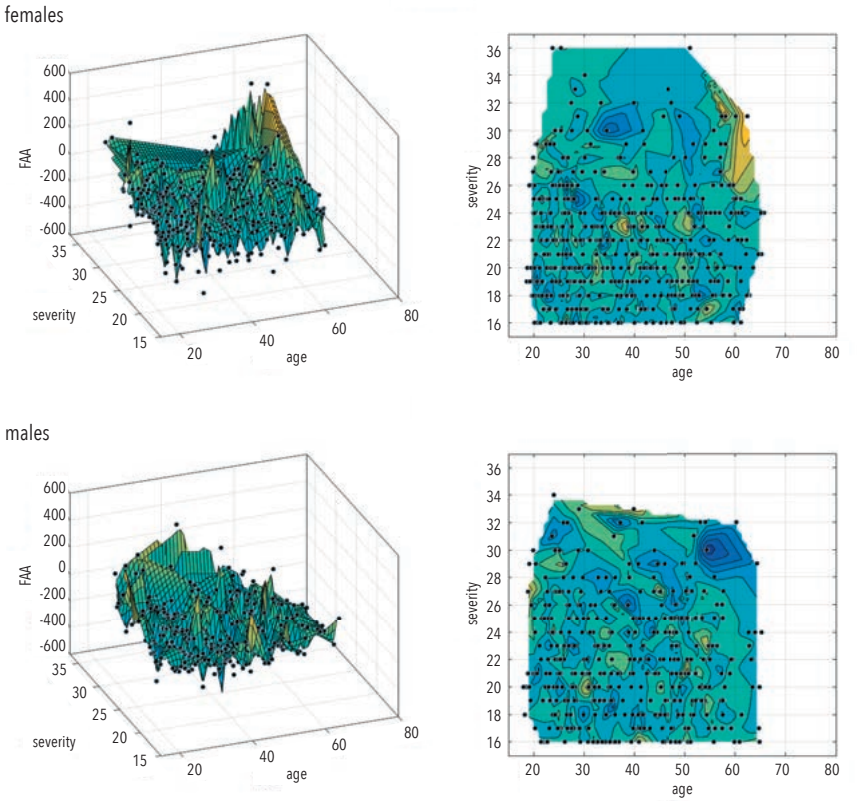


Figure 2.4: *Linear fitted surface graph, visualizing the three-way interaction effect of frontal alpha asymmetry (FAA), age, and depression severity, separately for females and males.*

The four defined groups were subsequently used to address the original question: Can a diagnosis of MDD be predicted using FAA? Univariate ANOVAs including only severely depressed, separately for males and females, and younger and older subjects (split up at 53 years) showed strikingly different results. While no differences between controls and depressed were found for the younger groups ($n = 243$ and $n = 288$ respectively), significant differences were found in the older groups with a severe depression, both for males and females (respectively $F(1,34) = 4.806$, $p = 0.035$, Cohen's $d = 0.71$ and $F(1,59) = 0.6791$, $p = 0.012$, Cohen's $d = -0.69$). Figure 2.5 illustrates that the direction of this effect is reversed for males and females, with relatively more left-sided FAA in de-

pressed males, and more right-sided FAA for females. Repeating these ANOVAs by replacing severely depressed with moderately depressed yielded no significant effects.
Comparing the different reference montages in the cross-sectional dataset through multivariate ANOVA did not result in significant group differences on FAA in either montage (see figure 2.6A), nor did stratification by gender, suggesting that the lack of group effects cannot be simply explained by the EEG-montage used.

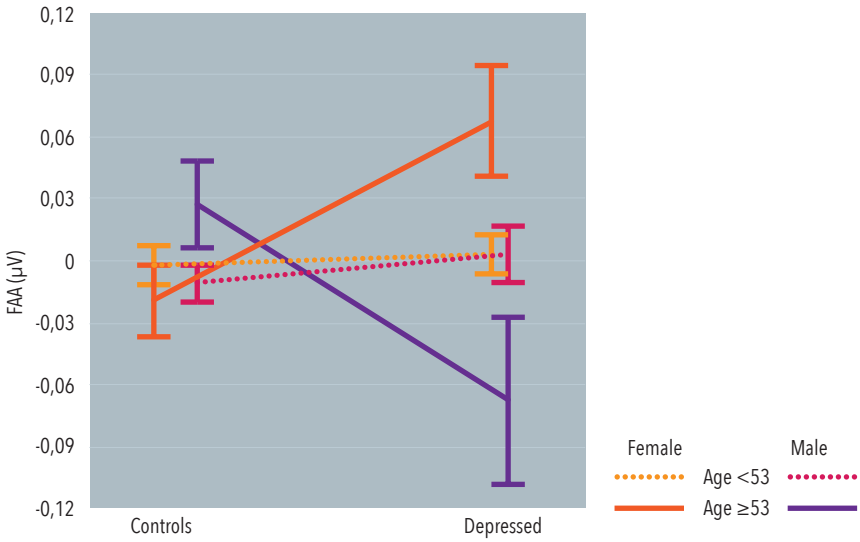


Figure 2.5: *Line graphs with error bars (representing standard error of the means) depicting the difference in frontal alpha asymmetry (FAA) between controls and severely depressed patients, separately for males and females, and <53 years and ≥53 years. Positive values of FAA indicate greater alpha over right than left frontal site, negative values indicate the opposite.*

To normalize interindividual differences FAA F4 minus F3 can be divided by its sum. However, most included studies calculated FAA only by the difference score *F4 - F3* (see table 2.1). Although two multivariate ANOVAs comparing the different methods did not yield different results of FAA in depressed and controls, the absence of sum division can result in rather large differences in raw individual FAA scores, depending on which reference scheme is applied (see figure 2.6b). However, not dividing by the sum still did not render the non-significant group effect to significance.

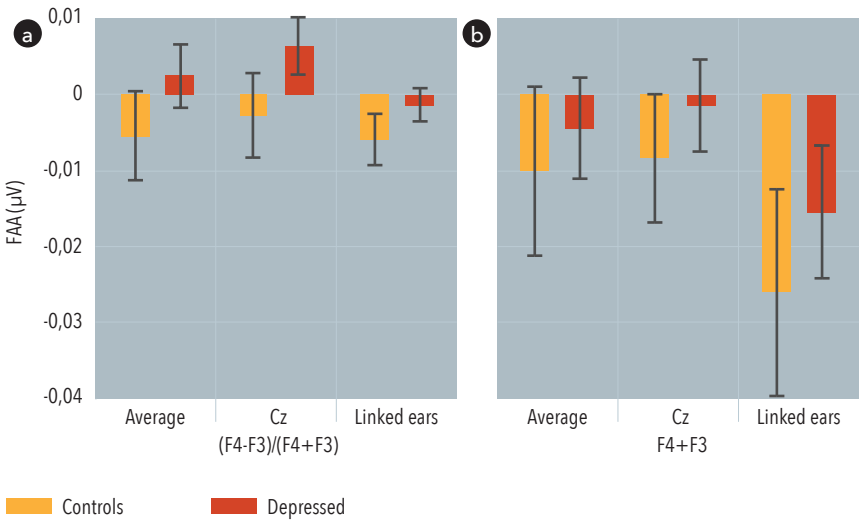


Figure 2.6a: *Illustration of frontal alpha asymmetry (FAA) means and 95% error bars (representing standard error of the mean) for controls and depressed patients, separately for the three different EEG reference schemes.* ***b:*** *Similar to* ***a****, except for the calculation of FAA without dividing by the sum of F4 and F3. Note the differences based on EEG reference montage, but also that none of these methodological changes changed the overall MDD-control contrast to a significant difference, illustrating that these methodological aspects could yield different outcomes, but do not explain the lack of 'diagnostic' effect of FAA in this large sample.*

2.4 DISCUSSION

In this meta-analysis, the diagnostic value of FAA was investigated. The small and non-significant effect size approaching zero extracted from this meta-analysis, accompanied by highly significant heterogeneity across studies, suggest that FAA is not a reliable diagnostic biomarker for MDD. Furthermore, the funnel plot in figure 2.3 suggests that at least 300 subjects need to be included to obtain a stable and biological plausible effect for FAA in MDD, confirming that most studies have been underpowered that investigated the diagnostic value of FAA (cf. table 2.1).

We could not identify a single variable that reliably explained a significant portion of the variance in FAA findings across studies. Cross-sectional analyses in the large iSPOT-D sample (Arns et al., 2016) were performed to explore possible candidate variables that have been suggested to explain differences between studies, e.g. EEG reference

montage, calculation of FAA with or without normalization, effects and interactions of gender, depression severity, anxiety severity, etc. A significant interaction effect of age, gender, and depression severity was found in depressed patients as visualized in figures 2.4 and 2.5, and also prospectively replicated in an independent sample (Gordon et al., 2010; Quinn et al., 2014). This interaction implicated more right-sided FAA (relatively more cortical activity on the left than right frontal site) in severely depressed women aged 53 years and older, in contrast to relatively more left-sided FAA in severely depressed men of the same age. This finding suggests that when unequal gender distributions, age ranges, and depression severity are studied, this may result in non-generalizable results. This confirms our hypothesis that a high level of heterogeneity in FAA in the depression population exists, which is in line with previous methodologically sound studies (Deldin & Chiu, 2005; Kentgen et al., 2000; Knott et al., 2001; Price et al., 2008; Quraan et al., 2014). Consequently, the lack of consistency in the results is not in line with the approach-withdrawal model, which was hypothesized to predict a meaningful relationship between the degree of approach behavior and affect on one hand, and FAA on the other hand. Note that Davidson (1998) emphasized that his previously developed model of approach and withdrawal systems *"...was never intended as a model of depression or any other form of psychopathology for that matter"*. Differences in frontal asymmetry may thus reflect individual differences in affective style rather than being a pure diagnostic marker for MDD (Davidson, 1998), thereby warranting its use more along the lines of Research Domain Criteria (RDoC) or Precision Medicine (Cuthbert, 2014). This is in line with the clear prognostic role of FAA, where a right frontal dominant FAA was associated with response to SSRIs and left frontal dominant FAA was associated with nonresponse to SSRIs in females (Arns et al., 2016). Translating this knowledge to prognostic methods in clinical practice, will allow health care professionals to personalize mental health treatments.

To our knowledge, our finding, comprising three different factors (gender, age, depression severity) has not been reported before. Previous studies have not always included all three variables, or sample sizes might have been too small to detect this three-way interaction. Interestingly, a closely related interaction effect between gender and

severity was recently reported by Jesulola et al. (2017), reflecting an FAA pattern in severely depressed females, that is opposite to the traditionally hypothesized direction of FAA in MDD, which is lacking in males. Note that age was not taken into account here. Previous studies in elderly showed no group differences in FAA (Kaiser et al., 2016), even when controlling for depression severity (Carvalho et al., 2011; Deslandes et al., 2008). On the one hand, age effects are not ruled out because increased neural heterogeneity in older adults has been found (Karch et al., 2015). Albeit, a large dataset of 6029 subjects showed that FAA does not change across the lifespan in a healthy population (Hashemi et al., 2016). Furthermore, significant differences between healthy and depressed individuals were reported in younger samples from this meta-analysis (mean sample ages of 29.4 and 35.7; in Beeney et al., 2014, and Gollan et al., 2014), but most likely these studies were underpowered (see funnel plot in figure 2.3). Therefore, the literature regarding more left-sided alpha in young and middle-aged depressed cannot be explained by current results. Other explanatory variables must be assumed, as the high level of heterogeneity suggests (Rosenberg et al., 2000).

Although the most frequently used EEG montage in the included studies is the average reference, other montages like Cz, and linked ears referencing are also common practice. Davidson (1998) and Hagemann et al. (1998) both made a strong case for enabling more consistent study outcomes by using average reference in FAA research. A promising reference-free methodology in FAA research is current source analysis (e.g. Brzezicka et al., 2016). In particular Current Source Density (CSD) is recommended for advanced EEG analysis by both Kayser and Tenke (2015) and Stewart et al. (2014), avoiding the question how to reference data and by providing a more distinct topography. As the current meta-analysis contains no comparable CSD studies, we can only recommend the use of the average reference, based on our post hoc analyses. Not only its ability to correct for strong occipital alpha, but also the topographical proximity of Cz to F3 and F4, and the possible insensitivity to subtle but meaningful differences of the linked ears reference, make the average reference the best candidate. An additional advantage is its relative insensitivity for the choice whether or not the FAA difference score ($F4 - F3$) is divided by its sum ($F4 + F3$), as visualized in figure 2.6. This choice has

considerably more effect when applied to linked ears referenced data. Although the relative difference in FAA between depressed patients and controls is similar in any combination of reference scheme and FAA measure, we recommend the use of ($F4$ - $F3$)/($F4$ + $F3$). Not dividing by its sum has large consequences for the degree of negativity of FAA in both groups.

A strong element in this meta-analysis was the calculation of an ES based on each study's FAA means and SDs, as well as the application of clear inclusion criteria improving the consistency across studies. Unfortunately, this resulted in considerably fewer included studies than Thibodeau et al. (2006), albeit this meta-analysis included a substantially larger overall sample size (k = 16 vs k = 24 and n = 4044 vs n = 1614 for our meta-analysis and the meta-analysis by Thibodeau respectively). Furthermore, a consequence of excluding most previously included studies by Thibodeau and colleagues, the current is a completely new meta-analysis with respect to the entered datasets (apart from one study), instead of an extended meta-analytic database. This might have caused a difference in findings. Nonetheless, we consider the consistency across studies superior to the quantity of studies. The inclusion of only sixteen studies did make it difficult to compare studies based on several characteristics, regularly leaving us with groups too small to come to reliable conclusions. In part, this was overcome by performing cross-sectional analyses on the largest dataset in this meta-analysis (Arns et al., 2016) and cross-validation in a second dataset (Gordon/Quinn; Gordon et al., 2010, and Quinn et al., 2014). This enabled us to unravel patterns that would not have become visible in a meta-analysis only, making the value of gender, age and MDD severity in relation to FAA evident, and need to be taken into account in future studies investigating FAA in MDD. The current data did not allow for identifying additional subgroups showing symptoms such as anhedonia, comorbid anxious apprehension, panic or social phobia, but further studying contribution of these specific clusters of symptoms to FAA, could benefit the personalization of mental health treatments. Furthermore, a few state emotion manipulations in EEG paradigms show greater FAA group differences than resting state EEG paradigms (Stewart et al., 2014). Although the number of these studies is too small to include in the current study, this method could enlarge the chance of determining subgroups. Finally,

reliability and consistency of measuring FAA might be improved by EEG recording across multiple sessions, as FAA is found to be moderately stable across time (Allen et al., 2004; Vuga et al., 2006), as originally suggested by Davidson (1998). For a detailed and recent overview of studies on hemispheric asymmetry in depression, please see the review by Bruder et al. (2017).

The importance of replication of results has become increasingly evident, because many scientific claims in psychology and psychiatry are rebutted, or more intricate systems appear to be implicated. New insights suggest a different application of FAA, actually utilizing the interindividual variation in this biomarker. For instance, the prediction of antidepressant treatment outcome using gender specific alpha asymmetry was first reported by Bruder et al. (2001) and replicated by Arns et al. (2016). Future studies into the use of FAA as a biomarker could help improve understanding of the basic dimensions underlying human behavior, and ultimately lead to improving treatment. This being one of the purposes of the use of Research Domain Criteria (RDoC), future studies should be in line with this approach, in order to demonstrate the clinical relevance of FAA more as a domain criterion or prognostic biomarker, rather than a 'diagnostic' marker. We emphasize that individual differences should not be ignored, but rather embraced, thereby potentially leading to optimized characterization of relevant subgroups and subsequent implications for a personalized treatment for the increasing number of depressed patients.

ACKNOWLEDGEMENTS

Preliminary results of this meta-analysis were presented at the 19th Biennial Conference of the International Pharmaco-EEG Society, October 2016. We acknowledge the support of Michelle Wang, Chris Spooner and Donna Palmer in obtaining all relevant data of the cross-sectional datasets, Ryan Thibodeau in further clarifying the first meta-analysis, as well as Aneta Brzezicka and Olga Kamińska, Andrea Cantisani, Stewart Shankman, Jennifer Stewart and John Allen, Andréa Deslandes and Alessandro Carvalho, Natalia Jaworska, Andreas Kaiser, Daniel Keeser, Joseph Beeney, and Peter Anderer for providing additional data requested.

2.5 SUPPLEMENT

SUPPLEMENTARY INFORMATION ON THE POST HOC TESTING OF POTENTIAL MODERATORS OF FAA IN MDD

Since several included singular studies reported on multiple reference montages, we performed analyses on both "unique" data and "non-unique" data. For example: for the initial grand mean, average referenced data from Stewart and colleagues (2) was used. Their Cz data however, were also used for analysis of the Cz reference montage "non-unique" subgroup. When the heterogeneity test $\chi 2$ was significant, analyses where repeated without the study contributing most to heterogeneity (indicated with "excluding one study" column 2 in table S2.1). Calculating grand means per reference montage resulted in non-significant heterogeneity tests (see Prob ($\chi 2$) in table S2.1) for all montages (after excluding one study for the average and Cz montages that contributed most to the significance). First analyses show that corresponding ESs were negligible (max. -0.07). Only for linked ears referenced data a small effect size of -0.23 with non-significant heterogeneity was found, after excluding one study contributing most to the significance. Corresponding fail-safe analyses (both Rosenthal's and Orwin's methods) showed that it would take only a few null-findings (up to 6.5) to reduce ES to irrelevant.
Additional analyses (with combined montages as well as separate analyses per type of montage) showed no significant correlations between study ES and average age, gender distribution, sample size, MDD severity (ESs where calculated due to different depression severity measures), % medicated, % of comorbidity, % of comorbid anxiety, EEG recording length, mean alpha frequency, and width of the alpha frequency band. Handedness was not investigated, because either the subjects were all right-handed, or no information on this variable was available. One-way ANOVAs showed no significant differences in study ES between resting EEG conditions (EO/EC) or continent of residence (both including and excluding international studies). A significant correlation was found for year of publication (r = 0.502, p = .048), although data was not collected continuously through the years. This effect was mainly driven by the two oldest studies, which when removed deemed this effect reversed and non-significant.

Table S2.1 ESs per reference montage group and corresponding statistics as provided by MetaWin 2.1 (1)
Analyses were first performed on "unique" studies (such as average referenced data from Stewart et al (2)), including studies only once across the reference montages. However, to utilize all available data, "non-unique" studies were included in separate non-unique analyses (such as Cz and linked ears referenced data by Stewart et al (2)). Based on these findings, aiming for non-significant heterogeneity, one study contributing most to the significance was excluded.

Uniqueness	Reference montage	Selection studies	K studies	Qtotal	df	Prob (X^2)[1]	Mean ES	95% CI[2]	Fail safe analysis Rosenthal's	Fail safe analysis Orwin's	Included studies[3]
Unique studies	Average reference[4]	All	8	16,873	7	.018*	-0.012	-0.109-[0.083]	0	0	Arns, Cantisani, Gordon/Quinn, Jaworska, Keeser, Saletu, Segrave, Stewart (2-10)
		Excluding one study	7	10,164	6	.118	0.007	-0.095-[0.108]	0	0	Arns, Cantisani, Gordon/Quinn, Jaworska, Keeser, Segrave, Stewart (2-8,10)
	Cz	All	2[5]								Baehr, Beeney (11,12)
	Linked ears	All	5	10,586	4	.031*	-0.109	-0.387-[0.170]	0	0	Carvalho, Deslandes, Gollan, Kaiser, Liu (13-17)
		Excluding one study	4	3,228	3	.358	-0.235	-0.304-[0.208]	0	0	Carvalho, Deslandes, Kaiser, Liu (13,14,16,17)
Non-unique studies	Cz	All	5	11,536	4	.021*	-0.118	-0.371-[0.135]	6.5	0	Baehr, Beeney, Jaworska, Segrave, Stewart (2,7,10-12)
		Excluding one study	4	3,903	3	.272	-0.067	-0.362-[0.229]	0	0	Beeney, Jaworska, Segrave, Stewart (2,7,10,12)
	Linked ears	All	7	11,940	6	.063	-0.033	-0.206-[0.140]	0	0	Carvalho, Deslandes, Gollan, Kaiser, Liu, Jaworska, Stewart (2,7,13-17)

** Significant at the p=.05 level; [1] A significant X^2 indicates a level of heterogeneity that is too large to be able to interpret the corresponding ES; [2] CI=confidence interval [3] Note that numbers do not correspond to table 1 of the manuscript; [4] Because the average reference was first priority in our analyses, this group did not contain a "unique" and "non-unique" subgroup; [5] Because of a limited number of available studies, this analysis was not performed*

3

STABILITY OF FRONTAL ALPHA ASYMMETRY IN DEPRESSED PATIENTS DURING ANTIDEPRESSANT TREATMENT

N. van der Vinne, M. A. Vollebregt, M. J. A. M. van Putten and M. Arns
NeuroImage: Clinical, 2019; 102056

ABSTRACT

Introduction

Frontal alpha asymmetry (FAA) is a proposed prognostic biomarker in major depressive disorder (MDD), conventionally acquired with electroencephalography (EEG). Although small studies attributed trait-like properties to FAA, a larger sample is needed to reliably asses this characteristic. Furthermore, to use FAA to predict treatment response, determining its stability, including the potential dependency on depressive state or medication, is essential.

Methods and materials

In the international Study to Predict Optimized Treatment in Depression (iSPOT-D), a multi-center, randomized, prospective open-label trial, 1008 MDD participants were randomized to treatment with escitalopram, sertraline or venlafaxine-extended release. Treatment response was established eight weeks after treatment initiation and resting state EEG was measured both at baseline and after eight weeks (n = 453).

Results

FAA did not change significantly after eight weeks of treatment (n = 453, p = .234), nor did we find associations with age, sex, depression severity, or change in depression severity. After randomizing females to escitalopram or sertraline, for whom treatment response could be predicted in an earlier study, FAA after eight weeks resulted in equivalent response prediction as baseline FAA (one tailed p = .028).

Conclusion

We demonstrate that FAA is a stable trait, robust to time, state and pharmacological status. This confirms FAA stability. Furthermore, as prediction of treatment response is irrespective of moment of measurement and use of medication, FAA can be used as a state-invariant prognostic biomarker with promise to optimize MDD treatments.

3.1 INTRODUCTION

Frontal alpha asymmetry (FAA) is a proposed biomarker conventionally acquired with electroencephalography (EEG). FAA has been studied for over three decades in major depressive disorder (MDD), anxiety, and other psychiatric diseases. Several studies stated, in a traditional framework of FAA, that it reflects the approach-withdrawal motivation system, i.e. the diathesis model (Davidson, 1984; Harmon-Jones & Allen, 1997; Henriques & Davidson, 1991; Kelley et al., 2017). Left-sided FAA (i.e. more right-sided frontal cortical activation than left-sided) was correlated more to withdrawal behavior than to approach, which was in turn associated with a vulnerability to developing MDD. However, our meta-analysis showed that FAA cannot be used as a generic diagnostic biomarker in MDD and does not reliably differentiate MDD from non-MDD patients (Van der Vinne et al., 2017), providing evidence against the diathesis model. Only a small subgroup of severely depressed females over 53 years of age showed more right-sided alpha activity and severely depressed males over 53 years of age more left-sided alpha than control peers.

When regarding FAA as a *prognostic* rather than *diagnostic* biomarker, alpha asymmetry may be more promising. Bruder and colleagues (2008) found SSRIs (selective serotonin reuptake inhibitors) treatment responders to have more right-sided alpha asymmetry while non-responders showed opposite asymmetry, primarily over the occipital region. This was confirmed in the large international Study for Predicting Optimized Treatment – Depression sample, where specifically female SSRI responders had more right-sided FAA, and non-responders the opposite (iSPOT-D, Arns et al., 2016). To further assess properties of FAA as a prognostic biomarker, knowledge on its reliability, stability, and sensitivity to other factors, such as medication or severity of depression, needs to be established.

A predominant view in affective neuroscience is that FAA in depressed patients consists of mostly *trait*-like features, not changing over time with *state* and independent of interventions, although some studies have suggested otherwise: both longitudinal and cross-sectional designs have been used to test FAA stability (see table 3.1 on page 50 for a summary, and supplement 3A for a table with detailed overview of studies). With an exception of Debener et al. (2000), most studies report FAA to be stable with minor or no changes between baseline and assessment later, both in patients and healthy controls (Allen et al., 2004; Bruder et al., 2008; Davidson et al., 2003; Deldin & Chiu, 2005; Gollan et al., 2014; Keune et al., 2011; Spronk et al., 2008; Sutton & Davidson, 1997; Tomarken et al., 1992).

Cross-sectionally, several studies showed that FAA is independent of depression severity, both between patients (Allen et al., 2004; Arns et al., 2016; Feldmann et al., 2018; Gollan et al., 2014; Nusslock et al., 2018; Van der Vinne et al., 2017; Vuga et al., 2006) and within patients, including remission (Carvalho et al., 2011). This contrasts the findings by Grünewald et al. (2018) and Keune et al. (2011), where a higher level of depression complaints correlated with more left-sided FAA (albeit only in the control group of Grünewald et al.). In other cross-sectional studies on FAA stability between depressed patients and patients remitted from depression, no differences were found (Carvalho et al., 2011; Feldmann et al., 2018; Gotlib et al., 1998).

Despite some inconclusive results, the majority of findings indicate that FAA is predominantly a trait, only partially or not affected by changes in depressive state. Our meta-analysis on FAA as a diagnostic marker of depression (Van der Vinne et al., 2017) demonstrated that bias is strongly reduced from 300 cases onwards. Studies investigating FAA stability until now always studied smaller samples ($n \leq 85$). This may explain part of the conflicting results on FAA in these studies.

This has motivated our current work that aims to replicate longitudinal results on the temporal stability of FAA by using data from the iSPOT-D dataset (baseline $n = 1008$, week-8 $n = 453$). The primary hypothesis was that FAA is reliable, and remains stable over

time, with limited changes as a result of antidepressant treatment, time and state change. We therefore assessed FAA after eight weeks of antidepressant drugs and consequential state changes in mood. As age, sex, and depression severity have had a significant influence on FAA-related outcomes in iSPOT-D and other studies (e.g. Arns et al., 2016; Bruder et al., 2001; Stewart et al., 2010; Van der Vinne et al., 2017), we extended analyses by investigating possible mediation of FAA by these variables. We specifically studied MDD patients versus healthy controls differentiating subgroups identified in our previous meta-analysis, i.e. severely depressed patients over 53 years old (Van der Vinne et al., 2017). As in earlier iSPOT-D reports on FAA anxiety was not found to be of influence, we did not add this variable to our analyses.

For clinical use of FAA as a biomarker for treatment response, it is relevant to assess stability and robustness to medication. Stability is particularly an advantage when patients are already on an AD preceding baseline (that often have long half-life times requiring washout periods of weeks) and FAA remains unaffected. We therefore also assess outcome prediction with FAA recorded after eight weeks treatment. In our previous report (Arns et al., 2016), at baseline, right-sided FAA in females was associated with favorable outcome to the SSRIs escitalopram and sertraline, whereas left-sided FAA was not. If FAA *is* prognostic for AD treatment outcome in specific subsamples, and FAA is indeed a stable *trait*, FAA after eight weeks on an AD should still be able to predict treatment outcome for females in agreement with our previous study (Arns et al., 2016). We hypothesized that analysis of week-8 medicated EEG data would result in the same treatment prediction results as baseline unmedicated data did.

Table 3.1 Summary of studies on state/trait properties of frontal alpha asymmetry

Study	Study type[1]	Mostly trait	Subjects	EEG methods	Intervention
Allen et al., 2004	1	•	MDD, female	3 to 5 Ax., 8 or 16 weeks apart	Acupuncture
Bruder et al., 2008	1	•	MDD and HC	2 Ax., 12 weeks apart	Fluoxetine treatment
Debener et al., 2000	1		MDD and HC	2 Ax., 2-4 weeks apart	Several antidepressants
Deldin & Chiu, 2005	1	•	MDD and HC	4 Ax. On 1 day	Cognitive restructuring
Gollan et al., 2014	1	•	MDD and HC	2 Ax., 16 weeks apart	Behavioral activation
Keune et al., 2011	1	•	MDD	2 Ax., 8 weeks apart	Mindfulness
Spronk et al., 2008	1	•	MDD	2 Ax., pre/post-treatment	rTMS
Vuga et al., 2006	1	•	Childhood onset MDD and HC	2 Ax., 1-3.2 years apart	Some patients on ADs (n = 13 of 49)
Davidson et al., 2003	2	•[2]	HC	3 Ax., 8 weeks, 4 months	Mindfulness meditation
Hagemann, et al., 2002	2	•	HC	4 Ax., all 4 weeks apart	None
Hagemann, et al., 2005[3]	2	•	HC	3 Ax., all 5 weeks apart	None
Sutton & Davidson, 1997	2	•	HC	2 Ax., 6 weeks apart	None
Tenke et al., 2017[3]	2	•	HC	2 Ax., 5-16 days apart	None
Tomarken et al., 1992	2	•	HC	2 Ax., 3 weeks apart	None
Carvalho et al., 2011	3	•[2]	MDD, remitted, and HC	1 Ax.	None
Feldmann et al., 2018	3	•[2]	MDD, remitted, and HC	1 Ax.	None
Gotlib et al., 1998	3	•	MDD, remitted, and HC	1 Ax.	None
Grünewald et al., 2018	3	•[2]	MDD and HC	1 Ax.	None
Nusslock et al., 2018	3	•	MDD and HC	1 Ax.	None

[1] Type 1: Multiple assessment moments with depressed patients. Type 2: Multiple assessment moments, only healthy controls. Type 3: Cross-sectional study; [2] No explicit statements on state or trait were made by the authors (on electrode F3/F4 or F7/F8 based FAA), based on other literature we suggest our own conclusion to these results; MDD = major depressive disorder, HC = healthy controls, Ax. = assessment(s); [3] not trait, or mostly state

3.2 METHODS AND MATERIALS

3.2.1 DESIGN

This is an international multi-center, randomized, prospective open-label trial (Phase-IV clinical trial) in which MDD patients were randomized to escitalopram, sertraline, or venlafaxine-XR treatment in a 1:1:1 ratio. The study protocol details, including a power calculation, have been published by Williams et al. (2011). This design was deliberately chosen to mimic real-world practice with the aim of optimizing the translatability to real world settings.

3.2.2 MDD PATIENTS AND TREATMENT

We included 1008 MDD patients, recruited between October 2008 and January 2011. A detailed description of the study assessments, inclusion/exclusion criteria, diagnostic procedures and treatment is available in Williams et al. (2011). In summary, the primary diagnosis of nonpsychotic MDD was confirmed before randomization using the Mini-International Neuropsychiatric Interview (MINI-Plus, Sheehan et al., 1998), according to DSM-IV criteria, and a score ≥16 on the 17-item Hamilton Rating Scale for Depression ($HRSD_{17}$). Additional measuring of depression complaints was done with the Very Quick Inventory of Depressive Symptomatology – Self Report ($VQIDS\text{-}SR_5$, De La Garza et al., 2017). Comorbid anxiety disorders were allowed (present in 6.2% [specific phobia] to 10.5% [social phobia] of patients). All patients were either medication-naive or, if previously prescribed an antidepressant medication, had undergone a washout period of at least five half-lives before the baseline visit clinical and EEG assessments. After the baseline visit, patients were randomized to one of three antidepressant medication treatments. After eight weeks of treatment, patients were tested again using the HRSD17, the VQIDS-SR5 and an EEG assessment (figure 3.1, page 52). This study was approved by the institutional review boards at all of the participating sites and this trial was registered with ClinicalTrials.gov. Registration number: NCT00693849; URL: http://clinicaltrials.gov/ct2/show/NCT00693849.

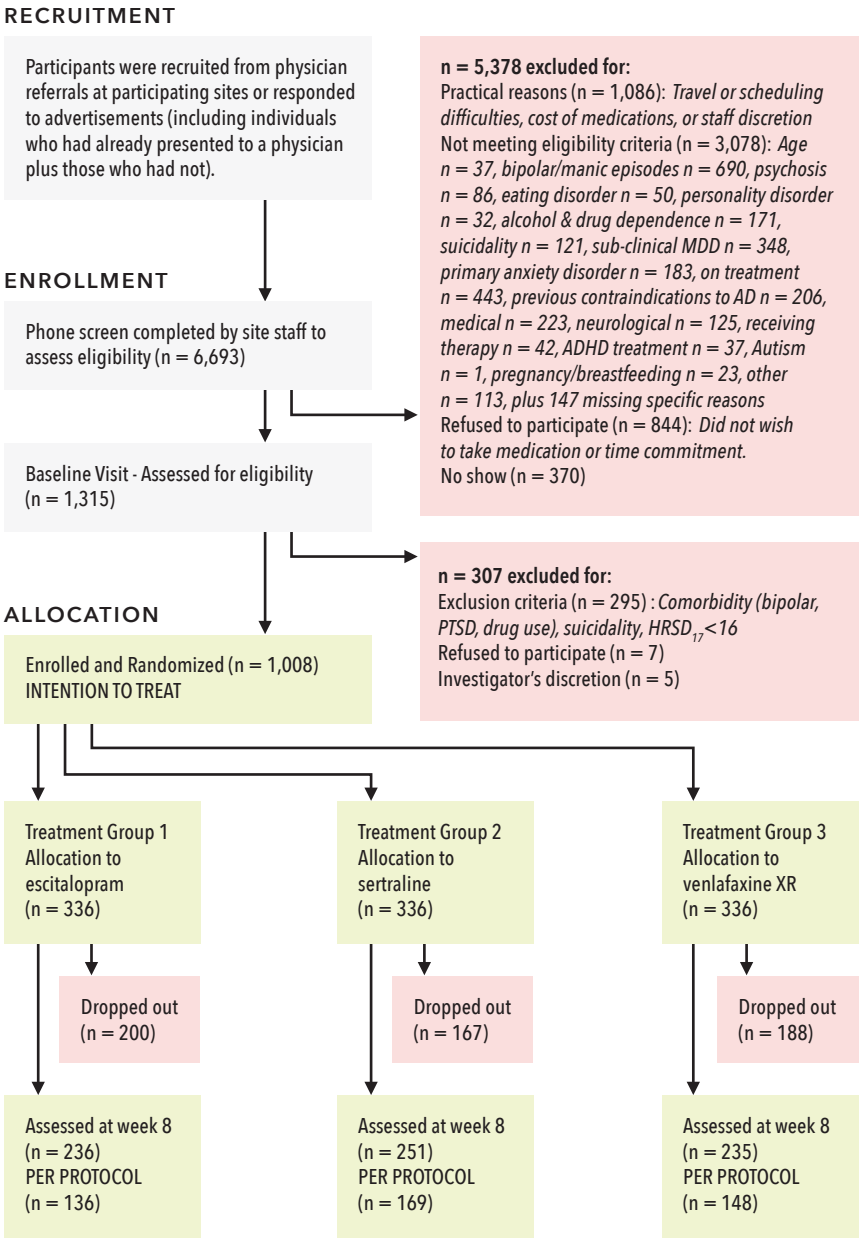


Figure 3.1: *Consort flow diagram of the iSPOT-D study. Abbreviations: ADHD, attention deficit hyperactivity disorder; AD, antidepressant treatment; $HRSD_{17}$, 17-item Hamilton rating scale for depression; MDD, major depressive disorder; PTSD, post-traumatic stress disorder; XR, extended release.*

3.2.3 PRE-TREATMENT ASSESSMENTS

EEG recordings were performed using a standardized methodology and platform (Brain Resource Ltd., Australia). Details of this procedure (Arns et al., 2008; Williams et al., 2011) and of its reliability and across-site consistency have been published elsewhere (Paul et al., 2007; Williams et al., 2005). In summary, subjects were seated in a sound and light attenuated room that was controlled at an ambient temperature of 22 °C. EEG data were acquired from 26 channels: Fp1, Fp2, F7, F3, Fz, F4, F8, FC3, FCz, FC4, T3, C3, Cz, C4, T4, CP3, CPz, CP4, T5, P3, Pz, P4, T6, O1, Oz and O2 (Quik-cap; NuAmps; 10-20 electrode international system). EEG was assessed for two minutes with eyes open (EO) (with the subject asked to fixate on a red dot on the screen) and two minutes with eyes closed (EC). The subject was instructed to remain relaxed for the duration of the recording. The operator did not intervene when drowsiness patterns were observed in the EEG. Data were referenced to averaged mastoids with a ground at AFz. Horizontal eye movements were recorded with electrodes placed 1.5 cm lateral to the outer canthus of each eye. Vertical eye movements were recorded with electrodes placed 3 mm above the middle of the left eyebrow and 1.5 cm below the middle of the left bottom eyelid. Skin resistance was <5 K Ohms for all electrodes. The sampling rate of all channels was 500 Hz. A low pass filter with an attenuation of 40 dB per decade above 100 Hz was employed prior to digitization.

3.2.4 EEG ANALYSIS

A detailed overview of the data-analysis can be found in Arns et al. (2016). In summary, data were 1) filtered (0.3-100 Hz and notch); 2) EOG-corrected using a regression-based technique similar to that used by Gratton, Coles, and Donchin (1983), segmented in 4 s epochs (50% overlapping), and an automatic de-artifacting method was applied. This EEG processing pipeline was also validated against an independent manual-processing pipeline (Arns et al., 2016). For further analysis, an average reference was applied, data were filtered (alpha power [μV^2]: 8-13 Hz) and FAA was calculated between F3 and F4 as $(F4 - F3) / (F4 + F3)$.

3.2.5 STATISTICS

Normal distribution was inspected, and appropriate transformations performed in case of non-normality. Non-log transformed alpha power was used to calculate FAA. Remission was defined as a score ≤7 on the $HRSD_{17}$ eight weeks after starting treatment (current endpoint), and response was defined as a ≥50% decrease in $HRSD_{17}$ score from baseline to eight weeks. To control for antidepressant side-effects, we employed the $VQIDS\text{-}SR_5$, developed specifically to focus on the core symptoms of depression. This enabled us to measure true depression severity, ruling out antidepressant side-effects such as physical complaints. We repeated ANOVAs from paragraphs 3.3.2 and 3.3.3 and replaced all $HRSD_{17}$ variables with VQIDS-SR5 equivalents. Results are reported in supplement 3D.

Differences in age, sex, education, and depression severity at baseline were tested using one-way ANOVA or non-parametric tests, depending on its distribution. We only included patients who returned for their week-8 visit while on their assigned medication, having followed this treatment for a minimum of 6 weeks ('per-protocol' grouping, also see the CONSORT flow diagram in figure 3.1).

FAA reliability analysis was performed by calculating intraclass correlations (ICCs) across baseline and week-8 measurements. A full-factorial Repeated Measures ANOVA was conducted with the within–subject factor FAA Change Eyes Closed (FAA at baseline and after eight weeks) and between-subject factor Treatment arm (comparing drug effects of respectively escitalopram, sertraline, and venlafaxine). Given the large sample size we set the significance level for main effects found for FAA Change in the main analyses at $p \leq .01$, for interaction effects this remained at a conventional level of $p \leq .05$. When significant interactions were found prompting subgroup analyses, again a level of $p \leq .05$ was used. Effect sizes (ES) of main effects are reported in Cohen's *d*. FAA stability was also tested through Pearson correlations between FAA Change and $HRSD_{17}$ Change.

Post hoc, we repeated the Repeated Measures and Pearson correlations analyses in the subgroups of moderately and severely depressed ($HRSD_{17}$ score of ≥24) over the age of 53, separately for males and

females (conform our meta-analysis, Van der Vinne et al., 2017). However, as these groups might lead to underpowered tests, we also performed a custom Repeated Measures ANCOVA on the whole dataset, now also including covariates Age and Depression severity, separately for males and females.

When a null hypothesis was not rejected by any of the ANOVAs or correlational analyses, we utilized Bayesian alternatives. This was done for testing evidence of *absence* of a change in FAA, using the Bayesian Repeated Measures ANOVA framework (based on work by Jeffreys (1961) and Rouder et al. (2009)). We analyzed the data with JASP (JASP Team, 2017). The first null hypotheses states that there is no difference in FAA between baseline and after eight weeks. The second that FAA Change is not correlated to $HRSD_{17}$ Change. The two-sided alternative hypotheses state that FAA changed after eight weeks, or that FAA is correlated to $HRSD_{17}$ Change.

Through a Repeated Measures model (Arns et al. 2016), we again predicted treatment outcome in females taking an SSRI (escitalopram or sertraline), while this time replacing baseline FAA with week-8 FAA (within subjects variable FAA Condition [EC and EO], and between subjects variable Response, and covariate Age). We tested effects one-tailed (halved p-values were reported) because we specifically expected more right-sided FAA in SSRI responders than in non-responders, implying that a result in the unexpected direction would lead to the same conclusion as finding no differences at all (Ruxton & Neuhäuser, 2010). In supplement 3B, we explain why we compare the smaller sample containing only patients who were present for the assessment after eight weeks, to the larger sample with *all* baseline patients from the previous study.

3.3 RESULTS

Of the 1008 MDD patients enrolled, the final MDD sample for the FAA Change analyses consisted of 453 MDD patients. The remaining 555 patients were left out of the study: they either never started treatment, had less than six weeks of medication, or had no week-8 assessment (or it was of insufficient quality) (see figure 3.1). Table 3.2 shows demographic information and response and remission rates for included patients. There were no differences between the three treatment groups regarding age, sex, baseline MDD, anxiety severity, remission and response rates, or number of rejected EEG epochs. Approximately 5.3% of EEG epochs were rejected due to artifacts for the MDD group during EC.

Table 3.2 Demographic features and treatment outcomes for patients who completed treatment

	escitalopram	sertraline	venlafaxine-XR	Total
n	136	169	148	453
Females	71	96	80	247
				Mean
% Female	52.5	56.8	54.1	54.5
Average age (years)	38.27	38.72	37.98	38.34
$HRSD_{17}$ baseline	21.45	21.74	21.45	21.56
$HRSD_{17}$ week-8	8.62	9.25	9.01	8.98
$VQIDS\text{-}SR_5$ baseline	8.01	8.34	7.99	8.13
$VQIDS\text{-}SR_5$ week-8	3.26	3.35	3.21	3.28
% Remission ($HRSD_{17}$)	51.5	46.7	44.6	47.5
% Response ($HRSD_{17}$)	66.2	66.9	66.2	66.4

3.3.2 FAA CHANGE OVER TIME

ICCs for FAA with both continuous and dichotomous (leftward or rightward FAA) variables were 0.276 and 0.256, respectively. The Repeated Measures ANOVA revealed no evidence for change in FAA after AD treatment ($F(1,450) = 1.421$, $p = .234$), nor an interaction with Treatment Arm ($F(2,450) = 0.690$, $p = .502$). FAA Change was neither significantly correlated to the change score in $HRSD_{17}$ ($r = 0.039$, $p = .410$), nor to the percentage change in $HRSD_{17}$ ($r = 0.047$, $p = .323$).

Results of Bayesian Repeated Measures testing of invariant (constant) FAA revealed a Bayes factor indicating evidence for the null hypothesis. The models with the factors FAA Change and Treatment Arm showed that the data occur >7.4 times more likely under the null hypothesis, than under any alternative model with (a combination of) the factors. Bayesian Pearson correlations between FAA Change and the difference score $HRSD_{17}$/the percentage difference of $HRSD_{17}$ reveal moderate to strong results. The data are respectively 12.1 and 9.3 times more likely to occur under the null hypothesis than under the model assuming a correlation between the variables. See supplement 3F for an elaboration on results and JASP tables.

3.3.3 EXTENDED REPEATED MEASURES MODEL AND CORRELATIONS

Focusing on variables known to have an influence on FAA, specifically in the subgroup we thought to be prone to changes in FAA (severely depressed females and males over 53 years old), we did not find significant changes, although subsample sizes were small (total $n = 27$). Furthermore, in these subgroups the FAA Change score was not significantly correlated to the change score in $HRSD_{17}$ (see table in supplement 3C for all statistics). Bayesian Repeated Measures ANOVAs for the two sex groups of severely depressed over the age of 53 reveal anecdotal (i.e. worth no more than a bare mention, a customary description for BFs ranging 1-3) to moderate results. Most models therefore provided no conclusive evidence for either the null or the alternative hypotheses, although some models indicated moderate evidence of the data being more likely to occur under the null hypothesis. See supplement 3F for an elaboration on results and JASP tables.

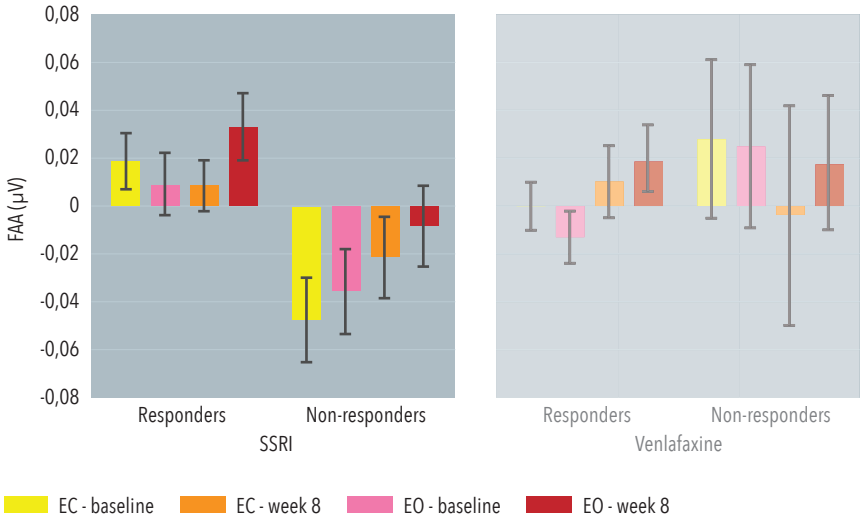
Extending the Repeated Measures model from paragraph 3.3.2 showed that - irrespective of sex - baseline severity and age are not significantly contributing to FAA Change. Bayesian Repeated Measures alternatives for the extended ANOVAs showed similar results to paragraph 3.3.2. For females, the data are ≥6.6 times more likely to occur under the null hypothesis, than under any alternative model with (a combination of) the factors, and ≥4.7 times more likely in case of males. See supplement 3F for an elaboration on results and JASP tables.

3.3.4 TREATMENT PREDICTION USING MEDICATED WEEK-8 DATA IN FEMALES

Treatment outcome prediction with week-8 data, revealed a similar prediction pattern as baseline data reported in Arns et al. (2016): one-tailed testing of the prediction of response in females taking an SSRI for depression (escitalopram or sertraline), treatment response effects remained significant with week-8 FAA on group level ($F(1,150) = 3.725$, $p = .028$). Furthermore, the response effect of FAA was again lacking after eight weeks in the venlafaxine group.

The week-8 SSRI data in figure 3.2 visualize how responders were significantly more right-sided than non-responders (based on female FAA means reported in supplement table 3E). Figure 3.2 also shows how the response effect was similar to the baseline assessment. This was despite the confidence interval (CI) of FAA in figure 3.2 (SSRI non-responders) showing no significant difference from 0 when measured with EO after eight weeks. No interactions with age were observed. The equivalent of figure 3.2 data for males is available in supplement 3G.

Cohen's *d* comparing FAA change scores of female SSRI responders and non-responders was .304. When using the direction of week-8 FAA alone to prescribe an SSRI or SNRI would have improved the overall remission rate from 47% to 56-58% for an SSRI.



3.4 DISCUSSION

We investigated the stability of FAA in MDD patients during antidepressant treatment. We hypothesized that FAA is a robust metric, insensitive to time, antidepressant drug treatment and state changes. FAA did not change significantly after eight weeks of escitalopram, sertraline, or venlafaxine treatment, despite a relatively low reliability of the FAA measurements. Additional Bayesian testing revealed that a stable FAA is more likely than a change in FAA over time after antidepressant treatment. Furthermore, post-hoc tests with variables known to have influence on FAA (in earlier iSPOT-D studies), revealed no differential temporal changes in FAA in depressed patients differing on age, sex, depression severity, or change in depression severity. Focusing on core depression symptoms only (as measured by the VQIDS-SR_5, see supplement 3D), we found similar results.

To further confirm FAA temporal stability, we hypothesized that predicting treatment outcome in females taking SSRIs would lead to similar outcome when using *week-8* FAA instead of the previously studied *baseline* FAA (Arns et al., 2016). This re-analysis indeed confirmed an overall response in the SSRI group with right-sided FAA, and a non-response with left-sided FAA. Although the effect size was less pronounced with week-8 data, week-8 FAA yielded the same conclusions as the baseline measurements, with a Cohen's *d* of .547 in the previous analyses vs our current .304. Furthermore, we yielded the same improvement in remission rates when week-8 FAA had been used for 'prescribing' medication: previous SSRI remission rates improved from 46% to 53-60% using baseline FAA, the current from 47% to 56-58% using week-8 FAA. This extends the use of FAA as a prognostic biomarker, as response prediction was neither modified by moment of assessment, nor by AD treatment.

Figure 3.2: *(facing page) Mean values of female frontal alpha asymmetry (FAA, eyes open and eyes closed [EO and EC]), for the SSRI and venlafaxine groups, split up for responders and non-responders. Error bars represent standard error of the mean. The means and error bars indicate that baseline and week-8 FAA were not significantly different in predicting treatment outcome in females; SSRI responders showed right-sided, non-responders left-sided FAA. No differences were, yet again, observed for the venlafaxine group. The equivalent of this data for males is available in supplement 3G.*

The low reliability was unexpected, and implies that FAA following treatment was not as stable as in previous studies. In several studies, FAA was found to be relatively reliable and consistent, based on ICCs and Cronbach's alpha (Allen et al., 2004; Debener et al., 2000; Keune et al., 2011; Sutton & Davidson, 1997; Towers & Allen, 2009). Especially Towers and Allen (2009) demonstrated FAA consistency, through several methods. An important difference is the use of a single FAA statistic per assessment time (two in total) in our study vs several other studies using (fictive) multiple time points. This could account for our lower reliability. Despite the low ICC, we did replicate no evidence for a significant change in FAA over time, in a large sample (n = 453).

To our knowledge, this is the first study to assess the temporal stability of FAA in a large sample. This supports previous studies showing that FAA mainly depends on a considerable number of trait-like features, insensitive to antidepressant treatment, age, sex or depression severity (Allen et al., 2004; Arns et al., 2016; Bruder et al., 2008; Carvalho et al., 2011; Deldin & Chiu, 2005; Feldmann et al., 2018; Gollan et al., 2014; Keune et al., 2011; Nusslock et al., 2018; Spronk et al., 2008; Sutton & Davidson, 1997; Tomarken et al., 1992; Van der Vinne et al., 2017; Vuga et al., 2006). Similarly, Segrave and colleagues (2011) showed no evidence for antidepressant elicited changes in FAA when comparing a small group of depressed patients on ADs with unmedicated patients. In other small cohorts, FAA was not modified by the use of antidepressive medication either (Bruder et al., 2008; Vuga et al., 2006), in agreement with our observations.

In the prevailing approach-withdrawal motivation system hypothesis, it is assumed that FAA is associated with lifetime MDD (having had at least one depressive episode in one's life), and not specifically current MDD. This is an important distinction, and our results initially support this theory. The motivation system hypothesis states that FAA is not expected to change as a result of changes in MDD status, and ultimately not with MDD remission. However, with establishing FAA (in)stability, our study would neither provide evidence for, nor against the theory. That is, if we would have found the opposite result (a change in FAA), this could have been explained as well,

by the related capability model (Coan, Allen, & McKnight, 2006). This model states that resting state FAA is more prone to fluctuations than FAA measured after inducing positive or negative mood. Because we measured resting state FAA, either outcome could be explained within the approach-withdrawal motivation system, given the capability model. Therefore, it is difficult to unambiguously place our results in the existing theories. Note that our earlier findings were less compatible with the motivation system: Firstly, in the approach-withdrawal motivation system, left-sided FAA is theorized to be more associated with withdrawal behavior and depression. But brain asymmetry was found not to be different in these groups as measured both through EEG FAA (Van der Vinne et al., 2017), and through fMRI in a recent large ENIGMA consortium study (de Kovel et al., 2019). Secondly, prognostic results for females in the FAA iSPOT-D study (Arns et al., 2016) revealed heterogeneity in MDD patients, not consistent with assuming a homogenic FAA related vulnerability for MDD. In sum, the current study was not designed to directly investigate the approach-withdrawal motivation theory, and cannot provide support in favor of or against the theory.

We show that FAA is a robust metric, suitable for sex specific treatment prediction under challenging circumstances, such as state, time, the use of common antidepressive agents and drug changes. This suggests reliable implementation in clinical practice as a prognostic biomarker in both medicated and unmedicated patients.

CONCLUSIONS

In an adequately powered sample, we demonstrate that 1) neither antidepressant medication, 2) nor MDD state and severity, have systematic effects on FAA. This confirms FAA stability. Furthermore, as prognosis of treatment response is irrespective of the moment of measurement, FAA may serve as a robust biomarker to optimize MDD treatments.

Table S3.1 Overview of FAA stability related studies

Study	Year	Study type[1]	n =	Subjects	EEG methods	Intervention	Relevant factors
Allen et al.	2004	1	30	MDD (females)	3 to 5 Ax., 8 or 16 weeks apart	Acupuncture (specific and non-specific)	HRSD change score
Bruder et al.	2008	1	18	MDD and HC	2 Ax., 12 weeks apart	Fluoxetine treatment	Response ("CGI-I rating much or very much improved")
Debener et al.	2000	1	15 and 22	MDD and HC	2 Ax., 2-4 weeks apart	Several antidepressants	BDI-score
Deldin & Chiu	2005	1	15 and 18	MDD and HC	4 Ax. On 1 day	Cognitive restructuring	Happiness change score
Gollan et al.	2014	1	37 and 35	MDD and HC	2 Ax., 16 weeks apart	Behavioral activation	IDS-SR
Keune et al.	2011	1	78	MDD	2 Ax., 8 weeks apart (neutral vs sad state)	Mindfulness	BDI and BDI-Change, gender
Spronk et al.	2008	1	8	MDD	2 Ax., pre/post-treatment	rTMS	None that was associated with frontal asymmetry
Vuga et al.	2006	1	49 and 50	Childhood onset depression and HC	2 Ax., 1-3.2 years apart	13 of 49 cases on ADs	Age, sex, BDI
Davidson et al.	2003	2	41	HC	2 Ax., 8 weeks, 4 months	Mindfulness meditation	
Hagemann et al.	2002	2	59	HC	4 Ax., all 4 weeks apart	None	
Hagemann et al.	2005	2	59	HC	3 Ax., all 5 weeks apart	None	
Sutton & Davidson	1997	2	46	HC	2 Ax., 6 weeks apart	None	No depression scores (only BIS/BAS and PANAS)
Tenke et al.	2017	2	39	HC	2 Ax., 5-16 days apart	None	
Tomarken et al.	1992	2	85	HC	2 Ax., 3 weeks apart	None	
Carvalho et al.	2011	3	12, 8 and 7	MDD, remitted and HC	1 Ax.	None	BDI
Feldmann et al.	2018	3	22, 16 and 34	MDD, remitted and HC (also other groups, with comorbid anxiety)	1 Ax.	None	BDI
Gotlib et al.	1998	3	16, 31 and 30	MDD, remitted and HC	1 Ax.	None	
Grünewald et al.	2018	3	28 and 31	MDD and HC	1 Ax.	None	BDI
Nusslock et al.	201	3	37 and 69	MDD and HC	1 Ax.	None	BDI

[1]Type 1: Multiple assessment moments with depressed patients. Type 2: Multiple assessment moments, only healthy controls. Type 3: Cross-sectional study. abbreviations used :MDD = major depressive disorder, HC = healthy controls, Ax. = assessment(s), HRSD = Hamilton Rating Scale for Depression, CGI = Clinical Global Impression, BDI = Beck Depression Inventory, IDS-SR = Inventory of Depressive Symptomatology-Self Report, BIS/BAS = Behavioral Avoidance/Inhibition Scales, PANAS = Positive and Negative Affect Scale.

Table S3.1 Overview of FAA stability related studies, continued (1)

Study	Relevant analyses	Calculation FAA
Allen et al.	FAA solely: ICCs. FAA in relation to symptoms: Correlations & multivariate Repeated Measures analysis of variance (HRSD-score) with changing covariates (3 asymmetry measures)	ln[right] - ln[left]
Bruder et al.	1. Interaction Response-Nonresponse-HC * Hemisphere * Site. 2. Interaction Response * Site * Hemisphere * Session. 3. Test-retest correlations.	Interaction of site and hemisphere
Debener et al.	1. Cronbachs alpha FAA for internal consistency. 2. Correlation FAA-BDI. 3. Temporal stability through Pearson correlations.	1. Interaction of session * site * region (posterior-anterior) * Hemisphere 2. ln(right) - ln(left)
Deldin & Chiu	1. Interaction Diagnosis * Block * Region * Laterality 2. Correlation FAA-Happiness score	1. Interaction with region and laterality 2. ln F4 - ln F3
Gollan et al.	1. Rep Measures ANOVA FAA over time. 2. Correlation pre FAA-pre IDS/pre FAA-post IDS/post FAA-pre IDS/post FAA-post IDS	log F4 - log F3
Keune et al.	1. Cronbachs alpha. 2. ANOVA interaction FAA time 1 * time 2. 3. Correlation FAA-affect scores (to test state). 4. Test-retest reliability with Pearson product moment correlations. 5. Correlation FAA-BDI time 1. 6. Correlation FAA Change-BDI Change. 7. Interactions with gender	subtracting power values in the left hemisphere from the values in the right hemisphere
Spronk et al.	Interaction of Time * Hemisphere (left: F3, FC3, F7. Right: F4, FC4, F8)	Interaction of time and hemisphere
Vuga et al.	1. Cronbach's alpha. 2. ANOVA with Age, apart for the sexes. 3. ICC 4. ANOVA with group, sex, and FAA at Time 2 as dependent variable, on Time 1 as dependent variable. 5. Same analyses, apart for with and without medication. 6. Regression on FAA Time 2, with FAA Time 1, BDI and BDI-change.	ln(F4) - ln(F3)
Davidson et al.	Group (mindfulness/waitlist) * Time.	ln(F4) – ln(F3), ln(F8) – ln(F7)
Hagemann et al.	Model of LST theory	ln power density F4 – ln power density F3
Hagemann et al.	Model of LST theory	ln F4 - ln F3
Sutton & Davidson	Averaged over 13 asymmetry measures (from 13 homologous electrode pairs): Cronbach's alpha (.87) and ICC (.57)	log F4 - log F3
Tenke et al.	Test-retest correlations	F4-F3
Tomarken et al.	1. t-tests of asymmetry measures. 2. ICCs and Pearson correlations. 3. Cronbach's alphas	log R minus log L power density
Carvalho et al.	1. ANOVA on FAA, with hemisphere and group. 2. Correlation FAA-BDI.	ln[right]— ln[left]
Feldmann et al.	1. One-way ANOVAs on FAA between the HC, Mda- and rMDa- . 2. Correlation FAA-BDI	ln[right ROI]-ln[left ROI]
Gotlib et al.	Regression with 1st predictor "Never vs Ever depressed " and 2nd predictor "Currently depressed vs remitted"	log R - log L
Grünewald et al.	Correlation FAA-BDI	ln[right]-ln[left]
Nusslock et al.	Correlation FAA-BD	right - left

ICC = Intraclass Correlation Coefficient, HRSD = Hamilton Rating Scale for Depression, HC = healthy controls, BDI = Beck Depression Inventory, LST theory = latent state-trait theory, Avg Ref = average reference

Table S3.1 Overview of FAA stability related studies, continued (2)

Study	Conclusion
Allen et al.	Stable across time and independent of depression severity.
Bruder et al.	Overall stable. 1. No FAA differences between groups. 2. No change of FAA over time. 3. Moderate test-rest correlations of FAA after treatment.
Debener et al.	Overall: Not a stable measure. 1. Good internal consistency. 2. No correlation with BDI, so state-independent. 3. Unstable temporal stability of frontal regions (not posterior).
Deldin & Chiu	Overall: no changes between assessments on the same day. 2. No correlation.
Gollan et al.	Overall: FAA = stable, trait-like. 1.No changes in FAA. 2. No significant correlations.
Keune et al.	1. Stable. 2. Change (in sad condition, not in neutral). 3. No correlation with affect, so stable. 4. Correlates, so reliable. 5. Significant correlation FAA-BDI in sad condition (not in neutral, this was only for other sites). 6. No correlation. 7. No gender interaction effects.
Spronk et al.	No interaction for alpha1, alpha2 and alpha. Alpha seems a trait.
Vuga et al.	Overall: Moderate to high long term stability. 1. Consistent. 2. No influence of Age or Sex. 3. Moderately stable. 4. Stable after p-value correction, without correction there would be differences. 5. Same results with and without medication, so stable. 6. Depressive symptom severity and change in symptoms did not affect EEG asymmetry stability.
Davidson et al.	Changes in asymmetry were only observed in other electrode pairs, not in F3/4 or F7/8.
Hagemann et al.	60% trait - 40% state
Hagemann et al.	40-50% state
Sutton & Davidson	High to modest test-retest reliability
Tenke et al.	Variation in outcome: low in general, but it is correlating. Test-retest correlations differ per research site, but is in general significant for F3-F4: .371 (see supplement).
Tomarken et al.	1. Stable: "…asymmetry measures tended to be associated with nonsignificant shifts in mean values over time". 2. Not very high, buy significant ICCs (.66 for Avg Ref baseline only, .79 for Avg Ref across time). 3. Acceptable-to-excellent Cronbach's alphas.
Carvalho et al.	Overall: doubtful, whether cross-sectional data is sufficient to establish trait properties. But there are no indications against FAA having trait properties. 1. No interaction especially between MDD and remitted, but also controls. 2. No correlation.
Feldmann et al.	The most relevant results: 1. No differences between MDD and remitted. 2. No correlations.
Gotlib et al.	Most relevant results: No difference between currently depressed and remitted. FAA seems to be a state independent marker.
Grünewald et al.	Overall no specific conclusion on state or trait, but no correlations were found in the MDD group.
Nusslock et al.	Overall: "[our results] highlight the trait like quality of reduced relative left frontal EEG activity." 1. No correlation.

3.5.2 SUPPLEMENT 3B: COMPARISON BASELINE AND WEEK-8 DATA

To justify the use of a follow-up sample that is supposed to contain the same MDD patients as the baseline data (paragraph 3.3.4), but does not due to incomplete assessments, we performed the baseline analysis from Arns et al. (2016) on only those who *did* have a complete week-8 assessment. The effect within the SSRI group was the same (p = .001, $F(1,150)$ = 10.619, see table S3.2 for all statistics).

Table S3.2 P-values of mentioned interaction effects in the re-analysis of Arns et al. (2016) with data only of MDD patients who had measurements after eight weeks (thus excluding FAA baseline measurements of patients who did not return for follow-up)

	Original analysis	Original analysis without patients with no follow-up	Re-analysis with week-8 FAA[1]
Females SSRI: Response	p = .001	p = .001	p = .028
Females venlafaxine: Response	p = .070	p = .011	p = .411

[1]Halved p-values due to one-tailed analysis

3.5.3 SUPPLEMENT 3C

Table S3.3 Statistics paragraph 3.3.3

			F(df)	p(F)	r	p(r)
A[1]	Females	FAA Change	2.080 (1,14)	.171		
		FAA Change * Treatment arm	2.425 (2,14)	.125		
	Males	FAA Change	0.092 (1,7)	.771		
		FAA Change * Treatment arm	0.061 (2,7)	.941		
	Females	FAA Change * $HRSD_{17}$ Change			0.259	.316
	Males	FAA Change * $HRSD_{17}$ Change			-0.070	.849
B[1]	Females	FAA Change	0.355 (1,235)	.552		
		FAA Change * Treatment arm	0.714 (2,235)	.491		
		FAA Change * Age	0.889 (1,235)	.344		
		FAA Change * Depression severity	0.645 (1,235)	.423		
		FAA Change * Treatment arm * Age	0.849 (2,235)	.429		
		FAA Change * Treatment arm * Depression severity	0.846 (2,235)	.430		
		FAA Change * Age * Depression severity	1.254 (1,235)	.264		
		FAA Change * Treatment arm * Age * Depression severity	1.148 (2,235)	.319		
	Males	FAA Change	0.029 (1,194)	.864		
		FAA Change * Treatment arm	0.282 (2,194)	.755		
		FAA Change * Age	0.024 (1,194)	.878		
		FAA Change * Depression severity	0.022 (1,194)	.881		
		FAA Change * Treatment arm * Age	0.292 (2,194)	.747		
		FAA Change * Treatment arm * Depression severity	0.471 (2,194)	.625		
		FAA Change * Age * Depression severity	0.052 (1,194)	.820		
		FAA Change * Treatment arm * Age * Depression severity	0.352 (2,194)	.704		

[1]A: Severely depressed ≥53 years old only. B: Whole dataset.

3.5.4 SUPPLEMENT 3D: VQIDS-SR5

To control for AD side effects, we repeated analyses from paragraphs 3.3.2 and 3.3.3 and replaced all $HRSD_{17}$ variables with VQIDS-SR_5 equivalents. Correlational analyses showed that FAA Change was neither significantly correlated to the change score in VQIDS-SR_5 (r = 0.059, p = .225), nor to the percentage change in VQIDS-SR_5 (r = 0.060, p = .219).

Focusing on variables known to have an influence on FAA, specifically in the subgroup we thought to be prone to changes in FAA (severely depressed females and males over 53 years old), we did not find the FAA Change score to be significantly correlated to the change score in VQIDS-SR_5, although subsample sizes were small. Extending the Repeated Measures model from paragraph 3.3.2 showed that VQIDS-SR_5 baseline severity and age are not significantly contributing to FAA Change, both in males and females (see table S3.4 for all statistics).

Table S3.4 VQIDS-SR_5 statistics for paragraphs 3.3.2 and 3.3.3

			F(df)	p(F)	r	p(r)
A[1]	Females	FAA Change * VQIDS Change			-0.121	.644
	Males	FAA Change * VQIDS Change			0.127	.381
B[1]	Females	FAA Change	0.530 (1,225)	.467		
		FAA Change * Treatment arm	0.002 (2,225)	.998		
		FAA Change * Age	0.930 (1,225)	.336		
		FAA Change * VQIDS Depression severity	0.125 (1,225)	.724		
		FAA Change * Treatment arm * Age	0.066 (2,225)	.936		
		FAA Change * Treatment arm * VQIDS Depression severity	0.145 (2,225)	.865		
		FAA Change * Age * VQIDS Depression severity	0.384 (1,225)	.536		
		FAA Change * Treatment arm * Age * VQIDS Depression severity	0.351 (2,225)	.705		
	Males	FAA Change	0.991 (1,225)	.321		
		FAA Change * Treatment arm	1.491 (2,225)	.228		
		FAA Change * Age	0.407 (1,225)	.524		
		FAA Change * VQIDS Depression severity	1.214 (1,225)	.272		
		FAA Change * Treatment arm * Age	0.773 (2,225)	.463		
		FAA Change * Treatment arm * VQIDS Depression severity	1.739 (2,225)	.179		
		FAA Change * Age * VQIDS Depression severity	0.654 (1,225)	.420		
		FAA Change * Treatment arm * Age * VQIDS Depression severity	1.158 (2,225)	.316		

[1] *A: Severely depressed ≥53 years old only. B: Whole dataset.*

3.5.5 SUPPLEMENT 3E

Table S3.5 FAA means of the different subgroups reported in paragraph 3.3.4. Split on sex, medication type, EEG condition, response group, and time of assessment

			Baseline		Week-8	
Sex	Medication type	EEG condition[1]	Response	Non-response	Response	Non-response
Female	SSRI	EC	0.019	-0.048	0.009	-0.022
		EO	0.009	-0.036	0.033	-0.008
	SNRI	EC	0.000	0.028	0.010	-0.004
		EO	-0.013	0.025	0.020	0.018
Male	SSRI	EC	0.003	0.017	0.013	0.030
		EO	0.015	0.036	0.044	0.036
	SNRI	EC	-0.015	-0.028	-0.031	-0.023
		EO	-0.010	-0.045	-0.036	0.002

[1] *EC = eyes closed, EO = eyes open*

3.5.6 SUPPLEMENT 3F: BAYESIAN REPEATED MEASURES ANOVA AND CORRELATIONS

F1 ELABORATED BAYESIAN ANALYSES PARAGRAPH 3.3.2

Results of Bayesian testing of an absence of change in FAA, revealed a Bayes factor indicating evidence for the null hypothesis: the models with the factors FAA Change and Treatment Arm showed that the data occur >7.4 times more likely under the null hypothesis, than under any alternative model with (a combination of) the factors. This means that moderate evidence for the null hypothesis was found with only FAA Change in the model (BF_{01} = 7.483), increasing to (very) strong evidence when adding a combination of the two main effects (BF_{01} = 240.356) and including their interaction effect (BF_{01} = 5,109.119). The error percentage was <2.5%, which indicates sufficient stability of the numerical algorithm that was used to obtain the result. For each factor, the $BF_{inclusion}$ reflects how well the factor predicts the data by comparing the performance of all models that include the factor to the performance of all the models that do not include the factor. For both the factors FAA Change and Treatment Arm, there is weak evidence in favor of their inclusion ($BF_{inclusion}$ = 0.134 and 0.031 respectively), as well as a weak evidence in favor of the inclusion of the interaction effect ($BF_{inclusion}$ = 0.047). This implies that these factors are not providing evidence for change in FAA. See table S3.6.1 on page 68 for all results.

Table S3.6.1 Bayesian Repeated Measures ANOVA main analysis

Model comparison

Models	P(M)	P(M|data)	BF_M	BF_{01}	error %
Null model (incl. subject)	.200	.856	23.749	1.000	
FAA Change	.200	.114	0.517	7.483	1.276
Treatment	.200	.026	0.107	32.853	0.604
FAA Change + Treatment	.200	.004	0.014	240.356	2.282
FAA Change + Treatment + FAA Change * Treatment	.200	1.675e-4	6.702e-4	5109.119	2.471

Note: All models include subject

Table S3.6.1 ***continued***

Analyses of effects

Effects	P(incl)	P(incl|data)	$BF_{Inclusion}$
FAA Change	.400	.118	0.134
Treatment	.400	.030	0.031
FAA Change * Treatment	.200	1.675e-4	0.047

Note: Compares models that contain the effect to equivalent models stripped of the effect. Higher-order interactions are excluded.

Bayesian Pearson correlations between FAA Change and the difference score $HRSD_{17}$/the percentage difference $HRSD_{17}$ reveal moderate to strong results, where the data are respectively 12.1 and 9.3 times more likely to occur under the null hypothesis than under the model assuming there is a correlation between the variables. See table S3.6.2 for all results.

Table S3.6.2 Bayesian Pearson correlations
FAA Change vs $HRSD_{17}$ Change/$HRSD_{17}$ % Change

			r	BF_{01}
FAA Change	vs	$HRSD_{17}$ Change	0.039	12.111
FAA Change	vs	$HRSD_{17}$ % Change	0.052	9.275

F2 ELABORATED BAYESIAN ANALYSES PARAGRAPH 3.3.3

Bayesian Repeated Measures ANOVAs for the two sex groups of severely depressed over the age of 53 reveal anecdotal (i.e. worth no more than a bare mention, a customary description for BFs ranging 1-3) to moderate results. Males: BF_{01} = 1.351-2.715 for models with only main effects, BF_{01} = 6.195 for the model with the interaction; $BF_{inclusion}$ = 0.438-0.748; *error%* = 0.701-2.327. Females: BF_{01} = 1.864-2.944 for most models, BF_{01} = 4.304 for the model with only main effects of FAA Change and Treatment Arm; $BF_{inclusion}$ = 0.434-1.462; *error%* = 0.922-1.372. Most models therefore provided no conclusive

evidence for either the null or the alternative hypotheses, and BFinclusions indicate that there is (very) weak evidence in favor of including the factors. However, some models indicated moderate evidence of the data being more likely to occur under the null hypothesis. See tables S3.6.3 and S3.6.4 for all results.

Table S3.6.3 Bayesian Repeated Measures ANOVA for severely depressed males ≥53 years old

Model comparison

Models	P(M)	P(M|data)	BF_M	BF_{01}	error %
Null model (incl. subject)	.200	.363	2.282	1.000	
FAA Change	.200	.175	0.851	2.070	0.701
Treatment	.200	.269	1.472	1.351	0.687
FAA Change + Treatment	.200	.134	0.618	2.715	1.744
FAA Change + Treatment + Time * Treatment	.200	.059	0.249	6.195	2.327

Note: All models include subject

Table S3.6.3 ***continued***

Analyses of effects

Effects	P(incl)	P(incl|data)	$BF_{Inclusion}$
FAA Change	.400	.309	0.489
Treatment	.400	.403	0.748
FAA Change * Treatment	.200	.059	0.438

Note: Compares models that contain the effect to equivalent models stripped of the effect. Higher-order interactions are excluded.

Table S3.6.4 Bayesian Repeated Measures ANOVA for severely depressed females ≥53 years old

Model comparison

Models	P(M)	P(M|data)	BF_M	BF_{01}	error %
Null model (incl. subject)	.200	.393	2.592	1.000	
FAA Change	.200	.211	1.069	1.864	1.400
Treatment	.200	.171	0.825	2.299	0.528
FAA Change + Treatment	.200	.091	0.402	4.304	0.922
FAA Change + Treatment + FAA Change * Treatment	.200	.134	0.617	2.944	1.372

Note: All models include subject

Table S3.6.4 ***continued***

Analyses of effects

Effects	P(incl)	P(incl|data)	$BF_{Inclusion}$
FAA Change	.400	.302	0.536
Treatment	.400	.262	0.434
FAA Change * Treatment	.200	.134	1.462

Note: Compares models that contain the effect to equivalent models stripped of the effect. Higher-order interactions are excluded.

Bayesian Repeated Measures alternatives for the extended ANOVAs showed similar results to paragraph 3.3.3: for females, the data are ≥6.6 times more likely to occur under the null hypothesis than under the alternative hypothesis (only models including factor FAA Change: $BF_{inclusion}$ FAA Change and FAA Change X Treatment Arm 0.152 and 0.102, *error*% ≤ 8.576), and ≥4.7 times more likely in case of males (only models including factor FAA Change: $BF_{inclusion}$ Time and Time X Treatment Arm 0.132 and 0.151, *error*% ≤ 5.582). See tables S3.6.5 and S3.6.6 for all results.

Table S3.6.5 Bayesian Repeated Measures ANOVA for females, with factors and covariates Treatment Arm, Age and Baseline $HRSD_{17}$

Model comparison

Models	P(M)	P(M|data)	BF_M	BF_{01}	error %
Null model (incl. subject)	.050	.547	22.983	1.000	
FAA Change	.050	.083	1.720	6.596	1.069
Age	.050	.092	1.935	5.922	1.199
FAA Change + Age	.050	.014	0.268	39.377	1.598
Baseline $HRSD_{17}$	.050	.097	2.036	5.657	1.928
FAA Change + Baseline $HRSD_{17}$	.050	.015	0.286	36.858	1.939
Age + Baseline $HRSD_{17}$	.050	.027	0.534	20.007	1.962
FAA Change + Age + Baseline $HRSD_{17}$	.050	.004	0.077	134.758	2.073
Treatment	.050	.073	1.490	7.526	0.651
FAA Change + Treatment	.050	.011	0.216	48.653	1.854
Age + Treatment	.050	.013	0.243	43.438	1.488
FAA Change + Age + Treatment	.050	.002	0.039	268.859	4.110
Baseline $HRSD_{17}$ + Treatment	.050	.013	0.255	41.259	1.331
FAA Change + Baseline $HRSD_{17}$ + Treatment	.050	.002	0.040	263.804	1.689
Age + Baseline $HRSD_{17}$ + Treatment	.050	.004	0.076	137.616	3.325
FAA Change + Age + Baseline $HRSD_{17}$ + Treatment	.050	5.979e-4	0.011	915.659	1.734
FAA Change + Treatment + FAA Change*Treatment	.050	.001	0.022	472.071	5.124
FAA Change + Age + Treatment + FAA Change * Treatment	.050	1.915e-4	0.004	2858.225	2.712
FAA Change + Baseline $HRSD_{17}$ + Treatment + FAA Change * Treatment	.050	2.204e-4	0.004	2483.772	8.576
FAA Change + Age + Baseline $HRSD_{17}$ + Treatment + FAA Change * Treatment	.050	5.817e-5	0.001	9410.129	2.373

Note: All models include subject

Table S3.6.5 *continued*

Analyses of effects

Effects	P(incl)	P(incl|data)	$BF_{Inclusion}$
FAA Change	.400	.132	0.152
Age	.500	.157	0.187
Baseline $HRSD_{17}$	.500	.163	0.195
Treatment	.400	.119	0.135
FAA Change *Treatment	.200	.002	0.102

Note: Compares models that contain the effect to equivalent models stripped of the effect. Higher-order interactions are excluded.

Table S3.6.6 Bayesian Repeated Measures ANOVA for males, with factors and covariates Treatment Arm, Age and Baseline $HRSD_{17}$

Model comparison

Models	P(M)	P(M|data)	BF_M	BF_{01}	error %
Null model (incl. subject)	.050	.189	4.416	1.000	
FAA Change	.050	.025	0.492	7.471	3.978
Treatment	.050	.303	8.262	0.622	0.600
FAA Change + Treatment	.050	.040	0.787	4.740	1.459
FAA Change + Treatment + FAA Change * Treatment	.050	.006	0.118	30.614	2.419
Age	.050	.047	0.929	4.045	1.842
FAA Change + Age	.050	.006	0.111	32.350	1.464
Treatment + Age	.050	.060	1.203	3.166	1.480
FAA Change + Treatment + Age	.050	.008	0.152	23.809	2.818
FAA Change + Treatment + Age + FAA Change * Treatment	.050	.001	0.022	162.929	2.264
Baseline $HRSD_{17}$	.050	.081	1.684	2.316	2.516
FAA Change + Baseline $HRSD_{17}$	.050	.010	0.201	18.048	1.736
Treatment + Baseline $HRSD_{17}$	.050	.130	2.832	1.454	1.023
FAA Change + Treatment + Baseline $HRSD_{17}$	.050	.018	0.339	10.743	2.659
FAA Change + Treatment + Baseline $HRSD_{17}$ + FAA Change*Treatment	.050	.003	0.049	73.444	2.043
Age + Baseline $HRSD_{17}$	.050	.028	0.547	6.740	1.240
FAA Change + Age + Baseline $HRSD_{17}$	.050	.004	0.070	51.141	2.066
Treatment + Age + Baseline $HRSD_{17}$	.050	.037	0.728	5.113	1.253
FAA Change + Treatment + Age + Baseline $HRSD_{17}$	.050	.005	0.097	37.061	5.852
FAA Change + Treatment + Age + Baseline $HRSD_{17}$ + FAA Change*Treatment	.050	7.334e-4	0.014	257.148	2.230

Note: All models include subject

Table S3.6.6 ***continued***

Analyses of effects

Effects	P(incl)	P(incl|data)	$BF_{Inclusion}$
FAA Change	.400	.116	0.132
Age	.400	.600	1.538
Baseline $HRSD_{17}$	.500	.195	0.243
Treatment	.500	.316	0.462
FAA Change *Treatment	.200	.011	0.151

Note: Compares models that contain the effect to equivalent models stripped of the effect. Higher-order interactions are excluded.

3.5.7 SUPPLEMENT 3G: MALE DATA EQUIVALENT TO FIGURE 3.2 WITH FEMALE DATA

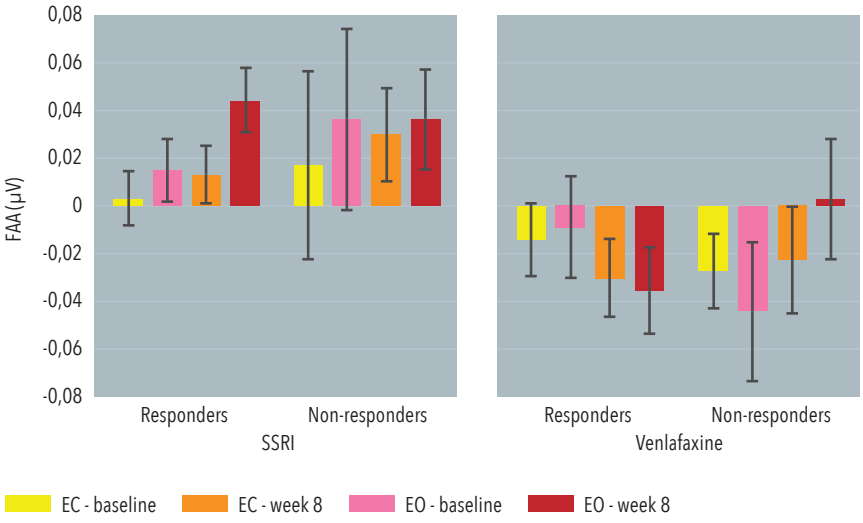


Figure S3.1: *Mean values of male frontal alpha asymmetry (FAA, eyes open and eyes closed [EO and EC]), for the SSRI and venlafaxine groups, split up for responders and non-responders.*

4

NORMALIZATION OF EEG IN DEPRESSION AFTER ANTIDEPRESSANT TREATMENT WITH SERTRALINE? A PRELIMINARY REPORT

N. van der Vinne, M. A. Vollebregt, N. N. Boutros, K. Fallahpour, M. J. A. M. van Putten and M. Arns
Journal of Affective Disorders, 2019; 259: 67–72

ABSTRACT

Introduction

MDD patients with abnormal EEG patterns seem more likely to be non-responsive to the antidepressants escitalopram and venlafaxine, but not sertraline, than patients without EEG abnormalities. This finding suggests that patients with both MDD and abnormal EEGs may differentially respond to antidepressant treatment. In the current study, we investigated whether depressed patients with an abnormal EEG show a normalization of the EEG related to antidepressant treatment and response and whether such effect is drug specific, and whether having had early life stress (ELS) increases the chance of abnormal activity.

Methods and materials

Baseline and week 8 EEGs and depression symptoms were extracted from a large multicenter study (iSPOT-D, n = 1008) where depressed patients were randomized to escitalopram, sertraline, or venlafaxine-XR treatment. We calculated odds ratios of EEG normalization and depression response in patients with an abnormal EEG at baseline, comparing sertraline versus other antidepressants.

Results

Fifty-seven patients with abnormal EEGs were included. EEGs did not normalize significantly more with sertraline compared to other antidepressants (OR = 1.9, p = .280). However, patients with a normalized EEG taking sertraline were 5.2 times more likely to respond than subjects taking other antidepressants (p = .019). ELS was not significantly related to abnormal activity.

Limitations

Neurophysiological recordings were limited in time (two times 2 min EEGs) and statistical power (n = 57 abnormal EEGs).

Conclusions

Response rates in patients with normalized EEG taking sertraline were significantly larger than in subjects treated with escitalopram/venlafaxine. This adds to personalized medicine and suggests a possible drug repurposing for sertraline.

4.1 INTRODUCTION

Abnormal activity in the electroencephalogram (EEG) in absence of clinical events (such as seizures) includes paroxysmal activity (e.g. isolated epileptiform discharges (IEDs), intermittent focal slowing) or diffuse slowing of the background pattern. Abnormal EEG activity is not exclusively associated with disorders such as epilepsy and may occur without obvious clinical signs or symptoms. However, it has been hypothesized that particular EEG abnormalities are associated with multiple mental disorders (Boutros 2018; Inui et al., 1998; Shelley & Trimble, 2009; Yasuhara 2010), opting the possibility that patients with an abnormal EEG may benefit from medication targeting both the mental disorder and the EEG abnormalities. Assessing pre-treatment EEG abnormalities therefore, could both benefit prognosis reliability and treatment outcome and would adhere to the standards set by the NIMH with the Research Domain Criteria framework ([RDoC] Insel et al., 2010, under Arousal (Physiology)).

EEG abnormalities are reported in 3–5% of patients with MDD, similar to controls (1–6%: Arns et al., 2017; Arns et al., 2008; Goodwin, 1947; Lennox-Buchtal et al., 1960; Monin et al., 2018; Oh et al., 2018; Richter et al., 1971; Shelley et al., 2008). Furthermore, an increased likelihood of developing epilepsy exists in MDD (Kanner et al., 2017b). Vice-versa, the most commonly reported psychiatric comorbidity in epilepsy is MDD (Bragatti et al., 2014), as 24% of epileptic patients are affected by MDD (Kanner, 2017a).

Ribot, Ouyang, and Kanner (2017) found a decrease in seizure frequency as a result of antidepressant treatment in depressed humans with co-morbid epilepsy. This was irrespective of how well their mood symptoms improved. In animal models, antidepressants (AD) showed a similar decrease in seizure frequency (Kamal, 2007).

Only few studies explored how abnormal EEGs, in the absence of epilepsy, relate to the treatment effects in affective disorders. Arns et al. (2017) showed that a subgroup of MDD patients with abnormal EEG patterns was more likely to be non-responsive to the ADs escitalopram and venlafaxine, whereas response to sertraline was not different for patients with or without EEG abnormalities. These findings suggest that patients with both MDD and abnormal EEGs may differentially respond to AD treatment.

In this light, sertraline might exert some anticonvulsant effects, given the observation that people with abnormal EEGs did respond well to sertraline. This would justify the consideration of repurposing the AD sertraline as a mild anticonvulsant for the treatment of MDD with paroxysmal activity. To explore such repurposing, we studied if sertraline treatment results in more EEG normalization after eight weeks of treatment, compared to venlafaxine and escitalopram, and if this is associated with clinical response. Sertraline responders were expected to show more normalization than responders to other ADs. This was complemented by evaluating if early life stress (ELS, a potential cause of an abnormal EEG) is associated with abnormal EEG patterns. For background and results, see supplement 4.5. Our study was not originally designed to detect EEG abnormalities (particularly IEDs) and recordings were limited to 2 min of eyes-open and 2 min of eyes-closed EEG. This is sufficient for detecting slow wave abnormalities (Struve & Boutros, 2005). However, short recordings increase the chances for a false negative recording for IEDs. Despite these limitations, we sought to capitalize on this large sample size of well-characterized patients with MDD to investigate whether normalization after AD treatment occurs in EEGs initially showing abnormalities.

4.2 METHODS AND MATERIALS

4.2.1 DESIGN

This study was an international multicenter, randomized, prospective open-label trial (phase IV clinical trial, international Study for Predicting Optimized Treatment in Depression [iSPOT-D]) in which MDD subjects were randomized to escitalopram, sertraline, or venlafaxine-XR treatment in a 1:1:1 ratio. The study protocol details have been published by Saveanu et al. (2015) and Williams et al. (2011).

The iSPOT-D sample consisted of 1008 patients with MDD and 336 healthy controls, the current study focused on 58 subjects having an abnormal EEG (identical to the sample from our previous report, Arns et al., 2017). After excluding data from one subject due to poor EEG quality at endpoint, the sample for our main analyses consisted of 57 patients with MDD displaying an abnormal EEG at baseline measurement. A complete description of the study assessments, inclusion/exclusion criteria, diagnostic procedures, treatment and characterization of this paroxysmal subgroup is available in Arns et al. (2017). In summary, the primary diagnosis of nonpsychotic MDD was confirmed at the baseline visit (before randomization) using the Mini-International Neuropsychiatric Interview (MINI-Plus), according to the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) criteria, and a score ≥16 on the clinician-rated 17-item Hamilton Rating Scale for Depression (HRSD). To measure the neurophysiological consequences of childhood trauma (as was reported in Williams et al., 2016), we used the Early-Life Stress Questionnaire (ELSQ, McFarlane et al., 2005). The ELSQ comprises 18 items, which assess exposure to specific traumatic events in the first 17 years of life (see supplement 4.6 for the entire questionnaire) and which are equivalent to the trauma items assessed by the Child Abuse and Trauma Scale (Chu et al., 2013; Williams et al., 2016). Each item is scored dichotomously for the presence/absence of exposure to each type of trauma. As reported by Williams et al. (2016), in the complete iSPOT-D sample the extent of exposure to traumatic events was not found to differ across site or across country. All MDD subjects were either AD medication naive or, if previously prescribed an AD, had undergone a washout period of at least 5 half-lives before the baseline visit clinical and EEG assessments. After the baseline

visit, MDD subjects were randomized to one of the three AD medications. After eight weeks of treatment, subjects were tested again using the HRSD and EEG, further referred to as endpoint. Subjects provided written informed consent. This study was approved by the local institutional review boards at all the participating sites and was registered at ClinicalTrials.gov (NCT00693849).

4.2.2 PRE AND POST-TREATMENT ASSESSMENTS

EEG recordings were performed at baseline and at endpoint, using a standardized methodology and platform (Brain Resource Ltd, Australia). Details of this procedure have been published elsewhere (Williams et al., 2011), and details of the reliability and across-site consistency of this EEG procedure have also been published (Paul et al., 2007; Williams et al., 2005). In summary, subjects were seated in a sound and light attenuated room that was controlled at an ambient temperature of 22 °C. EEG data were acquired from 26 channels: Fp1, Fp2, F7, F3, Fz, F4, F8, FC3, FCz, FC4, T3, C3, Cz, C4, T4, CP3, CPz, CP4, T5, P3, Pz, P4, T6, O1, Oz, and O2 (NuAmps; 10–20 electrode international system). EEG data were collected for 2 min with eyes open (EO) and 2 min with eyes closed (EC). Data were referenced to averaged mastoids with a ground at AFz. Horizontal eye movements were recorded with electrodes placed 1.5 cm lateral to the outer canthus of each eye. Vertical eye movements were recorded with electrodes placed 3mm above the middle of the left eyebrow and 1.5 cm below the middle of the left bottom eyelid. Skin resistance was <5 kOhm for all electrodes. A low-pass filter with an attenuation of 40 dB per decade above 100 Hz was employed prior to digitization with a sampling rate of 500 Hz.

4.2.3 EEG ANALYSIS

A high-pass filter of 0.3 Hz, a low-pass filter of 100 Hz, and notch filters of 50 or 60 Hz (depending on the country in which the data were recorded) were applied. Data were EOG corrected using a regression based technique similar to the method described by Gratton, Coles, and Donchin (1983). No other artifact rejection was applied to the

data other than EOG correction and filtering. Data were visually inspected in Brain Vision Analyzer (Brainproducts, Germany) using a linked ears and queens-square montage. Visual inspection and classification were performed by author N.N.B. (a board-certified electroencephalographer, neurologist, and psychiatrist), who was blinded to the subject's status (patient vs control), clinical data, assessment moment (baseline vs endpoint) and treatment arm. Details on EEG data analysis and validation can be found elsewhere (Arns et al., 2016). Eyes closed awake EEG data were examined for the presence of any focal or generalized slowing (EEG slowing). Diffuse slowing was recorded if the background frequency was consistently below the alpha range (Niedermeyer, 2005). Focal slowing was recorded if rhythms slower than alpha (theta or delta, i.e. <8 Hz) were consistently detected in a particular location (Krauss et al., 2010). Epileptiform or paroxysmal activity were defined as any EEG pattern (with or without a sharp contour) that emerges and disappears paroxysmally from the ongoing background activity (Niedermeyer, 2005). Non-paroxysmal, focal or generalized, slow wave activity were more or less continuously recorded (note that records were almost entirely fully awake records) with some waxing and waning (Sharbrough, 2005). Finally, the presence of any of the so-called controversial waveforms (e.g., wicket spikes) was also recorded. These waveforms are paroxysmal but are of uncertain significance (Boutros, 2014). Supplement 4.7 contains descriptive and visual information of abnormalities on the individual level. Classification of all abnormalities was in accordance to the guidelines published by the International Federation of Clinical Neurophysiology (Noachtar et al., 1999).

4.2.4 STATISTICS

Our primary outcome measure was normalization of the EEG, defined as an absence of abnormalities after eight weeks of AD treatment, for those subjects who were classified as having an abnormal baseline EEG. Treatment response was defined as a more than 50 percent decrease in HRSD score from baseline to endpoint, identical to our previous study. Secondary, we focused on treatment response in patients with a normalized EEG. Differences in age, sex, and depression severity at baseline were tested using one-way analysis of

variance (or nonparametric tests if required). Although pharmacologically different presumed mechanisms of action, we chose to merge escitalopram and venlafaxine groups based on the prior finding that these two drugs show poor clinical response in patients with an abnormal EEG (as opposed to the sertraline group), thereby also equalizing statistical power. For the presence of abnormalities at endpoint, significance level was set at $p \leq .05$ and effect sizes of main effects were reported in odds ratios (OR) with 95% confidence intervals (CI). For assessing the relationship between ELS and EEG abnormalities, both ELS in general and abuse in particular, were tested binary (OR with CI) and continuously (summing up all experiences ELS events, testing through logistic regression), for their neurophysiological effect (having an abnormal EEG at baseline or not). Every questionnaire item was also singularly tested binary through the calculation of ORs.

4.3 RESULTS

In the abnormal sample, 45% achieved remission (HRSD score below MDD cutoff at endpoint) and 54% reached response to treatment (see table 4.1). The number of subjects with an abnormal EEG was significantly higher in the group prescribed with sertraline than other antidepressants, yielding unequal distributions for the number of responders among treatment groups. For the additional ELS analyses, the total sample of patients with complete EEG and ELS data (n = 1152) consisted of 878 patients with MDD and 274 healthy controls.

Responders (31.6 ± 10.1 years) did not significantly differ in age (37.3 ± 13.2 years; $p = .065$; $F(1,56) = 3.535$) or sex ($p = .541$; $\chi^2 = 0.374$) from non-responders. As previously reported (Arns et al., 2017), response rates in this abnormal subgroup were highest in the sertraline group (72%), followed by the other AD group (38%). Based on this sample, escitalopram and venlafaxine were associated with nonresponse in MDD patients with EEG abnormalities, but not so for sertraline. See table 4.1 for an overview of response rates.

Table 4.1 Response results after MDD treatment with antidepressants, and EEG outcome at endpoint

		Response				EEG at endpoint			
	Subgroup	Responders		Non-responders		Normal		Abnormal	
	total (*n*)	*n*	%	*n*	%	*n*	%	*n*	%
Sertraline	25	18	72%	7	28%	19	76%	6	24%
Other AD	32	12	37.5%	20	62.5%	20	62.5%	12	37.5%
Venlafaxine	*17*	*7*	*41%*	*10*	*59%*	*10*	*59%*	*7*	*41%*
Escitalopram	*15*	*5*	*33%*	*10*	*67%*	*10*	*67%*	*5*	*33%*
Total	57	30	53%	27	47%	39	68%	18	32%

Drug effect on abnormal EEGs. Out of 57 subjects with an abnormal EEG at baseline, 39 showed no more EEG abnormalities at endpoint (76% after sertraline, 62.5% after another AD). Results are shown in table 4.1. Comparing sertraline to the other ADs, OR analyses showed no significant differences in normalization of the EEG at endpoint (*OR* = 1.9, p = .280, 95% *CI* [0.593,6.084]).

Mediation of EEG normalization on clinical response to different drugs. Responders and non-responders were equally likely to have a normal EEG at endpoint (table 4.2). However, when specifically comparing sertraline to other ADs among subjects with a normalized EEG, subjects taking sertraline were 5.2 times more likely to be a responder than subjects taking other ADs (p = .019, 95% *CI* [1.317,20.539], see figure 4.1). For subjects with an unchanged EEG this difference was not significant (*OR* = 2.8, p = .325, 95% *CI* [0.361,21.727]).

Table 4.2 The number of responders and non-responders in the normal and abnormal endpoint EEG subgroups per treatment and across treatment arms.

		Endpoint normal		Endpoint abnormal	
		n	%*	*n*	%*
Sertraline	Responders	14	56%	4	16%
	Non-responders	5	20%	2	8%
Other AD	Responders	7	22%	5	16%
	Non-responders	13	40%	7	22%
Venlafaxine	*Responders*	4	23.5%	3	18%
	Non-responders	6	35%	4	23.5%
Escitalopram	*Responders*	3	20%	2	13%
	Non-responders	7	47%	3	20%
Total	Responders	21	37%	9	16%
	Non-responders	18	31%	9	16%

**Percentage of the total of all patients belonging to the same treatment arm (both responders and non-responders)*

Early-life stress. Having experienced ELS in general, abuse in general, and specific ELS events in early life, were not related to abnormal EEG. This was the case for 1) the entire iSPOT-D dataset (all ELS data available for MDD patients and controls in the iSPOT-D dataset n = 1152, $p \geq .116$), and 2) depressed subjects specifically (n = 878, $p \geq .091$): ORs and logistic regression models were non-significant.

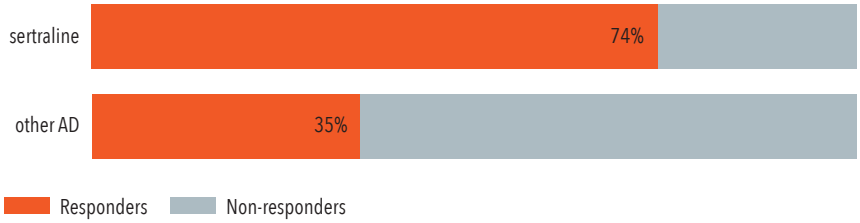


Figure 4.1: *Treatment response for subjects with a normalized EEG at endpoint, per AD type (sertraline n = 19, other AD n = 20). Percentages indicate response within each respective AD type.*

4.4 DISCUSSION

This study aims to extend our observations that depressed patients with an abnormal EEG at baseline are more likely to be a non-responder to venlafaxine or escitalopram, suggesting that sertraline is the preferred treatment in this subgroup (Arns et al., 2017). We investigated whether these patients show both an antidepressant (AD) treatment related and response related temporal change in the EEG and whether early life stress (ELS) increases the chance of observing abnormal EEG activity. We show that EEG normalization patterns – rated blind to diagnosis and treatment – did not significantly differ between sertraline and other AD treatment for eight weeks. However, when comparing sertraline treatment to the other ADs within the subjects who had a normalized EEG at endpoint, response was more likely to be achieved with sertraline than with other treatments. A mediating effect of ELS in developing an abnormal EEG (with ELS increasing the chance of such an EEG) was neither found in the whole sample of controls and depressed, nor in the depressed subsample.

To the best of our knowledge, this is the first study to explore the effects of ADs on abnormal EEGs of depressed patients. Several research groups did study the effects of a range of ADs on seizures (in-

cluding sertraline), finding no adverse influence on seizure frequency (Hovorka et al,, 2000; Kanner et al., 2000; Maguire et al., 2014; Okazaki et al., 2011; Thomé-Souza et al., 2007). Moreover, beneficial treatment effects of selective serotonin (and norepinephrine) reuptake inhibitors (SSRIs and SNRIs) on patients with epilepsy have been published before (Alper et al. , 2007; Favale et al., 2003; Hamid & Kanner, 2013; Kanner, 2016; Specchio et al., 2004). Abnormal EEGs however, were not subject to such investigations before. Our results suggest that depression response with concurrent EEG normalization occurs more often after sertraline treatment than escitalopram or venlafaxine. Note that this entails an assessment of abnormalities normally deemed subclinical by neurologists. Our inconclusive findings on ELS (whole sample n = 1152 and MDD sample n = 878) provided no additional clarification that would help elucidate underlying mechanisms involved in both developing MDD and displaying an abnormal EEG (see supplement 4.5 for background and results).

Remarkable is the differential effect of sertraline in terms of treatment response, compared to escitalopram and venlafaxine in subjects who had a normalized EEG, which was discussed before (Arns et al., 2017). Only escitalopram also acts on the allosteric sites of the serotonin transporter and has higher selectivity for this transporter. Sertraline on the other hand, has the most pronounced dopamine active transporter inhibitory activity (Sanchez et al., 2014). However, new literature on the different antidepressant profiles has barely clarified how this might have consequential effects on the (abnormal) human EEG. In a rodent model, sertraline prevented or diminished induced seizures (presumably mostly through effective inhibition of brain presynaptic Na+ channel permeability), comparable to antiepileptic drug carbamazepine (Sitges et al., 2012). Kanekar et al. (2018) did suggest a differential working mechanism in sertraline compared to other ADs, considering the increased inhibition of dopamine active transporter by sertraline. According to Bozzi and Borelli (2013), altered D2R signaling (part of the dopamine system), leading to decreased D2R function, might be involved in epileptogenesis, adding to earlier findings that dopamine plays a role in juvenile myoclone epilepsy and general tonic clonic seizures (Ciumas et al., 2010; Ciumas et al., 2008). Although the generalization of our results on EEG

abnormalities to epileptogenesis is uncertain, it might help explain in part how EEG abnormalities in depressed patients are best treated with sertraline, but further research is necessary to understand the mechanisms that may be underlying the differential effects of treatments.

LIMITATIONS

As the frequency of IEDs is highly variable and in part modulated by sleep (Askamp & van Putten, 2014; Geut et al., 2017), our short EEG recordings will have limited sensitivity to detect IEDs (Boutros, 2018). However, the most observed abnormalities in this sample, diffuse and focal slow activity, have a relatively high chance of being observed in one recording session as these findings are relatively stationary. Another limitation is that all EEG epochs were assessed by visual analysis. While this is still the gold standard for detection of paroxysms, global features (e.g. mean frequency) may be assessed more reliably with quantitative EEG. Further, we did not assess spontaneous fluctuations in these findings, nor differentiate between the different types of abnormalities (different types of IEDs and slowing).

In the search of treatment optimization, future studies involving larger samples (e.g. from consortia) allow for investigating whether the degree of abnormality is indicative of the chance of treatment success. New methods in neuroscience such as deep learning, could assist in this quest (Tjepkema-Cloostermans et al., 2018) as well as using a multimodal and integrative approach where data from various domains is combined in order to optimize prediction, including cognitive, psychological and genetic information (e.g. Spronk, et al., 2011). For future similar studies, we propose that records are interpreted by two independent EEG specialists, both blinded to group and treatment. Possible changes in EEG of participants that showed no abnormalities at baseline have not been investigated in the current study. This may, however, help explain the normalization of EEG after treatment or lack thereof. Findings from Sitges et al. (2012) imply that EEG normalization could occur after sertraline treatment, irrespective of treatment response. As our findings were not based on the ratio of normalization in responders and non-responders, future

studies could further investigate EEG normalization of sertraline irrespective of response status. Future studies should also address the relation between duration of the EEG and likelihood of detection of paroxysms to define the optimal recording time within practical limits. Although in previous epileptiform EEG studies SSRIs and SNRIs have been discussed as a treatment in general, our results show the importance to discriminate between specific ADs and incorporate several ADs within one study.

In conclusion, albeit in a small sample, our data demonstrate that patients showing a normalized EEG after MDD treatment comprised of more sertraline responders, than responders to escitalopram or venlafaxine. This first exploration of the relationship between EEG paroxysms and MDD treatment response points towards the suggestion that differentiation within a psychiatric patient group that seems homogenous at first (with respect to its symptoms), may improve treatment efficacy. To this end, a more routinely use of EEG in psychiatry could assist in personalized medicine.

4.5 SUPPLEMENT: EARLY LIFE STRESS AND PAROXYSMAL ACTIVITY

INTRODUCTION

Traumas, early life stress (ELS), and lifetime assaultive violence have been linked to cortical mal-development and increased EEG abnormalities (Davies, 1978; Ito et al., 1998; Ito et al., 1993; Teicher et al., 2003, 1997). The influence of ELS on depression treatment involving abnormal EEGs might help elucidate possible underlying mechanisms involved in both developing MDD and displaying an abnormal EEG.

The number of both healthy and depressed subjects reporting an overall trauma, or specifically childhood abuse as operationalization of ELS, is higher in the abnormal EEG subgroup. To this end, like Arns et al. (2015), we used data from iSPOT-D (see Saveanu et al., 2015 and Williams et al., 2011 for details). The size of this sample including well-characterized patients with MDD (full sample of $n = 1008$) allowed for a reliable investigation of the effects of antidepressants, in a subsample of patients with a paroxysmal EEG (showing either IEDs or EEG background slowing, abnormal $n = 58$).

METHODS

To measure the neurophysiological consequences of childhood trauma (as was reported in Williams et al., 2016), we used the Early-Life Stress Questionnaire (ELSQ, McFarlane et al., 2005). The ELSQ comprises 18 items, which assess exposure to specific traumatic events in the first 17 years of life (see supplement 4.6 for the entire questionnaire) and which are equivalent to the trauma items assessed by the Child Abuse and Trauma Scale (Chu et al., 2013). Each item is scored dichotomously for the presence/absence of exposure to each type of trauma. As reported by Williams et al. (2016), in the complete iSPOT-D sample the extent of exposure to traumatic events was not found to differ across site or across country. For assessing the relationship between ELS and EEG abnormalities, both ELS in general and abuse in particular, were tested binary (OR with CI) and continuously (summing up all experiences ELS events, testing through

logistic regression), for their neurophysiological effect (having an abnormal EEG or not). Every questionnaire item was also singularly tested binary through the calculation of ORs.

RESULTS

Having experienced ELS in general, abuse in general, and specific ELS events in early life, were not related to abnormal EEG. This was the case for 1) the entire iSPOT-D dataset (all ELS data available for MDD patients and controls in the iSPOT-D dataset $n = 1152$, $p \geq .116$), and 2) depressed subjects specifically ($n = 878$, $p \geq .091$): ORs and logistic regression models were non-significant. Only for depressed, an opposite trend was visible where subjects with an abnormal EEG were 2.2 times *less* likely to have had experienced ELS than those with a normal EEG ($p = .091$).

DISCUSSION

The absence of a confirmed association between ELS and abnormal EEGs was not in line with the literature, where abused and traumatized children showed an increased chance of presenting with electrophysiological abnormalities (Davies, 1978; Ito et al., 1998; Ito et al., 1993; Teicher et al., 2003, 1997). Inconsistencies may result from different methods of categorizing deviant EEGs. Our findings comprise events that might not be severe enough, or do not contain particular characteristics, to be linked to ELS. Importantly, in the same full-sample dataset as currently investigated, Williams et al. (2016) showed that specifically the ELS type abuse (either physical, sexual, or emotional) was associated with non-response to AD treatment, especially to sertraline. Both these findings are opposite of the expected effect based on the ELS literature: After ELS would have increased the chance of an abnormal EEG, these depressed presumably should in turn benefit from sertraline (Arns et al., 2015). Perhaps ELS could account for the less frequent non-response to sertraline, in the abnormal EEG subgroup. Our sample size, however, does not allow for further division into subgroups to investigate such a hypothesis.

4.6 SUPPLEMENT: EARLY LIFE STRESS QUESTIONNAIRE ITEMS

Table S4.1 Items of the Early Life Stress Questionnaire

ELSQ items: (Yes/No)	Yes	No
Were you born prematurely- or experience other birth complications?	☐	☐
Were you adopted?	☐	☐
Did you undergo major surgery or repeated hospitalization?	☐	☐
Did you experience a life-threatening illness or injury?	☐	☐
Did you experience sustained bullying or rejection by schoolmates?	☐	☐
Were you physically abused?	☐	☐
Were you sexually abused?	☐	☐
Were you emotionally abused?	☐	☐
Did you experience extreme poverty or neglect?	☐	☐
Did you witness first-hand a natural disaster such as earthquake, flood or fire?	☐	☐
Was your house destroyed by fire or other means?	☐	☐
Did you witness warfare?	☐	☐
Did your parents divorce or separate?	☐	☐
Were you separated for a long period from a parent, brother or sister?	☐	☐
Was there sustained conflict within your family?	☐	☐
Did one of your parents, a brother or sister die?	☐	☐
Did one of your parents, a brother or sister experience a life-threatening illness?	☐	☐
Did you witness domestic violence within your family?	☐	☐

4.7 SUPPLEMENT: EEG ABNORMALITIES

Table S4.2 Characteristics of abnormalities on the individual level

	Slow[1]		IED[1]		Any abnormal activity[1]			Slow[1]		IED[1]		Any abnormal activity[1]	
ID	Base-line	End-point	Base-line	End-point	Base-line	End-point	ID	Base-line	End-point	Base-line	End-point	Base-line	End-point
1	1	1	0	0	1	1	31	1	0	0	0	1	0
2	0	1	1	0	1	1	32	0	1	1	1	1	1
3	1	0	0	0	1	0	33	1	0	0	0	1	0
4	1	1	0	0	1	1	34	1	0	0	0	1	0
5	1	1	1	1	1	1	35	1	0	0	0	1	0
6	1	1	0	0	1	1	36	1	0	0	0	1	0
7	1	1	0	0	1	1	37	1	0	0	0	1	0
8	1	0	0	0	1	0	38	1	0	0	0	1	0
9	0	1	1	1	1	1	39	1	0	0	0	1	0
10	1	1	0	0	1	1	40	1	0	0	1	1	1
11	1	0	0	0	1	0	41	1		0		1	
12	1	0	0	0	1	0	42	0	0	1	0	1	0
13	1	0	0	0	1	0	43	0	1	1	1	1	1
14	0	0	1	0	1	0	44	0	1	1	0	1	1
15	0	1	1	0	1	1	45	1	0	0	0	1	0
16	1	0	0	0	1	0	46	0	0	1	0	1	0
17	1	0	0	0	1	0	47	0	1	1	1	1	1
18	1	0	0	0	1	0	48	1	0	0	0	1	0
19	1	0	0	0	1	0	49	1	0	0	0	1	0
20	1	0	0	0	1	0	50	1	0	0	0	1	0
21	0	1	1	0	1	1	51	1	0	0	0	1	0
22	1	0	0	0	1	0	52	0	1	1	0	1	1
23	1	0	0	0	1	0	53	1	0	0	0	1	0
24	0	0	1	0	1	0	54	0	1	1	0	1	1
25	1	0	0	0	1	0	55	0	0	1	0	1	0
26	1	0	0	0	1	0	56	1	0	0	0	1	0
27	1	0	0	0	1	0	57	1	1	0	0	1	1
28	0	0	1	0	1	0	58	0	0	1	0	1	0
29	0	0	1	0	1	0							
30	0	0	1	0	1	0							

[1] *Binomial scoring of having a normal (0) or abnormal (1) EEG*

Table S4.2, *continued*

	Baseline				
ID	Type of slow[2]	Site of slow[3]	Type of IED	Site of IED[3]	Comments
1	DIFF 7				Background borderline slow
2				LT	Focal mild slowing
3	DIFF 7, theta				Background borderline slow, mild
4	Transients	LT			Mild
5	Bilateral	Maximal LT	LT		Moderate severity
6	Transients	RT			Mild
7	Both temp	Independent	Independent		Mild
8	Transients, theta	RT			Mild
9			Theta	Maximal LT	Paroxysmal/moderate severity
10	Transients	RT			Mild
11	Transients	RT			Mild
12	Transients	RT			Mild
13	Transients	LT			Mild
14			Wickets	LT	Barely normal background/mild
15			Theta burst	LT	Theta burst
16	Transients	RT			Mild
17	Transients	LT			Mild
18	Transients	RT			Mild
19	Slow focus	RT			Moderate severity
20	Transients	RT			Mild
21			Theta bursts	Diffuse	Frequent, moderate severity
22	Transients	RT			Mild
23	Transients	RT			Mild
24			Sharp	LT	Moderate severity
25	Transients	RT			Mild
26	Slow	Diffuse			Slow background/moderate severity
27	Transients	LT			Mild
28			Wicket spikes	Wicket spikes	Mild
29			Sharp	LT	Moderate severity
30			Sharp	LT	Moderate severity

[2]DIFF 7 = Mild slowing of the background
[3]Sites: LT = left temporal, RT = right temporal

Table S4.2, ***continued***

	Baseline				
ID	Type of slow	Site of slow	Type of IED	Site of IED	Comments
31	Theta	Diffuse			Moderate severity
32			Theta bursts	Diffuse	Many episodes. Very abnormal
33	Transients	RT			Mild
34	Transients	LT			Mild
35	Theta	Diffuse			Barely normal background, mild
36	Theta	Diffuse			Mild
37	Theta	Diffuse			Barely normal background, mild
38					Borderline diffuse slowing, mild
39	Transients	LT			Mild
40	Transients	LT			Sharp transient, moderate severity
41	Transients	RT			Mild
42			RMTD	Both temp	Controversial paroxysms, mild
43			Theta burst	Maximal RT	Theta paroxysm, moderate severity
44			Delta burst	RT	Significant abnormality
45	Theta				Borderline diffuse slowing, mild
46			Theta burst	Theta burst	One theta burst, mild
47			Delta burst	Diffuse	Maximal LT
48	Focal slowing	Both temp	Both temp		Foci are independent, moderate severity
49	Theta	Diffuse			Moderate severity
50	Transients	LT			Mild
51	Theta	Diffuse			Moderate severity
52			Theta bursts	Diffuse	Mild slowing
53	Transients, theta	RT			Mild
54			Theta bursts	Maximal LT	Mild focal slow
55			Theta bursts	Diffuse	Moderate severity
56	Focal slowing	Both temp			Slowing on both temp. regions, moderate severity
57	Theta	Diffuse			Mild slow
58					

Table S4.2, ***continued***

	Endpoint, week 8				
ID	EEG quality	Type of slow	Site of slow	IED/other comments	Site of IED
1		Theta	LT	Mild	
2		Theta	LT	Mild	
3					
4		Theta	LT	Mild	
5		Theta	Both temp	Moderate severity	
6		Theta	RT	Mild	
7		Theta	LT	Mild	
8					
9		Theta	Both temp	Paroxysmal/moderate severity	Paroxysmal/moderate severity
10		Theta	RT	Mild	
11					
12					
13					
14				Barely normal background, mild	Barely normal background, mild
15		Theta		Theta burst	LT
16					
17					
18					
19					
20					
21		Delta		Delta burst, moderate severity	Diffuse
22					
23					
24					
25					
26					
27					
28					
29					
30					

Table S4.2, ***continued***

	Endpoint, week 8				
ID	EEG quality	Type of slow	Site of slow	IED/other comments	Site of IED
31					
32		Theta		Bursts, significant abnormal	Diffuse
33					
34					
35					
36					
37					
38					
39					
40				Sharp transient, moderate severity	LT
41	Bad, excluded				
42					
43		Theta	Diffuse	Theta paroxysm, moderate severity	
44		Theta	LT		
45					
46					
47		Delta bursts	Delta bursts	Delta bursts, significant abnormal	Diffuse, but maximal LT
48					
49					
50					
51					
52		Theta	Diffuse	Mild slowing	
53					
54		Theta	LT	Mild focal slow	
55					
56					
57		Theta	Diffuse	Mild slow	
58					

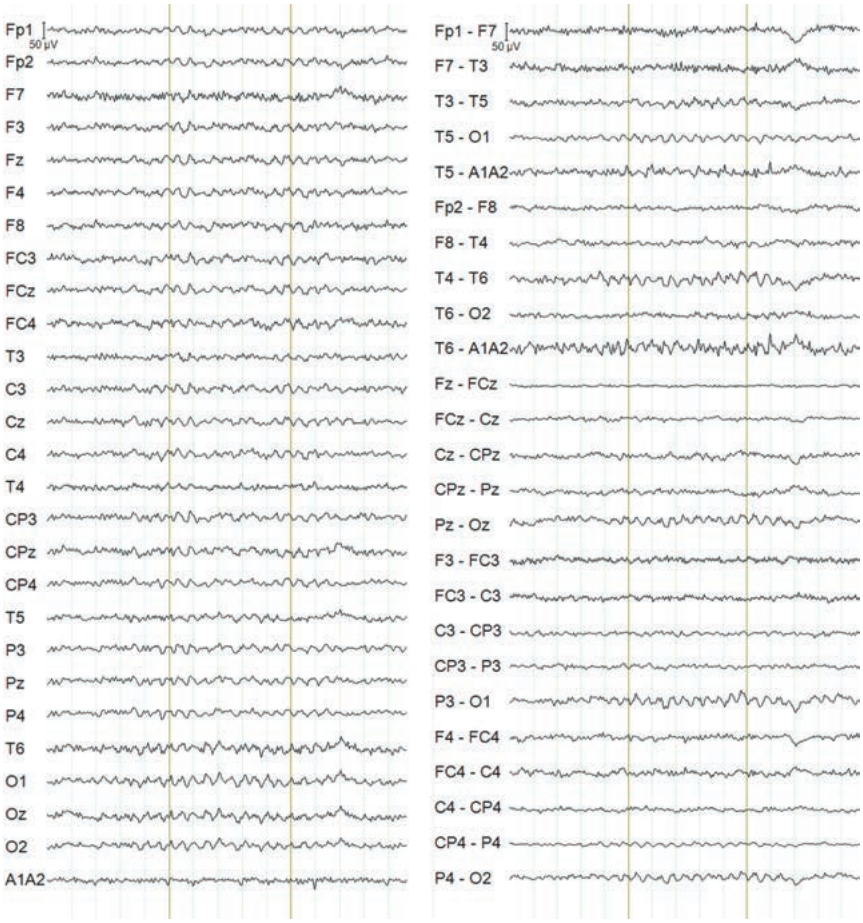


Figure S4.1: *Example of an EEG segment without abnormal events, displayed in two montages.*

Filter settings: 1-45 Hz.
Timescale: 1 s per vertical line.

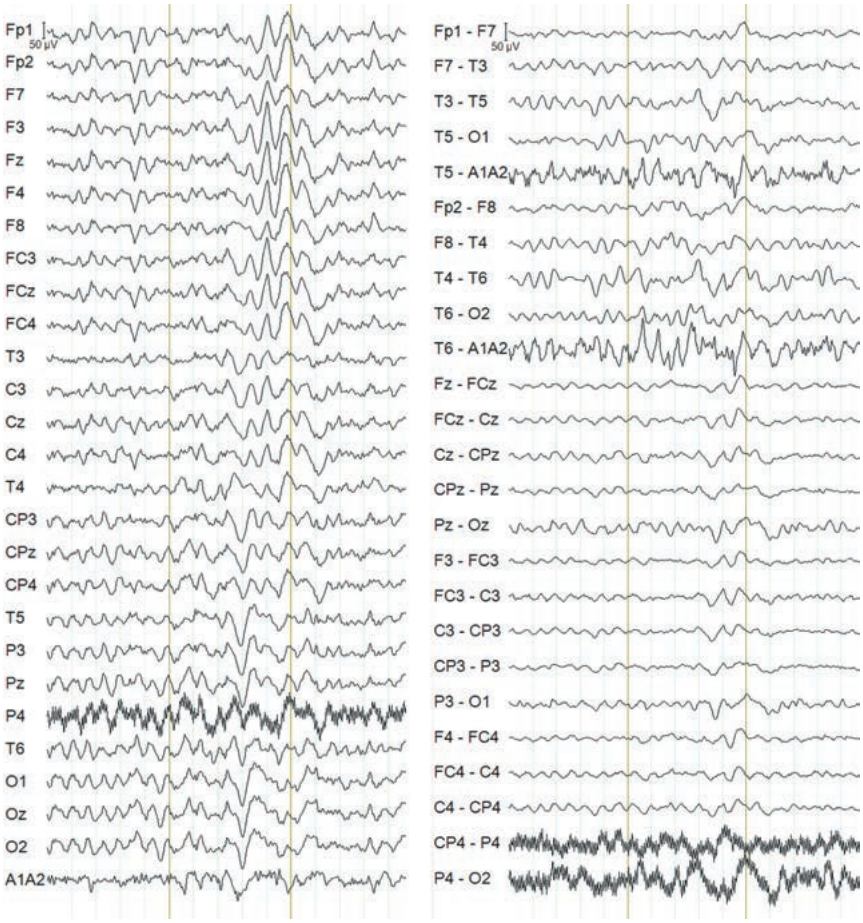


Figure S4.2: *Example of an EEG segment with a theta burst, displayed in two montages. Note that in the bipolar montage the burst is mainly located to the temporal regions.*

Filter settings: 1-45 Hz.
Timescale: 1 s per vertical line.

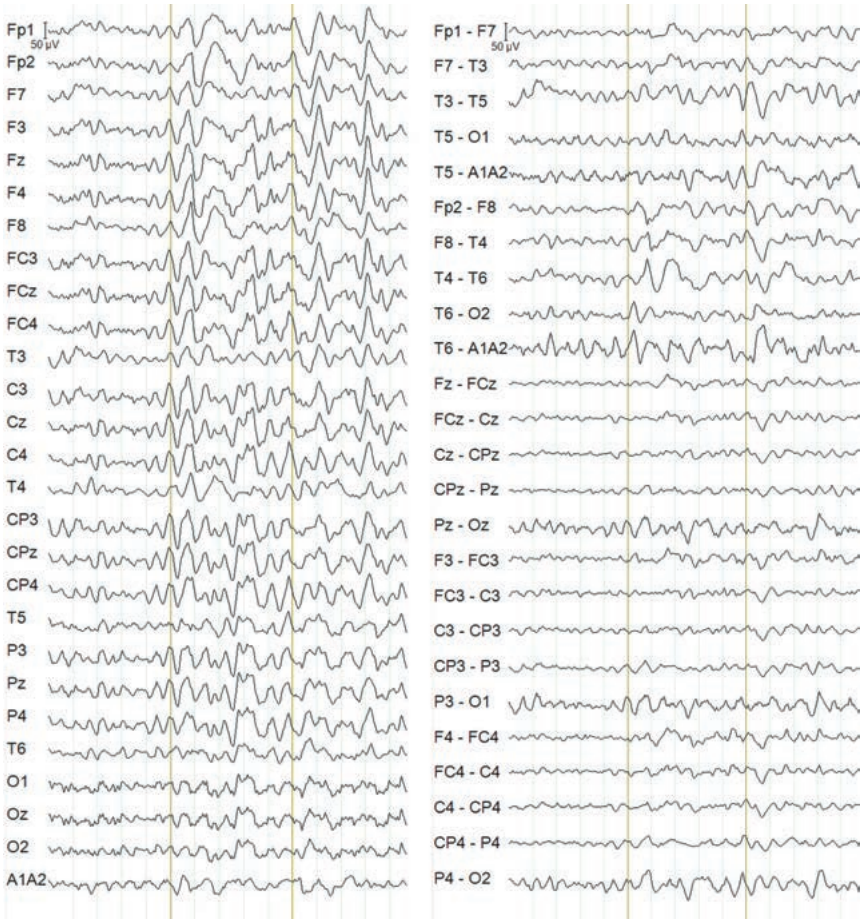


Figure S4.3: *Example of an EEG segment with a delta burst, displayed in two montages.*

Filter settings: 1-45 Hz.
Timescale: 1 s per vertical line.

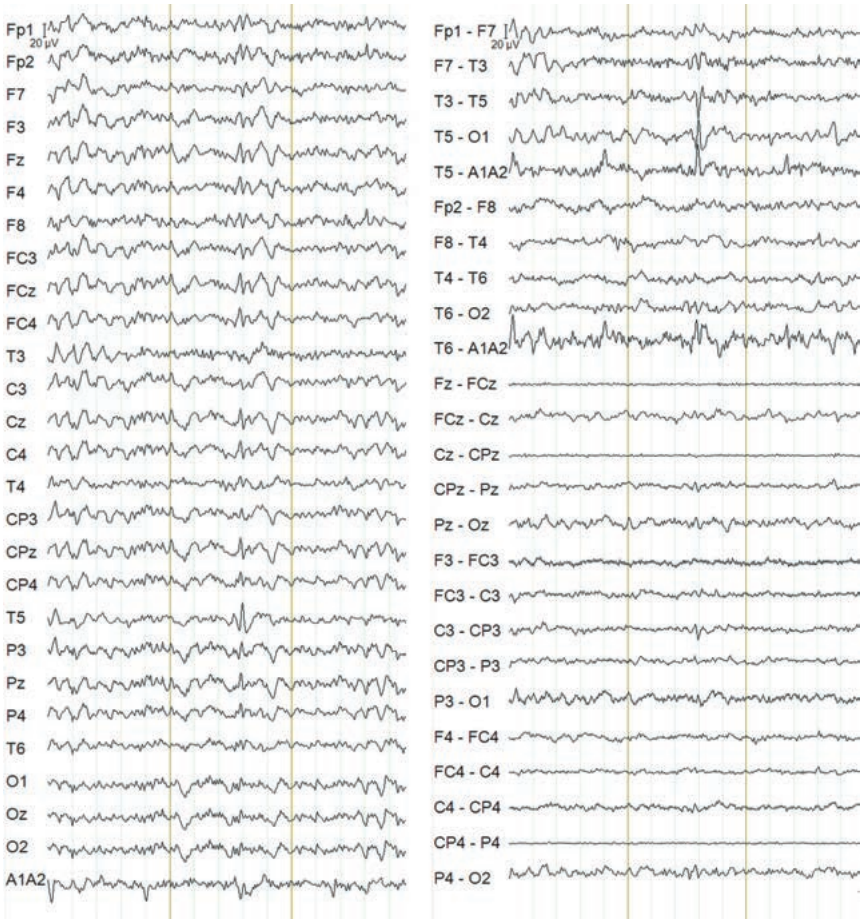


Figure S4.4: *Example of an EEG segment with sharp activity (a spike) over the left temporal region, displayed in two montages.*

Filter settings: 1-45 Hz.
Timescale: 1 s per vertical line.

5

COMPUTER ASSISTED EEG ABNORMALITY DETECTION AND TREATMENT RESPONSE PREDICTION IN MAJOR DEPRESSIVE DISORDER

N. van der Vinne*, M. C. Tjepkema-Cloostermans*, C. da Silva Lourenço, M. A. Vollebregt, M. Arns and M. J. A. M. van Putten
* *Shared first authorship*

ABSTRACT

Introduction

Patients with major depressive disorder (MDD) and abnormal EEG patterns were more likely to be non-responsive to the antidepressants escitalopram and venlafaxine in comparison to patients without EEG abnormalities. Response to sertraline was not different for patients with or without EEG abnormalities. Assessment of EEG is typically performed by visual inspection, characterized by low inter-rater reliability. We aim to utilize computed features for automated assessment of abnormal EEGs in the depressed population to assist in treatment prediction.

Methods and materials

EEG epochs (2 min, eyes closed) from depressed patients and healthy controls (n = 1241) were visually assessed for abnormal components (slowing, asymmetry, paroxysms). In addition, three quantitative features were extracted: the dominant frequency, the temporal and spatial brain symmetry index (tBSI and sBSI), complemented with the likelihood that the epoch contained an epileptiform discharge, using the output of a convolutional neural network (CNN). The four features were logistically regressed on EEG type (visually normal vs abnormal) and combined in a random forest model, trained on predicting treatment outcome per antidepressant drug (for the depressant sample only: n = 935).

Results

Outcomes from the CNN, the dominant frequency and the tBSI significantly differed between normal and abnormal EEGs (all ps < .001). Our attempts to replicate response prediction (using visual analysis) with the random forest model was not successful, with areas under the curve of 0.49-0.51.

Discussion

CNN probability, dominant frequency and tBSI successfully identified EEG abnormalities in agreement with visual analysis. The random forest model did not reliably predict treatment outcome.

5.1 INTRODUCTION

Arns and colleagues (2017) showed that patients with major depressive disorder (MDD) with abnormal EEG patterns are more likely to be non-responsive to the antidepressants (ADs) escitalopram and venlafaxine in comparison to patients without those EEG abnormalities. In the same study, it was shown that response to sertraline was not different for patients with or without EEG abnormalities. Abnormal activity included paroxysmal activity (e.g. isolated epileptiform discharges, intermittent focal slowing) or diffuse slowing of the background pattern. Furthermore, in a follow-up study, depression response with concurrent EEG normalization occurred more often after sertraline treatment than escitalopram or venlafaxine (van der Vinne et al., 2019). These findings suggest that the EEG can be used to guide patients to the right antidepressant.

Assessment of EEG is typically performed by visual analysis. A limitation of this is the need for human experts and low interrater reliability. As an example, six board-certified neurophysiologists achieved an agreement (Fleiss's kappa) of 55% when classifying 300 EEGs as normal or containing epileptiform discharges or seizures (Grant et al., 2014). It is likely that classifying less evident abnormalities are even more prone to a high inter- and intra-rater variability. This motivates exploration of more advanced forms of detection of abnormal EEG patterns using quantitative techniques.

Extracting (spatio-)temporal features of the EEG using e.g. Fourier transforms, wavelets or coherence is well established (Lodder & Van Putten, 2011; Van Putten 2008). More recently, machine learning and deep learning have been explored for detection of anomalies. This includes detection of interictal epileptiform discharges (Tjepkema-Cloostermans, de Carvalho, & van Putten, 2018; Van Leeuwen et al., 2019; da Silva Lourenço, Tjepkema-Cloostermans, Teixeira, & van Putten, 2020) or assessment of EEG patterns for prediction of neurological outcome in patients after cardiac arrest (Tjepkema-Cloostermans et al., 2019, 2017).

As the abnormal EEG patterns assessed by visual analysis in our previous work (Arns et al., 2017) contained both paroxysms and diffuse slowing, we combined two different quantitative techniques. For the paroxysmal abnormalities in the EEGs the output of a deep neural net was used (da Silva Lourenço et al., 2020), complemented with quantitative assessment of mean frequency, asymmetry (through the spatial Brain Symmetry Index) and intermittent slowing (through the temporal Brain Symmetry Index).
In this work, we aim to predict treatment response using these quantitative metrics. We further explore how this approach performs in comparison with earlier work described in Arns et al. (2017) using visual analysis.

5.2 METHODS AND MATERIALS

5.2.1 DESIGN

Data were extracted from the international Study for Predicting Optimized Treatment in Depression (iSPOT-D). This study was an international multicenter, randomized, prospective open-label trial (phase IV clinical trial), in which MDD subjects were randomized to escitalopram, sertraline, or venlafaxine-XR treatment in a 1:1:1 ratio. The study protocol details have been published by Saveanu et al. (2015) and Williams et al. (2011).
The iSPOT-D sample consisted of 1008 patients with MDD and 336 healthy controls, with 58 subjects having an abnormal EEG (identical to the sample from in Arns et al. (2017)). The same study provides a complete description of the study assessments, inclusion/exclusion criteria, diagnostic procedures, treatment and characterization of this paroxysmal subgroup. In summary, the primary diagnosis of nonpsychotic MDD was confirmed at the baseline visit (before randomization) using the Mini-International Neuropsychiatric Interview (MINI-Plus), according to the Diagnostic and Statistical Manual of Mental Disorders, fourth edition (DSM-IV) criteria, and a score ≥16 on the clinician-rated 17-item Hamilton Rating Scale for Depression (HRSD). All MDD subjects were either AD medication naive or, if previously prescribed an AD, had undergone a washout period of at least five half-lives before the baseline visit clinical and EEG assessments. After the baseline visit, MDD subjects were randomized to one of the three

AD medications. After eight weeks of treatment, subjects were tested again using the HRSD and EEG, further referred to as endpoint. Patients provided written informed consent. This study was approved by the local institutional review boards at all the participating sites and was registered at ClinicalTrials.gov (NCT00693849).

5.2.2 PRE AND POST-TREATMENT ASSESSMENTS

EEG recordings were performed at baseline and at endpoint, using a standardized methodology and platform (Brain Resource Ltd, Australia). Details of this procedure have been published elsewhere (Williams et al., 2011), and details of the reliability and across-site consistency of this EEG procedure have also been published (Paul et al., 2007; Williams et al., 2005). In summary, EEG data were acquired from 26 channels: Fp1, Fp2, F7, F3, Fz, F4, F8, FC3, FCz, FC4, T3, C3, Cz, C4, T4, CP3, CPz, CP4, T5, P3, Pz, P4, T6, O1, Oz, and O2 (NuAmps; 10–20 electrode international system). EEG data were collected for 2 min with eyes closed (EC). Data were referenced to averaged mastoids with a ground at AFz. Horizontal eye movements were recorded with electrodes placed 1.5 cm lateral to the outer canthus of each eye. Vertical eye movements were recorded with electrodes placed 3 mm above the middle of the left eyebrow and 1.5 cm below the middle of the left bottom eyelid. Skin resistance was <5 kOhm for all electrodes. A low-pass filter with an attenuation of 40 dB per decade above 100 Hz was employed prior to digitization with a sampling rate of 500 Hz.

5.2.3 EEG ANALYSIS

A high-pass filter of 0.5 Hz, a low-pass filter of 25 Hz, and notch filters of 50 or 60 Hz (depending on the country in which the data were recorded) were applied. Data were EOG corrected using a regression based technique similar to the method described by Gratton, Coles, and Donchin (1983). No other artifact rejection was applied to the data.

5.2.4 ORIGINAL VISUAL ASSESSMENT

For the distinction between normal and abnormal EEGs in Arns et al.

(2017), data were visually inspected and classified using Brain Vision Analyzer (Brainproducts, Germany) by a professional board-certified electroencephalographer, neurologist and psychiatrist, who was blinded to the subject's status (patient vs control), clinical data (response vs non-response to treatment), assessment moment (baseline vs endpoint) and treatment arm. Details on EEG data analysis and validation can be found elsewhere (Arns et al., 2016). Eyes closed awake EEG data with linked ears and queens-square montage were examined for the presence of any focal or generalized slowing (EEG slowing). Diffuse slowing was defined as a background frequency consistently below the alpha range (Niedermeyer, 2005). Focal slowing was defined as rhythms slower than alpha (theta or delta, i.e. <8 Hz) consistently present in a particular location (Krauss, Fisher, & Kaplan, 2010). Paroxysmal activity was defined as any EEG pattern (with or without a sharp contour) that emerges and disappears paroxysmally from the ongoing background activity (Niedermeyer, 2005). Non-paroxysmal, focal or generalized, slow wave activity was also defined, which could occur more or less continuously (note that records were almost entirely fully awake records) with some waxing and waning (Sharbrough, 2005). Finally, the presence of any of the so-called controversial waveforms (e.g., wicket spikes) was also documented. These waveforms are paroxysmal but are of uncertain significance (Boutros, 2014). Supplement 4.7 contains descriptive and visual information of abnormalities on the individual level. Classification of all abnormalities was in accordance to the guidelines published by the International Federation of Clinical Neurophysiology (Noachtar et al., 1999).

5.2.5 FEATURE EXTRACTION

EEGs of both depressed patients (n = 935) and healthy controls (n = 306) were exported to European Data Format (Kemp, Värri, Rosa, Nielsen, & Gade, 1992). Processing was done with in-house developed software on a MatLab platform (The MathWorks, Inc, Natick, MA), according to previously described methods (Tjepkema-Cloostermans et al., 2018; van Putten 2006a, 2007; van Putten et al., 2004a). All analyses were performed after re-referencing to the longitudinal bipolar montage.

5.2.5.1 DEEP LEARNING MODEL

For the detection of paroxysmal anomalies, we used an existing convolutional network that was previously trained for the detection of interictal epileptiform discharges in epilepsy patients. As described by da Silva Lourenço and colleagues (2020), this VGG C convolutional network was trained on interictal EEG recordings from patients with focal (50 patients) and generalized (49 patients) epilepsy. Normal EEGs (67 patients) were also included. The EEG data in the 0.5-35 Hz range were filtered and downsampled to 125 Hz. Each recording was split into 2 s non-overlapping epochs, yielding a 18×250 (channels × samples) matrix for each epoch. More details can be found in da Silva Lourenço et al. (2020). The model provides the probability p1 (range 0 – 1, with 0 being normal) for the presence of an epileptiform discharge per 2 s epoch. To rule out coincidental false findings, we chose to take the average of the probability p1 of the three most extreme epochs, instead of the one most extreme epoch.

5.2.5.2 DOMINANT FREQUENCY

Because the abnormalities in patients with MDD not only consist of short paroxysmal events, the dominant frequency was calculated as the largest peak in the power density spectrum, with a minimum peak height of 0.5 uV^2/Hz and a frequency between 5 and 15 Hz. Power spectrum was estimated using Welch's method, with half-overlapping 2 s windows. We extended the frequency range of the alpha peak frequency utilized in Arns et al. (2017) to prevent other dominant frequencies (especially below 8 Hz) to remain unnoticed.

5.2.5.3 SBSI AND TBSI

As the dominant frequency will not detect frequency fluctuations in the EEG, nor by construction, spatial asymmetries, two additional metrics were applied, the spatial and temporal brain symmetry index (BSI; Van Putten, 2007). The BSI was developed to assist in the visual interpretation of the EEG, by quantifying spatial (left-right hemisphere, sBSI) and temporal (tBSI) characteristics.

The revised spatial BSI (sBSI) is the absolute value of the relative difference in the average spectral density of the right and left hemispheres in the frequency range from 1 to 25 Hz.
The sBSI is defined as:

$$sBSI\,(t) = \frac{1}{K}\sum_{n=1}^{K} \left|\frac{R_n^*(t) - L_n^*(t)}{R_n^*(t) + L_n^*(t)}\right|$$

with

$$R_n^*(t) = \frac{1}{M}\sum_{ch=1}^{M} a_n^2(ch, t)$$

for the right hemisphere, and a similar expression for the left hemisphere. Here, a_n(ch,t) is the Fourier coefficient with index n of channel ch, evaluated at time t, corresponding to a particular epoch $[t - T, t]$ with duration T. In the current work we set T = 4 s. M is the number of channels per hemisphere and K is the number of Fourier coefficients. Additional details have been published previously (van Putten 2007; van Putten et al., 2004b). The sBSI is an index, ranging from 0 (perfect symmetry) to 1 (maximum asymmetry).

In order to quantify temporal changes in spectral characteristics of EEG, the temporal brain symmetry index (tBSI) was proposed and revised by Van Putten (2006). The revised tBSI is defined as the normalized difference between spectral estimates of two EEG epochs and thus provides a measure of temporal invariance or symmetry. The tBSI is defined as:

$$tBSI\,(t) = \sqrt{\left|(\Delta R\,(t) - \gamma) \cdot (\Delta L(t) - \gamma)\right|}$$

with

$$\Delta R\,(t) = \frac{1}{K}\sum_{n=1}^{K} \left|\frac{R_n^*(t) - R_n^*(t_o)}{R_n^*(t) + R_n^*(t_o)}\right|$$

for the right hemisphere, and a similar expression for the left hemisphere. Here, γ is an offset factor, and t_o a suitable reference. In this work we set $\gamma = 25$ and for t_o we chose a moving reference epoch of 8 s before the analyzed epoch (t). Similar to the sBSI, the tBSI ranges from 0 (no temporal spectral differences) to 1 (maximal temporal

spectral differences).
To avoid coincidental extreme scores, we averaged the three most extreme sBSI and tBSI segments for every subject.

5.2.6 STATISTICS

We compared the four predictors to the binary coded groups with normal/abnormal EEGs according to visual assessment as reported in Arns et al. (2017). To this end, we employed logistic regression for every single predictor on the binary EEG outcome. Normal distributions are not assumed in logistic regression; therefore, it poses no problem for the non-normally distributed predictors CNN probabilities, sBSI and tBSI. To correct for multiple testing, the alpha level was set to .025 (Bonferroni correction for four predictors).

5.2.7 TREATMENT PREDICTION

All individual features (CNN probability, dominant frequency, tBSI, sBSI) were included in a random forest model. Hereby, for each type of AD random forest classifiers were trained and evaluated using 5-fold cross-validation. Each random forest classifiers consisted of 500 individual decision trees, and the maximum number of terminal nodes was set to three. The output of the classifiers is the probability (ranging 0-1) for a patient responding to treatment (≥50% reduction on the HRSD). Random forest classification was done using the software package R (Liaw & Wiener, 2002; R Core Team 2014). Discrimination of the model was assessed with receiver operator characteristic (ROC) analyses.

5.3 RESULTS

5.3.1 SINGLE PREDICTORS OF ABNORMALITY

Logistic regression on the groups of visually classified normal and abnormal EEGs, revealed significantly different features outcomes (as measured by the mean of the three most extreme CNN probabilities per subject (figure 5.1a), the three epochs with the most extreme tBSI (figure 5.1b), and the dominant frequency (figure 5.1c). No differences

were found with the sBSI. Additional information on test outcomes and distribution of results can be found in supplement 5.5.

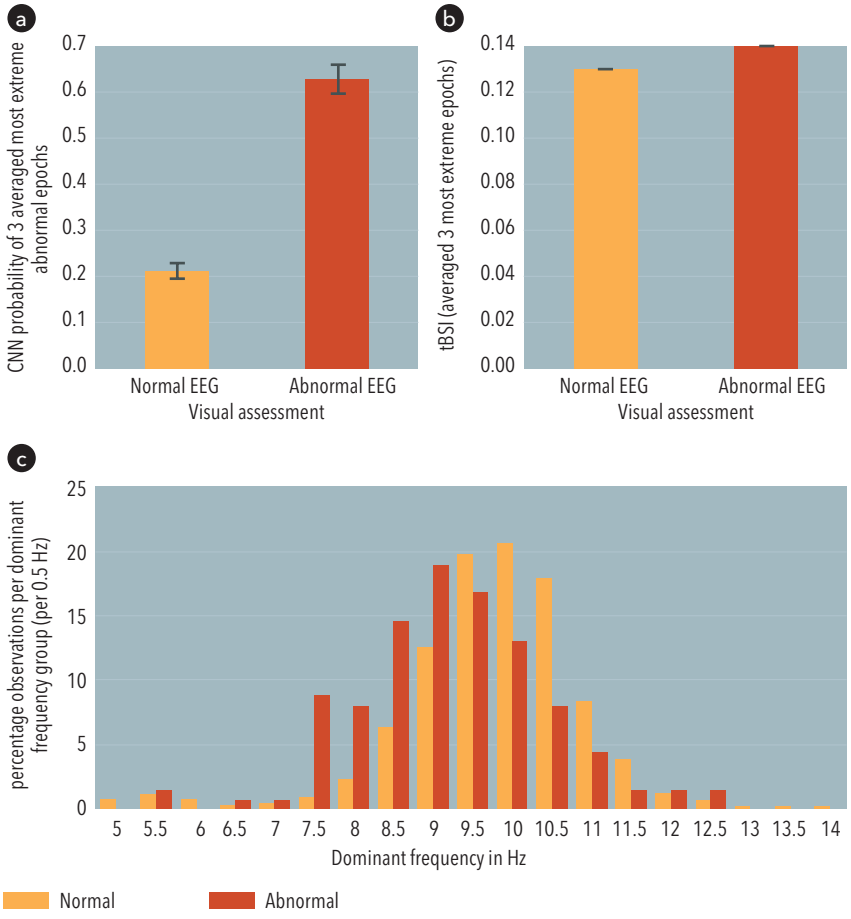


Figure 5.1: *Comparison of the individual features to visual analysis of EEGs conform Arns et al., 2017).* ***a:*** *means of convolutional neural network (CNN) probabilities of an epoch being abnormal (the three most extreme abnormal epochs; with SEM bars, representing 95% standard error of the mean).* ***b:*** *tBSI means (the three most extreme epochs; with SEM bars representing 95% standard error of the mean). Note that the significant difference is barely visible here.* ***c:*** *Histogram of the dominant frequency, per group (visually determined normal and abnormal EEGs), with weighted number of observations in percentage (due to large subsample size differences).*

5.3.2 TREATMENT PREDICTION

Training and validating random forest classifiers per medicament using 5-fold cross-validation resulted in ROC curves with a very poor AUC (ranging from 0.49 - 0.51), with wide confidence intervals (figure 5.2). This indicates that with the current method, in which the four predictors are combined into a random forest classifier, it is not possible to predict treatment response reliably.

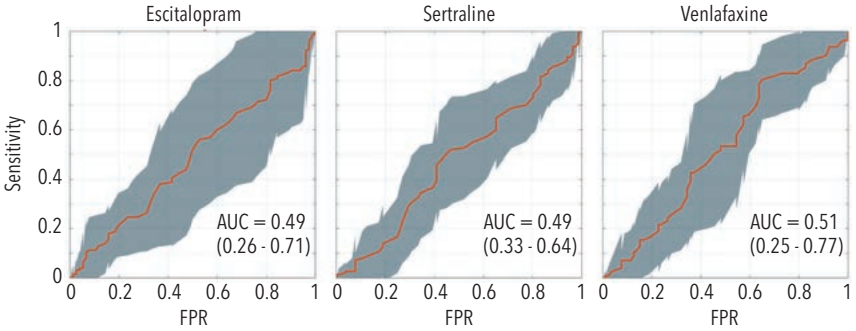


Figure 5.2: *ROC curves for the prediction of treatment response using random forest models. The red lines indicate the average ROC curve per antidepressant, while the grey zones indicate the 95% confidence intervals based on 5-fold cross validation.*

5.4 DISCUSSION

We aimed to predict treatment response using four quantitative metrics and explored how this approach performs in comparison with earlier work. Subjects with normal and abnormal EEGs assessed with visual analysis as reported in Arns et al. (2017) differed significantly on three out of four computed features for detecting abnormalities in our study. This confirms our hypothesis that the computed features can detect differences between the two groups that were visually identified earlier.

Interestingly, the most meaningful group difference was visible for the CNN, where group means of 0.63 (visually abnormal EEG) and 0.21 (visually normal) on a scale of 0 to 1 indicate a possibly useful clinical application. Less useful are the mean dominant frequencies in the groups of 9.2 (visually abnormal) and 9.8 (visually normal), as these mean frequencies are both generally interpreted as normal. The significant, but extremely small difference between the mean tBSIs renders clinical usage impossible.

We developed a new experimental method to create computer-assisted detection of abnormalities for the benefit of AD treatment prediction. The combination of the promising predictors in a random forest model did not result in the hypothesized treatment prediction, with AUCs indicating treatment success of practically 50%. Our method therefore unfortunately yielded no clinically relevant results up to now. As the importance of null-findings is increasingly emphasized in psychology and psychiatry – specifically for EEG biomarkers by Widge and colleagues (2019) – we report these results.

We note that the EEG abnormalities occurrence according to Arns et al. (2017; which data was used for the current study) concerned an estimated 5-10% of the population. This was sufficient for establishing an effect through statistics. Not only did they find the best treatment response for those with EEG abnormalities with the AD sertraline, normalization of the EEG in those taking sertraline was also associated with response (van der Vinne et al., 2019). However, our strongest extracted feature of detecting abnormalities, the CNN, typically needs a large sample in order to distinguish patterns. The lack of an effect might therefore be due to the limited number of abnormal EEG subsample, despite the large entire sample. As a result, the CNN might not be able to pick up patterns that held prognostic capabilities when assessed visually by a clinician (Arns et al; 2017; van der Vinne et al., 2019).

In contrast, in particular for machine learning algorithms, a small and unbalanced data set is known to limit performance (Bauder & Khoshgoftaar, 2018; Johnson & Khoshgoftaar, 2019). Furthermore, as Roy et al. (2019) state in a systematic review on deep learning, features learned through a deep net might be more effective or expressive. Therefore, different treatment prediction outcomes between Arns et al. (2017) and our current study might originate from the more sophisticated abnormalities detection by especially the CNN, picking up signals that are invisible to the human eye. As this contradicts our previous possible explanations, future studies are all the more important, clarifying how our visual methods could be better approached by computer-assisted methods.

As the deep net processes segments that were subjected only to EOG artifact rejection and filtering, the risk exists that non-neural artifacts influence the outcome. We should consider this having a con-

tribution to the null result for MDD response prediction. However, by trying to reduce artifacts through a narrow filtering range, and by reducing input size (and consequently computational complexity) through downsampling to 125 Hz, we expected this effect to be limited.

We were otherwise surprised that the CNN developed for the detection of interictal discharges showed good performance in identifying the specific and notably "light" abnormalities, in a different type of patient population. This deep net developed by da Silva Lourenço et al. (2020) initially seemed the best candidate for the detection of EEG abnormalities in our patient group, although its relevance is diminished with the established lack of capabilities in treatment outcome prediction. Our prediction efforts did not reach what the collaboration of Wu and colleagues did achieve (Wu et al., 2020). Their latent-space machine-learning algorithm trained on one dataset and validated on multiple independent datasets, identified a sertraline-predictive EEG signature, compared to placebo. Their method is in line with a future study suggestion of our own: not training an algorithm on biomarkers conform our current methods, but specifically on the EEG and response profiles themselves. This could be complemented by another method, the highly comparative time series (Fulcher et al., 2013). It allows to organize time-series datasets automatically, according to their properties. The authors demonstrated its utility on a dataset of EEGs, distinguishing between healthy and epileptic EEGs with a classification rate exceeding 95%.

After successfully extracting features for the detection of mild EEG abnormalities in a large depressed sample, our efforts to use these promising biomarkers – that are traditionally assessed visually – unfortunately did not facilitate treatment prediction. As we live in the era where deep learning algorithms are continuously being improved, this method could be explored further, paving the road towards optimized treatment prediction.

5.5 SUPPLEMENT: ADDITIONAL TEST RESULTS PAR. 5.3.1

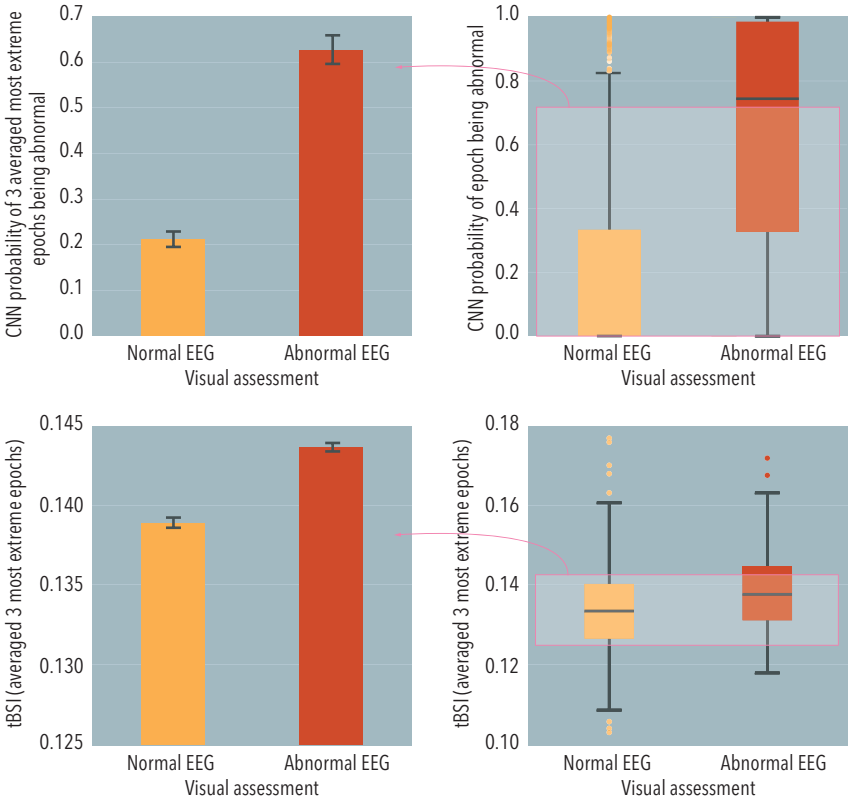
ADDITIONAL TEST RESULTS

Table S5.1 Test results of logistic regression of extracted features on the visual assessment of EEGs being normal or abnormal

		Group mean		Outcome logistic regression	
		Visually normal	Visually abnormal	B	*p*
CNN model	mean of 3 most extreme epochs	0.21	0.63	2.726	<.001
sBSI	mean of 3 most extreme epochs	0.127	0.129	0.473	.789
tBSI	mean of 3 most extreme epochs	0.134	0.139	45.011	<.001
Dominant frequency		9.751	9.184	-0.343	<.001

CNN – Convolutional Neural Network, sBSI: spatial Brain Symmetry Index, tBSI: temporal Brain Symmetry Index

COMPARISON OF EXTRACTED FEATURES TO VISUAL ANALYSIS – INCLUDING DISTRIBUTION PLOTS



Facing page:
Figure S5.1 (top): *Comparison of means (with SEM bars, left) and visualization of distribution through a boxplot (right) of convolutional neural network (CNN) probabilities of the averaged three most extreme abnormal epochs. Grouping conform visual analysis reveals significantly different CNN probabilities at the p<.001 level. Outliers are represented by dots (SD>2).*

Figure S5.2 (bottom): *Comparison of means (with SEM bars, left) and visualization of distribution through a boxplot (right) of the averaged three most extreme tBSIs, for the groups with visually determined normal and abnormal EEGs. Grouping conform visual analysis reveals significantly different tBSI means at the p<.001 level. Outliers are represented by dots (SD>2).*

6

EEG BIOMARKER INFORMED PRESCRIPTION OF ANTIDEPRESSANTS IN MDD: A FEASIBILITY TRIAL

N. van der Vinne, M. A. Vollebregt, A. J. Rush, M. Eebes, M. J. A. M. van Putten and M. Arns
Under review, European Neuropsychopharmacology

ABSTRACT

Introduction

Using pre-treatment biomarkers to guide patients to the preferred antidepressant medication treatment could be a promising approach to enhance its current modest response and remission rates. This open-label prospective study assessed the effectiveness and feasibility of using such pre-treatment biomarkers, by using previously identified EEG features (paroxysmal activity; alpha peak frequency; frontal alpha asymmetry) to inform the clinician in selecting among three different antidepressants (ADs; escitalopram, sertraline, venlafaxine) as compared to Treatment As Usual (TAU).

Methods and materials

EEG data were obtained from 195 outpatients with major depressive disorder prior to eight weeks of AD treatment. Primary outcome measure was the percentage change between before and after treatment on the Beck Depression Inventory-II. We compared TAU and EEG-informed prescription through AN(C)OVAs. Recruitment started with patients receiving TAU to establish baseline effectiveness, after which we recruited patients receiving EEG-informed prescription.

Results

108 patients received EEG-informed prescription and 87 patients received TAU. Clinicians and patients were satisfied with the protocol. Overall, 70 (65%) of the EEG-informed clinicians followed recommendations (compared to 52 (60%) following prescriptions in the TAU group), establishing feasibility. Depressive symptoms reduced significantly more in the EEG-informed (36.8%) compared to the TAU (23.9%) group (p = .024; between groups d = 0.42; 39% responders in EEG-informed vs 27% responders in TAU; number-needed-to-treat = 8). Intention-to-treat analysis including all patients (n = 195) yielded similar results.

Discussion

We here confirm that treatment allocation informed by EEG variables previously reported in correlational studies, was feasible and yielded significantly better clinical response than TAU.

6.1 INTRODUCTION

The treatment of major depressive disorder (MDD) is characterized by modest response and remission rates, while the disease affects increasing numbers of people worldwide, from 183 million in 2005, to 216 million in 2015 (Vos et al., 2016). Clinical efficacy ranges from 37% remission after a first antidepressant (AD) prescription to declining remission rates of respectively 31%, 14%, and 13%, after each consecutive AD trial, including augmentation strategies (Rush et al., 2006).

One way to improve response and remission rates in the early AD treatment steps for major depressive disorder (MDD) is to better target the medications to particular patients. In that regard, identification of pre-treatment biomarkers which can inform choices between or among treatments offers a promising approach, although a need for replication and out of sample validation have been suggested (Widge et al., 2019).

To develop such biomarkers, we initially collected pre-treatment EEG data, in the International Study to Predict Optimized Treatment (iSPOT-D; Saveanu et al., 2015; Williams et al., 2011). The initial phase of the study randomized 1008 patients with non-psychotic MDD to eight weeks of treatment with either escitalopram, sertraline or venlafaxine-XR. Overall response (64%) and remission (46%) rates did not distinguish among these three medication groups, indicating comparable clinical efficacy on the group-level based on randomized treatment allocation.

Furthermore, several EEG parameters were investigated as predictors for response and remission (using pre-registered hypotheses). Three promising biomarkers emerged that seemed to inform which patients are preferentially served by which AD medication, as both drug-specific as well as drug-class specific predictors, opening up the

possibility for EEG-guided treatment (or stratification). The first was frontal alpha asymmetry (FAA): Right FAA was found to be related to response and remission (and left FAA to non-response and non-remission) to the SSRIs escitalopram and sertraline in females only. No such effect was observed for the SNRI venlafaxine (Arns et al., 2016). A post-hoc simulation showed that assigning patients to an SSRI or SNRI merely based on their FAA, resulted in a 7-14% higher remission rate (Arns et al., 2016). The second biomarker was alpha peak frequency (APF): A low APF was associated with better response to sertraline and no effects for escitalopram and venlafaxine (Arns, Gordon, & Boutros, 2017). The third biomarker was abnormal EEG activity: abnormalities like isolated epileptiform discharges (IEDs) were associated with non-response to escitalopram and venlafaxine, and no such effect for sertraline (Arns, Gordon, & Boutros, 2017). In addition, EEG normalization after eight weeks on sertraline mediated AD response, suggesting sertraline specifically worked on the reported EEG abnormalities (van der Vinne et al., 2019a).

In summary, our prior work revealed drug-specific (sertraline), drug class-specific (selective serotonin reuptake inhibitor (SSRI) vs serotonin norepinephrine reuptake inhibitor (SNRI)) and sex-specific EEG parameters that could aid in resolving the heterogeneity in clinical response to ADs, and that could be used to choosing among ADs, going from a stepped-care approach to a biomarker-informed approach.

In order to determine whether medication could be prescribed based on these baseline EEG biomarkers, we conducted a prospective feasibility trial, in which we compared EEG-informed treatment recommendation with Treatment As Usual (TAU). We report our findings in developing and implementing this clinical decision-making tool and providing power calculations to inform future clinical trials. Our null hypothesis was that clinical response is not worse for EEG-informed treatment allocation, compared to TAU. If, alternatively, EEG-informed prescription was better, we expected group differences to be small, given the comparison of two active treatments.

6.2 METHODS AND MATERIALS

6.2.1 DESIGN

This was an open-label, naturalistic study, which was deliberately chosen to mimic real-world practice, with the aim of optimizing the translatability to real world settings. We investigated 195 outpatients with non-psychotic MDD, recruited between June 2015 and July 2019 in an outpatient clinic in Leeuwarden, the Netherlands. The primary diagnosis of nonpsychotic MDD was confirmed by a psychiatrist or specialized clinical psychologist, according to DSM-IV criteria, and a score ≥14 on the Dutch 21-item Beck Depression Inventory Second Edition (BDI-II-NL, Beck, Steer, & Brown, 1996; Van der Does, 2002). Only data from patients who were prescribed with ADs were included. All MDD patients were allowed to enter the study when already on an AD (since Van der Vinne et al. (2019b) demonstrated that the predictive value of FAA was not influenced by medication status).
To establish baseline effectiveness, recruitment started with MDD patients who had not received EEG-informed prescription, and received TAU. Once the biomarker algorithm based on our prior work was finished, we began to recruit patients for the second arm – those who were to receive EEG-informed prescription (see paragraph 6.2.2). After eight weeks of AD treatment, all patients were tested again using the BDI. Both EEG assessments and week-8 measurements were completed at a priori defined dates. Patients who did not complete eight weeks of treatment within the defined period, were excluded from analyses. Both dropouts and different recruitment periods accounted for different subsamples. As part of our feasibility report, we discuss reasons for dropout extensively in paragraph 6.3.1 and supplement 6.5.

6.2.2 EEG-INFORMED PROTOCOL VS TAU

Patients had to meet a DSM-5 classification for non-psychotic depression, and a BDI-score≥14. For those receiving EEG-informed prescription (for a full decision tree also see figure 6.1, page 123), the EEG outcome was shared with the designated nurse practitioner or psychiatrist. Together with the patient, it was decided whether the advice was to be followed. The clinician's decision on whether to follow the advice was leading, and the EEG recommendation was not binding. Without in-

formed consent, the advice was still offered, but data were not recorded for scientific purposes nor included in this report.
In case of TAU, AD choice resulted from the prescriptive decisions of a psychiatrist or nurse practitioner. In case of a first depressive episode, the prescription of an AD was based on national guidelines for prescribing first-choice ADs. In case of a recurrent depression, applicable national guidelines were followed. Where available, information on earlier (un-)successful drug treatments of either the patient or first-degree relatives was considered.
The patient signed an informed consent and data were coded anonymously. Treatment and progress were monitored. BDIs were filled out at fixed times: at intake, prior to each form of medical treatment, and eight weeks after each started medication. Registration of clinical patient data consisted of the following variables: previous treatments, current and/or previous diagnoses/classifications, EEG (advice) outcome, medication during EEG, following or ignoring the advice, which AD was prescribed, which psychological treatment was followed, and all BDI measurements.

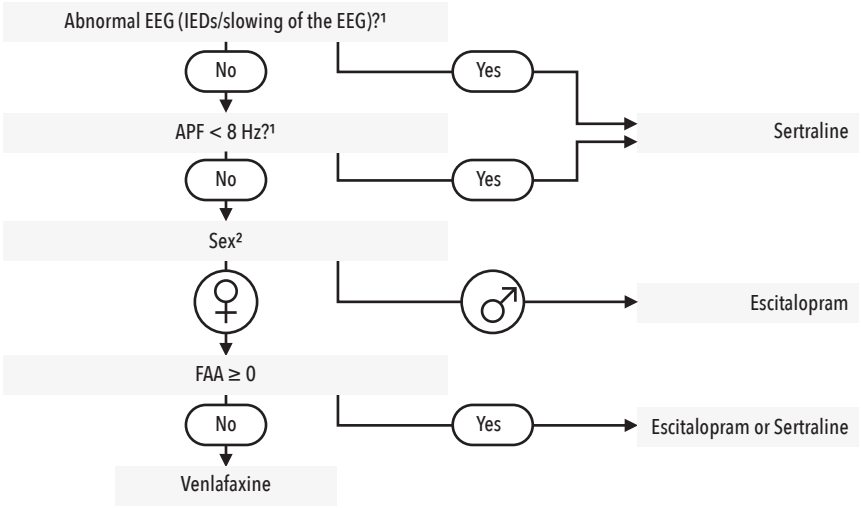
6.2.3 PRE-TREATMENT ASSESSMENTS

EEG recordings were performed using a standardized methodology and platform (Brain Resource Ltd., Australia). Details of this procedure (Arns et al., 2016) and of its reliability have been published elsewhere (Paul et al., 2007; Williams et al., 2005). In summary, patients were seated in a sound and light attenuated room. EEG data were acquired from 26 channels: Fp1, Fp2, F7, F3, Fz, F4, F8, FC3, FCz, FC4, T3, C3, Cz, C4, T4, CP3, CPz, CP4, T5, P3, Pz, P4, T6, O1, Oz and O2 (ANT Waveguard-cap; NuAmps; 10-20 electrode international system). EEG was assessed for two minutes with eyes open (EO, with the patient asked to fixate on a red dot on the screen) and two minutes with eyes closed (EC). The patient was instructed to remain relaxed for the duration of the recording. The operator did not intervene when drowsiness patterns were observed in the EEG. Data were referenced to averaged mastoids with a ground at AFz. Horizontal eye movements were recorded with electrodes placed 1.5 cm lateral to the outer canthus of each eye. Vertical eye movements were recorded with electrodes placed 3 mm above the middle of the left eyebrow

and 1.5 cm below the middle of the left bottom eyelid. Skin resistance was <10K Ohms for all electrodes. The sampling rate of all channels was 500 Hz. A low pass filter with an attenuation of 40 dB per decade above 100 Hz was employed prior to digitization.

6.2.4 EEG PROCESSING

EEG processing was conducted similarly to Arns et al. (2016), and performed by Brain Resource. The screening for subclinical EEG abnormalities was performed by trained psychologists (NvdV and MA), according to methods described in Arns et al. (2017). When an EEG provided inconclusive patterns or events, a board-certified neurologist/clinical neurophysiologist was consulted (MvP). EEGs of patients with TAU were not analyzed.
To determine EEG-informed prescription to AD, we employed the algorithm displayed in figure 6.1.



[1] *Arns, M., Gordon, E., & Boutros, N. N. (2017)*
[2] *Arns, M., Bruder, G., Hegerl, U., Spooner, C., Palmer, D. M., Etkin, A., ...Gordon, E. (2016)*

Figure 6.1: *Decision tree used for EEG-informed treatment allocation and clinical decision making. In the first step, sertraline was advised when a patient displayed subclinical abnormal EEG activity (IEDs or slowing of the EEG), or an alpha peak frequency (APF) below 8 Hz. For the second step, males were advised to start escitalopram. For females, escitalopram or sertraline were advised for a right-sided FAA, and venlafaxine was advised when a left-sided FAA was observed.*

6.2.5 STATISTICS

Clinical response was determined after eight weeks of treatment with the improvement as assessed by the BDI as primary outcome measure. The %BDI change from baseline to week-8 was chosen for analyses due to its normal distribution (as opposed to the non-normally distributed absolute difference). Remission was defined as a BDI score ≤12 and response was defined as a ≥50% reduction after eight weeks. Differences in age, sex, depression severity at baseline, and whether psychotherapy supplemented medication, were tested using one-way ANOVA or non-parametric tests, depending on distributions or variable type (continuous or binomial). When group differences were significant, the respective variable was added to the main analysis as covariate or fixed factor, depending on its nature.

The main analysis consisted of a univariate ANOVA, for investigating group differences in %BDI change. Significant effects were complemented with Cohen's *d* effect size and Number Needed to Treat (NNT, http://www.clincalc.com), based on %BDI change. In addition, odds ratios (ORs) were calculated for group differences in remission and response rates. Given this feasibility trial was expected to have insufficient statistical power to distinguish the groups and we had clear a priori expectations of the direction of effects (Ruxton & Neuhäuser, 2010), one tailed statistics were used to prevent overlooking relevant findings that require further investigation in future trials. Statistical significance was therefore set at $p < .10$. Power analysis to inform future trials were performed to derive minimum sample sizes for adequate study power (http://www.clincalc.com).

6.3 RESULTS

A total of 122 MDD patients who completed both baseline and week-8 assessment and met inclusion criteria, were included in our analyses, 70 with EEG-informed prescription and 52 controls with TAU. Dropout reasons and attrition are reported in great detail in supplement 6.5. Table 6.1 shows demographic information, and response and remission rates for included patients. There were no differences between the two treatment groups regarding age (Mann-Whitney U), sex and concurrent psychotherapy ($\chi 2$). Patients with TAU had

significantly higher BDI scores at baseline than those with EEG-informed prescription (Mann-Whitney U, $p = .048$).

Table 6.1 Demographic features of the two patient groups

	TAU	EEG-informed	Total
n	52	70	122
Females *n* (%)	25 (48%)	44 (62%)	69 (57%)
Average age (years M(SD))	37.2 (14.54)	40.3 (14.85)	38.9 (14.74)
BDI baseline (M(SD))	35.4 (9.61)	31.7 (10.56)	33.3 (10.28)
Supplemented psychotherapy	42 (71%)	50 (71%)	82 (71%)

6.3.1 FEASIBILITY

For an extensive qualitative analysis of the study's feasibility, we refer to supplement 6.5. In general, both clinicians and patients were satisfied with the new protocol. The practical implementation in the regular logistics of this Dutch mental health outpatient clinic proved to be feasible. Out of 195 initially included patients, we were able to analyze 122 (70 EEG-informed, 52 TAU). Attrition was mostly due to patients not filling out questionnaires at week 8 (10%), and patients stopping taking the AD (6%), which was not specific to the EEG-informed group. Other specific attrition reasons for the EEG-*informed* group were choosing a different AD due to earlier experience with the advised AD (4%), wanting to remain with their current AD, which was not the advised AD (3%), and the clinician choosing a different AD (5%).

6.3.2 TREATMENT OUTCOME

The EEG-informed prescription group demonstrated significantly better response (%BDI change) relative to the TAU group ($F(1,120) = 5.235$, $p = .024$), with a small to medium effect size of $d = 0.42$ and NNT of 8. A sensitivity analysis with the addition of covariate baseline BDI score to the model, yielded the same significant effect ($F(1,119) = 4.612$, $p = .034$) with the covariate itself being non-significant ($F(1,119) = 0.321$, $p = .572$). Treatment outcome can be found in table 6.2 on page 126. Interestingly, in the EEG-*informed* group the representation of the three ADs was quite comparable on the group

level (n = 22-24 per AD) and clinical benefit per individual drug was also quite comparable (range: 35.1-38.5% BDI change, not significantly different between drug groups) similar to the iSPOT-D study. Note that all patients in the EEG-informed group either had no earlier treatment failure experience with the advised AD, or did not receive the adviced AD before.

Table 6.2 Treatment outcomes for the two patient groups

	TAU	EEG-informed	Total
n	52	70	122
BDI-II baseline to week 8	35.4 – 27.1	31.7 – 20.2	33.3 – 23.1
% BDI-II change	23.9%*	36.8%*	31.3%
Escitalopram (n = 24)		*35.1%*	
Sertraline (n = 24)		*38.5%*	
Venlafaxine (n = 22)		*36.7%*	
Remission	17%	29%	34%
Response	27%	39%	24%

$*p < 0.05$

Intention To Treat analysis

With large dropout rates in both groups, we performed an intention to treat (ITT) analysis (n = 195). All dropout patients were assigned their baseline BDI value and thus assigned scores of 0 on %BDI change, response and remission. Repeating the treatment outcome analyses yielded similar group difference for %BDI change ($F(1,193) = 3.726$, $p = .022$).

6.3.3 FUTURE SAMPLE SIZE CALCULATION

Power calculation showed that with an alpha set on 5%, a power of 85%, and an enrollment ratio of 1:1.2 (TAU:EEG-informed, based on current subsample sizes) a total of at least 213 participants would be required to have sufficient statistical power in a future study (97 patients with TAU and 116 patients with EEG-informed prescription). This was based on the primary outcome measure, percentage improvement on the BDI.

6.4 DISCUSSION

We sought to prospectively test three previously identified EEG biomarkers using a feasibility trial. The implementation of the EEG-informed medication prescription algorithm proved feasible. The algorithm was sufficiently practical and satisfactory. The involved professionals were motivated to follow instructions and advice. It led to EEG-based prescription in 99 of 108 patients who received an EEG advice (89%), of which treatment outcome was known in 70 patients (65%). Major reasons for attrition were not specific to the EEG-informed algorithm, but related to patients not filling out questionnaires at week 8 (10%) or patients stopping taking the AD (6%). Most prominent attrition rates specific to the EEG-informed procedure comprised 12 subjects (11%) who either chose a different AD due to earlier experience with the advised AD, wanted to remain with their current AD which was not the advised AD, or for whom the clinician chose a different AD.

EEG-informed prescription resulted in significantly improved effectiveness with a small to medium effect size (d = 0.42), with response rates increasing from 27% to 39%, remission rates increasing from 17% to 29%, and an NNT of 8.

The TAU group presented with significantly higher baseline BDI scores. However, this is unlikely to change the above conclusion, since baseline BDI did not interact with the percentage BDI change. Furthermore, if higher baseline severity would have any effect, more room for improvement is expected with a regression to the mean in the TAU group (Mora, Nestoriuc, & Rief, 2011). A lack of this outcome strengthens our results for the EEG-informed group. It is highly unlikely that the group difference at baseline affected our main outcome.

To acquire sufficient statistical power for determining meaningful group differences in future controlled trials, we have established that it is feasible to recruit patients in new trials. Power calculation indicated that future trials using clinical improvement on the BDI over an eight week period of treatment as their primary endpoint, would require 218 patients. These numbers refer to patients with analyzable data, given the attrition rate of 35%. Attrition rates in future tri-

als will presumably be lower, given the naturalistic character of our study, and the average of 30% in a large review of AD trials (Woolley, Cardoni, & Goethe, 2009). Assuming an attrition rate of 30% suggests a future trial sample of 311 patients.

This first evidence from a prospective replication study, provides strong support for a future where EEG biomarkers can be used to inform clinical decision making, in this case treatment selection.

Recently, this particular application of EEG biomarkers was questioned, for publication bias, lack of proper replication studies, and out of sample validation (Widge et al., 2019). The present results provide the first prospective test of a priori defined EEG biomarkers, and thus provide strong evidence for the clinical use of these biomarkers, which should be confirmed in future controlled randomized controlled studies. In addition to the presence of EEG abnormality favoring response to sertraline (Arns et al., 2017), we previously have demonstrated that the degree of EEG normalization after eight weeks of treatment with sertraline, mediated clinical response (Van der Vinne et al., 2019a). These results suggest that sertraline impacts mechanistically on this EEG abnormality, i.e., by possibly having mild anticonvulsant properties. From a neurochemical perspective, it has been reported that sertraline, relative to other SSRIs, has the most pronounced dopamine transporter (DAT) inhibitory activity (Sanchez et al., 2014). Kanekar and colleagues also suggested a differential working mechanism in sertraline compared to other ADs, considering the increased inhibition of DAT by sertraline (Kanekar et al., 2018). However, further studies should investigate the exact neurobiological underpinnings further.

For the neurobiological underpinnings of FAA, the question is which role it has in the functional networks involved in depression, especially as a reflection of deeper nodes such as the subgenual anterior cingulate and other limbic structures. Given that FAA was associated to blood-oxygen-level-dependent (BOLD) activity in the left amygdala and emotion regulation (Zotev et al., 2016), FAA might reflect fronto-amygdala network activity. Previously reported amygdala AD response patterns (Carceller et al., 2018, in mice; Sheline et al., 2001,

in humans) as well as sex differences within the amygdala (Douillard-Guilloux et al., 2017), could help explain the sex-specific ability of FAA in predicting treatment outcome. The evidence however, remains indirect. The lack of multimodal studies that provide insight into causal relationships currently limits our ability to explain the relevant underlying neural circuitry behind the FAA-SSRI-response association.

An important limitation of this study, is the lack of random assignment to the groups, and possible non-specific and/or expectancy effects. Allocated patients were aware that the prescribed medication was based on their EEG, therefore expectancy effects might have played a role. On the other hand, the results derived from this naturalistic set-up are more likely to generalize to clinical practice, enabling us to investigate effectiveness (opposed to placebo-controlled efficacy studies where patients receive financial incentives to participate in studies, potentially boosting efficacy rates). Future studies should employ a randomized design where patients are unaware of how their assignment to treatment was done (e.g. TAU vs EEG-informed). Furthermore, longer follow-up periods are needed since non-specific (placebo) effects are usually brief. Results are nevertheless promising, given our preliminary effect size of 0.4, compared to placebo controlled effect sizes of ADs of 0.3 (Cipriani et al., 2018). Further investigation in prospective clinical trials of this EEG-informed medication prescription is warranted.

CONCLUSIONS

To the best of our knowledge, this was the first prospective EEG biomarker-based allocation feasibility study in MDD using a priori defined EEG Biomarkers. Our proposed protocol proved to be feasible, with more symptom improvement in patients allocated to ADs based on specific EEG biomarkers (FAA, APF, and paroxysmal activity). This improvement approached a medium effect size, despite the comparison of different patient groups taking the same AD treatments. Hence, it is very promising to follow up with clinical trials, on the road to personalized treatment, that might be preceded by EEG-informed treatment allocation as an intermediate step.

6.5 SUPPLEMENT: FEASIBILITY

In general, both clinicians and EEG-informed allocated patients were satisfied with the new protocol. They understood that actual improvement was unknown beforehand. The practical implementation in the logistics of this Dutch mental health outpatient clinic proved to be feasible, with an active role of the researchers in terms of planning and monitoring. With two trainings and annual update meetings, clinicians felt sufficiently comfortable with conveying the EEG advice to patients. When patients had questions, they were referred to the researchers (NvdV and ME). One nurse practitioner mentioned that this protocol particularly fits first-time prescriptions. Another indicated a feeling of lack of control, not making decisions as a clinician, although this was not perceived as very limiting.

The time to EEG outcome was experienced differently between clinicians, ranging from satisfyingly quick to very long. In the latter case, with severely depressed patients needing acute help, the clinician chose to skip the EEG and prescribe instantly. These were exceptional situations. With longer institutional waiting lists, patients could start with EEG-informed prescribed ADs similarly to patients with TAU. Theoretically, in times of shorter waiting lists, prescribing TAU would take place approximately 10 days earlier than following the protocol (which did not occur during this study, with current waiting lists).

One patient discontinued treatment due to severe side effects, after the clinician mistook instructions and quickly built up to a noradrenergic dosage of venlafaxine. We would like to stress that EEG-informed prescription of venlafaxine should not necessarily entail a noradrenergic dosage. Guidelines to venlafaxine prescription can be followed instead.

S6.5.1 ATTRITION

Detailed information is presented in the study sample overview in figure S6.1 on page 132. Numbering in the figure refers to the superscripted numbering in the following text. Out of eighty-seven patients planned to receive TAU, 80% actually started treatment and

60% completed treatment and measurements, compared to 89% of 108 patients receiving EEG-informed prescription, and 65% completing treatment and measurements. 18% of patients with TAU[1] and 3% of EEG-informed[7] chose to remain with their current AD, they wanted to try increasing dosage first. 4% of EEG-informed chose a different AD because to prior experiences[7].

Several patients dropped out before follow-up measurements could be executed. Twelve patients stopped taking their AD (6% of TAU[2], 7% of EEG-informed[8]). All five with TAU and three patients with EEG-informed prescription stopped due to side effects. One patient with EEG-informed prescription stopped after sleeping improved, which was sufficient for him. Of two patients, reasons are unknown. Nine patients went out of treatment in this clinic, it was unknown whether they were still taking ADs (6% of TAU[3], 4% of EEG-informed[9]). Nineteen did not fill out the questionnaire (8% of TAU[4], 11% of EEG-informed[10]). Four questionnaires were mistakenly not planned by the research team (1% of TAU[6], 4% of EEG-informed[13]).

Clinician's decisions lead to attrition as well. In 1% of the TAU group, the clinician chose to stop the AD the patient was already taking, because of side effects[5]. For EEG-informed patients, 3% of the clinicians chose to remain with the AD the patient was already taking[11]. Reasons for this were: wanting to try increasing dosage first and wanting to wait whether the small improvement with current AD would increase. For 2% of these patients, the clinician chose a different AD[12]. In one case, the prescription of an SSRI was deemed as not an option anymore. The other patient had a slow metabolizer for venlafaxine.

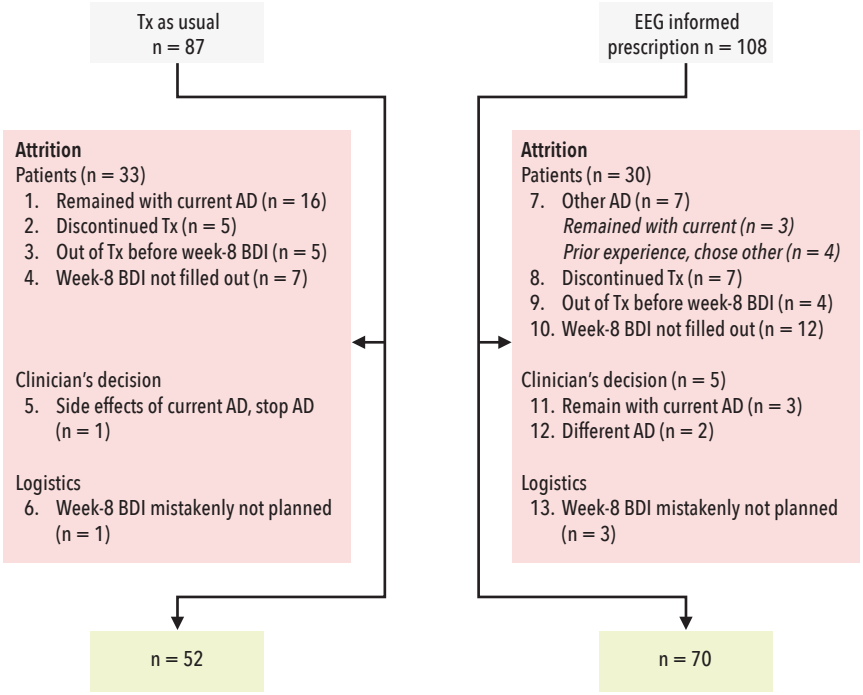


Figure S6.1: *Study sample overview of inclusion of patients.*
1-13 numbering refers to the superscripts in the paragraphs describing figure results.

ACKNOWLEDGEMENTS

We hereby acknowledge the support of and cooperation with Synaeda Psycho Medisch Centrum, in particular Anke Oosterbaan, Bianca Klop, Bob Goeree and Vera Veerman, and all clinicians who supported in the study; Bette Oterdoom, Lieuwe Krol, Dennis Bijlsma, Jenny Brouwer, Margreet de Jager, Marianne Wip-van Dijken and Jannie Bron. We thank Gonnie Toxopeus and Lize van der Meer for executing the first assessments, and all therapists and support staff for their efforts in including patients.
We also acknowledge supportive feedback from Guido van Wingen and Sebastian Olbrich at an earlier stage of the manuscript.

7

GENERAL DISCUSSION

In this dissertation, we described our first steps towards the implementation of an EEG biomarker informed protocol, with the aim to improve treatment outcome in depressed patients. Focusing on differences in treatment response enables us to reduce heterogeneity, which characterizes the targeted population. By establishing biomarkers related to treatment response in particular subgroups within the heterogeneous MDD group, we aim to improve treatment outcome.

In this last chapter, I will present our results in context of both older and current literature. I will discuss three EEG biomarkers in terms of: state and trait effects, underlying neurobiological explanations, their role in our successful prospective biomarker trial, and automated detection of the specific biomarker abnormalities in the EEG.

7.1 UNDERSTANDING BIOMARKERS

Frontal alpha asymmetry (FAA) has a long history in emotion research, with ensuing studies in depression. Over three decades ago, pioneering work started in FAA. The diathesis model by Davidson and Tomarken in 1989 (described in Henriques & Davidson, 1991; see also Davidson, 1998) linked a relative excess of alpha activity on the left side (compared to right) to more negative affect and withdrawal related emotion. Such left-sidedness has often been found in depressed patients throughout the years (Beeney et al., 2014; Bell et al., 1998; Debener et al., 2000; Fingelkurts et al., 2006; Gollan et al., 2014; Gotlib et al., 1998; Harmon-Jones & Allen, 1997; Henriques & Davidson, 1991; Jaworska et al., 2012; Kemp et al., 2010; Pizzagalli et al., 2002; Schaffer et al., 1983). In line with these prior studies, more recent research was focused on FAA as discriminant biomarker for (risk of) MDD and non-MDD. Given that studies since approximately the year 2000 became less conclusive about this distinction, we wanted to find out:

IS THE DEPRESSED POPULATION BEST CHARACTERIZED BY HOMOGENEITY OR HETEROGENEITY?

In the meta-analysis of **chapter 2**, we investigated the diagnostic value of frontal alpha asymmetry (FAA) for MDD. As discussed above, this relationship was demonstrated in several studies. Given some more recent studies (especially the largest study to date, iSPOT-D) could not corroborate this, we decided to conduct a systematic review and meta-analysis. In this meta-analysis, we found a small and non-significant mean effect size of the discriminant property of FAA, accompanied by highly significant heterogeneity across studies. This finding suggests that FAA is not a reliable diagnostic biomarker for MDD. Despite the characterization of FAA as the most clearly established marker of depression vulnerability (Reznik & Allen, 2018), the number of studies directly linking FAA to MDD is declining, as more recent studies tend to focus more on the relationship of FAA with emotion and affect, instead of MDD specifically.

Furthermore, the meta-analysis indicated that earlier study findings were most likely driven by underpowered studies. Our meta-analysis demonstrated that only studies with a sample of >300 subjects were consistent in their results. Results of these studies were characterized by a lack of a group difference. Furthermore, cross-sectional findings confirmed our hypothesis that a high level of heterogeneity in FAA in the depression population exists, which is in line with previous methodologically sound studies (Deldin & Chiu, 2005; Kentgen et al., 2000; Knott et al., 2001; Price et al., 2008; Quraan et al., 2014).
We cross-sectionally identified subgroups of severely depressed patients of 53 years and older, where males and females show opposite outcomes on FAA (comparable to Jesulola et al., 2017). This finding indicates that when studies are underpowered with overrepresentation of these older subgroups from either sex, the results can become biased. Although we split the MDD group based on statistical outcome, the resulting sample size of the subsamples was small, which is why this finding requires further replication.

In conclusion, with respect to FAA, the depressed population is best characterized by heterogeneity on the group level. These new insights in heterogeneity suggest a different application of FAA, actu-

ally utilizing the interindividual variation in this biomarker. This can potentially lead to optimized characterization of relevant homogenic subgroups and subsequent implications for a personalized treatment for the increasing number of depressed patients. Stratification based on these subgroups could eventually allow for precision psychiatry, which could replace the current expensive stepped-care model.

WHAT CAN WE LEARN FROM STATE, TRAIT, AND DRUG EFFECTS ON THESE BIOMARKERS?

For a protocol that relies on prognostic biomarkers, the prognostic value is not the only essential information we need to confidently employ this new method. Besides knowing which characteristics can separate responders from non-responders, we should have knowledge on the influence of the moment of measurement of these biomarkers. What can we learn from state, trait, and drug effects on these biomarkers? We addressed this in **chapters 3 and 4**.

We studied the stability of FAA in MDD patients during antidepressant treatment in **chapter 3**. FAA did not change significantly after eight weeks of escitalopram, sertraline, or venlafaxine treatment. FAA is therefore concluded to be a stable trait, not influenced by state and drug effects. This conclusion is in line with most of the literature (Allen et al., 2004; Bares et al., 2019; Bruder et al., 2008; Davidson et al., 2003; Deldin & Chiu, 2005; Gollan et al., 2014; Keune et al., 2011; Spronk et al., 2008; Sutton & Davidson, 1997; Tomarken et al., 1992), except for one study (Debener et al., 2000). Stability makes it suitable for sex specific treatment prediction under challenging circumstances such as state, time, the use of common antidepressive agents, and drug changes.

In **chapter 4**, we aimed to extend our earlier observations that depressed patients with a subclinically abnormal EEG at baseline are more likely to be a non-responder to venlafaxine or escitalopram. Abnormalities could consist of spike-wave discharges, diffuse slowing, or a slow background pattern operationalized using the frequency at which alpha waves are peaking (APF). This seems to suggest that sertraline is the preferred treatment in this subgroup (Arns et al., 2017).

We investigated whether abnormalities diminished after eight weeks of AD treatment, and how this related to response. EEG normalization patterns did not significantly differ between sertraline and other AD treatments. However, when comparing sertraline treatment to the other ADs within the subjects *who had a normalized EEG at endpoint*, response was more likely to be achieved with sertraline relative to the other treatments with a likelihood of 5.4 times. Hence, a subgroup was identified in which abnormalities were present prior to treatment, but disappeared afterwards, and showed a better response to sertraline than other AD treatments. The association with response indicates that we hereby identified a mediating biomarker (see Reznik and Allen (2018) for a discussion on mediating biomarkers).

New literature on the different antidepressant profiles has barely clarified how ADs might have consequential effects on the (abnormal) human EEG. Although the generalization of our results on EEG abnormalities to epileptogenesis is uncertain, it might help explain in part how depressed patients with EEG abnormalities are best treated with sertraline. But further research is necessary to understand the mechanisms that may be underlying the differential effects of treatments.
In the past, the importance of EEGs in psychiatry was suggested (Boutros, 2018; Inui et al., 1998; Shelley & Trimble, 2009; Yasuhara 2010). However, despite prior findings and those in the current paragraph, the EEG still barely has any clinical applications in psychiatry that are sufficiently science-based. This might be a consequence of abnormalities being linked to the disorders themselves rather than to treatment outcome. Our focus on improving treatments instead of diagnostics may facilitate the introduction of EEG assessment in psychiatry.

All in all, the influence of ADs on the EEG, and therefore EEG biomarkers extracted from it, is characterized by both stability and changes: Whereas FAA is reflecting a trait and is thus not sensitive to changes in state and drug effects, mediation of AD response by normalization of EEG abnormalities suggests a direct relation to the acute pathophysiology in those patients, being sensitive to state changes. This latter finding is also suggestive of a possible drug repurposing of sertraline, as a (mild) anticonvulsant, however, that will require further study.

WHAT CAN WE LEARN ABOUT THE UNDERLYING PATHOPHYSIOLOGY WHEN STUDYING THE TIME COURSE OF BIOMARKERS?

Although biomarker based treatment outcome results are promising, they do not explain why these biomarkers actually *have* value in treatment prediction. We sought to provide underlying explanations for the prognostic value of both FAA and subclinically abnormal EEG activity. This was discussed in **chapters 4 and 6**, and in earlier studies by Arns et al. (2017; 2016).

For working mechanisms of EEG abnormalities, we gained only little insight into the association with sertraline and treatment response. Sanchez et al. (2014) did find differing pharmacodynamic profiles in mice, with sertraline containing the most pronounced dopamine active transporter inhibitory activity, compared to escitalopram and paroxetine. Kanekar et al. (2018) corroborated this through animal models; they suggested improved dopaminergic transmission by sertraline, as opposed to paroxetine and escitalopram. However, it remains unclear how this could relate to the (subclinically abnormal) human EEG. This biomarker is considerably less investigated than FAA, and more research is necessary to obtain more insight.

Remarkable is the fact that FAA is found to be a trait, while it seems to predict at the same time: it can indicate change, while remaining stable itself. This seems contradictory. The question is whether this can be explained by which role alpha oscillations have in the depression network, especially when the activity is by definition measured at the surface of the skull. In a first study that measured resting state EEG (CSD) and fMRI in healthy males, Kaur and colleagues (2020) showed how several parts of the brain are either normally or reversely correlated to FAA. More specifically, Zotev and colleagues (2016) discovered a pathway by combining EEG and fMRI measurements, correlating FAA to blood-oxygen-level-dependent (BOLD) activity in the left amygdala. Their results suggested FAA variations to be strongly related to emotion regulation. Treatment with ADs resulted in a reduction of exaggerated reactivity to emotional faces in the left amygdala (Sheline et al., 2001). This was immediately visible after a single dose of the SSRI citalopram in another study (Murphy et al., 2009). On the cellular level, in mice, SSRI fluoxetine treatment coincided with the reorganization of inhibitory circuits in the baso-

lateral amygdala, through alterations in somatostatin interneurons (Carceller et al., 2018). Striking is the fact that a number of female depressed patients displayed reduced gene expression of this exact somatostatin, with no such display in males (Douillard-Guilloux et al., 2017). By integrating these results, we might provide possible components of a larger mechanism (i.e. depression network) that could underlie treatment prediction by FAA, specifically in females. However, such integration for understanding the role of alpha oscillations in the depression network is complicated: our confirmation of FAA being a *trait* in **chapter 3**, challenges the correlation between FAA and amygdala BOLD activity, when possibly state related (AD treatment) cellular *changes* do occur in the amygdala. A lack of insight might in part be a consequence of attempting to discern the depression network with multiple – and differing – single-modality studies. The low number of multimodal studies that addresses the entire underlying mechanism, currently limits the possibilities of coherently explaining our outcomes in FAA predicted treatment outcome. The obscured underlying pathophysiology warrants more future studies, and in particular multimodal studies.

7.2 BIOMARKERS IN CLINICAL PRACTICE

As I introduced in this dissertation, clinically accepted biomarkers for detection and treatment of disorders in the field of mental health are not readily available. In MDD, several biomarkers in the EEG have been proposed, including FAA, the APF, and subclinical EEG abnormalities. As reservations exist on true biomarker effectiveness in MDD (Widge et al., 2019), replication studies and more knowledge on biomarkers' characteristics are needed to understand their full potential. Pioneering in prospectively testing the predictive value of the most promising EEG biomarkers, we aimed to explore:

CAN EEG BIOMARKERS BE RELIABLY IMPLEMENTED FOR THE TREATMENT OF DEPRESSION?

Several reports on treatment outcome through EEG biomarkers that were identified in the international Study to Predict Optimized Treatment (iSPOT-D), revealed the three most promising: FAA, APF,

and subclinical EEG abnormalities. In FAA, a right frontal dominant FAA was associated with response to SSRIs, and left frontal dominant FAA was associated with non-response to SSRIs in females (Arns et al., 2016). Patients with a low APF or subclinical abnormalities, were more likely to be non-responsive to the ADs escitalopram and venlafaxine, whereas no such response effect existed for sertraline (Arns et al., 2017).

The identification of the specific subgroups in these reports enabled us to create a stepwise protocol, in which we integrated the biomarkers EEG abnormalities, APF and FAA. We discussed this extensively in **chapter 6**, where we presented the first results on a prospective feasibility trial. In one of two groups, ADs were prescribed according to this biomarker-based protocol. This is an alternative to the usual guideline informed, 'random' prescription by psychiatrists, for the same type of ADs, which was done in the second group. The stepwise protocol for prescribing antidepressants to patients appeared to be sufficiently practical and satisfactory. The involved clinicians and other professionals were motivated to follow instructions and advice. Moreover, patients with an EEG-informed prescription improved significantly more as measured with %BDI change (the only normally distributed outcome measure) with a remission rate of 29%, than patients with treatment as usual, with a remission rate of 17%. Based on the current results, we cannot disentangle specific AD effects from non-specific (placebo) effects. Our sample size calculation indicates that future studies will need an $n \geq 218$.

For useful testing in the clinic, biomarkers are expected to predict good treatment outcome with 90% accuracy (sensitivity) and indicate less than 10% false positives (specificity; Brower 2011). Generally, a binary choice is to be made between affected vs non-affected, or to treat vs not-to treat. This has major ethical implications, i.e. at what sensitivity/specificity is one allowed to discourage a treatment? However, in stratified psychiatry, we leverage the notion that there are multiple evidence based treatments in depression, such as cognitive behavioral therapy, SSRIs, SNRIs, repetitive transcranial magnetic stimulation, etc. (see figure 7.1, following pages). Randomized studies such as iSPOT-D have shown that by using randomization to three different antidepressants, no differences in efficacy are found on the group level (Arns et al., 2016). Therefore, if the biomarkers that we

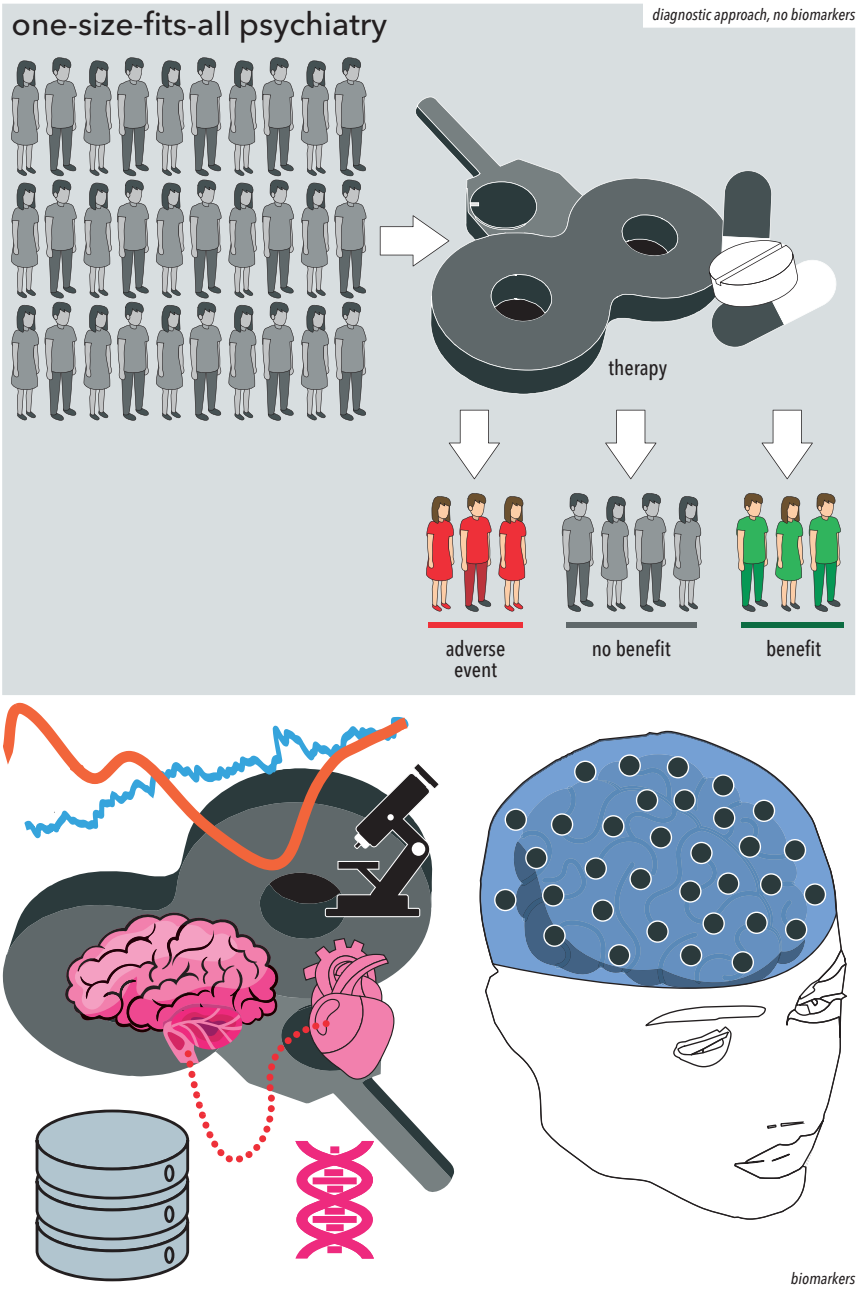
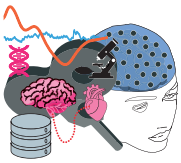
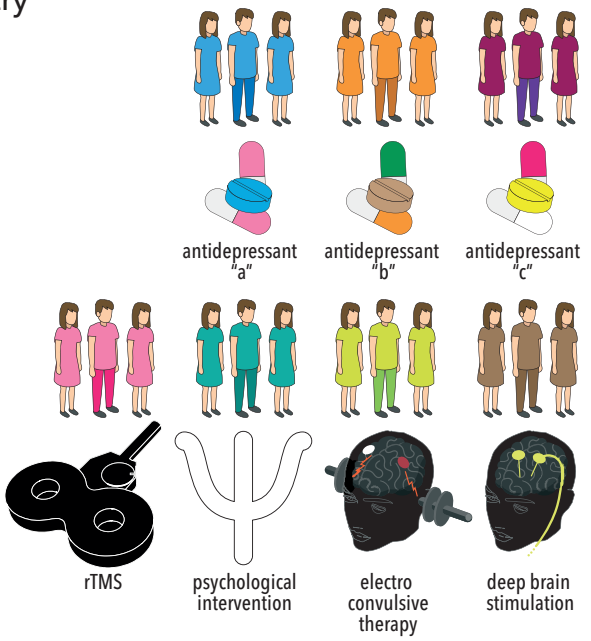


Figure 7.1: *Depiction of different types of treatment allocation, of which stratification in psychiatry is recommended in this dissertation, as opposed to one-size-fits-all psychiatry or precision psychiatry.*

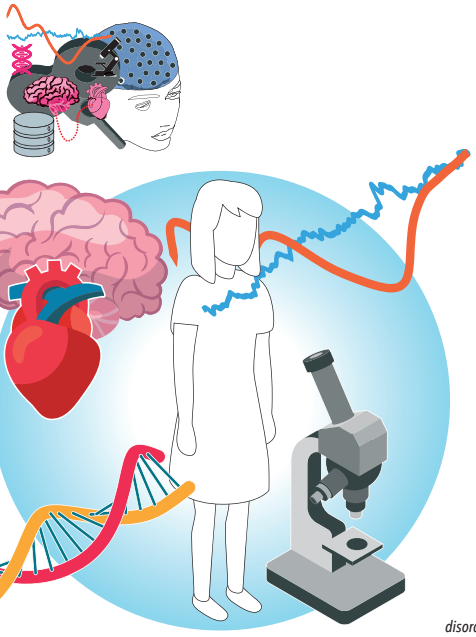
stratified psychiatry



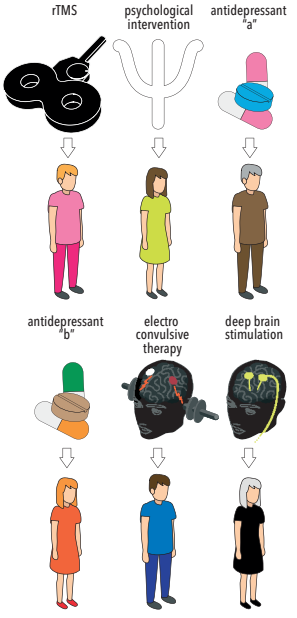
prognostic approach, using biomarkers



precision psychiatry



prognostic approach, using more biomarkers



disorder independent individualized therapy or prevention

discussed retrospectively have some value in explaining the heterogeneity between response and non-response, these could be used to guide an individual patient to their preferred treatment, or EEG biomarker informed treatment stratification. Instead of a binary go/no-go decision, the end result of this process is an evidence based treatment, where high sensitivity/specificity is not always required.

In sum, results from our feasibility trial indicate that the implementation of biomarkers remains worthy of further investigation. The naturalistic design is reflective of the clinical practice, enabling us to investigate real-world effectiveness. Wondering how we can improve the process, we asked ourselves:

WHICH CIRCUMSTANCES ARE NEEDED TO PROVIDE PATIENTS AND PROFESSIONALS WITH TRUSTWORTHY ADVICE?

The proposed biomarker protocol was initially developed for the benefit of drug-naive depressed patients being prescribed ADs for the first time, or after they quit the intake of ADs for a while. We assumed EEG assessments to be most optimal when brain function is uninfluenced by drugs. But in daily practice, a substantial part of patients enters a mental health facility while already on ADs. Either because it is not their first depressive episode and they are still taking medication since their previous episode, or because the GP already prescribed an AD. To serve this particular part of the depressed population, we need to establish the stability of EEG features irrespective of medication intake. Specifically of FAA, due to its largest influence on the prescription decision. As **chapter 3** shows, FAA can be regarded a trait. Its prognostic value is unaffected by state (different moments in time) or medication intake (with or without escitalopram, sertraline and venlafaxine intake). It suggests reliable implementation in clinical practice as a prognostic biomarker in both medicated and unmedicated patients. This enables us to serve a larger part of the depressed population. Although effects of medication on the EEG have been investigated before (i.e. in sleep stages and delta activity with agomelatine (Quera Salva et al., 2007), or EEG abnormalities or slowing with olanzapine (Degner et al., 2011)), these results are not directly translatable to frontal asymmetry in alpha oscillations. Future

studies should reveal whether FAA is stable on the individual level under any antidepressant, or any other medication for that matter. Current results suggest that circumstances of measurement with respect to moment of measurement and medication intake do not limit the prognostic advice. Next to establishing which patients can be offered trustworthy advice, we also aimed to improve methods by replacing visual assessment of the EEG by computational assessment, as far as current techniques allow us. We therefore asked:

CAN WE AUTOMATE THE DETECTION OF EEG ABNORMALITIES?

When working with the proposed biomarkers, a difference stands out: FAA and the APF are *calculated* with relatively established methods. However, detection of EEG abnormalities is done through *visual* assessment (for either severe abnormalities found in epilepsy or subclinically deviant activity). In our attempt to approach earlier outcomes of determining subclinical abnormalities in **chapter 5**, we utilized several features for the detection of severe deviations in the brain: a convolutional neural network and three spatiotemporal features (the sBSI, the tBSI and the dominant frequency). While we could achieve rather good overlap in ratings between subjective (visual) and objective (computed) ratings, unfortunately the objective methods did not significantly replicate the primary analysis of treatment prediction from Arns et al. (2017) through a combination of the features in a random forest model. While this may be partly related to a statistical power issue (i.e. even though the total sample consisted of 1008 patients, only 58 out of those presented with abnormalities), further optimizations of these methods could yield better and more reproducible results. This needs to be further investigated in future studies.

In short, automated detection approaches our visual assessment, but in its current form it does not yield prognostic value like visual detection. Further development of methods may hold promise in automating the detection of EEG abnormalities and potential in relieving work pressure of clinicians. In the meantime, we will still rely on the visual assessment of EEGs for studies in the near future, for which we recommend to have two independent EEG evaluators and only include EEGs when both evaluators agree.

7.4 CONCLUSION

The complexity of a heterogeneous depressed population makes it impossible to find one or few treatments that fit all. However, this heterogeneity can be embraced by employing biomarkers that are capable of identifying homogenic subgroups, and potentially predicting treatment outcome. Reliability and qualitative improvement can be achieved through state-of-the-art computerized automatization of this process, although these methods need further development before this reaches the psychiatric clinic. Established quantification methods however, already seem to allow us to predict response to treatment, based on alpha lateralization and alpha peak frequency. Combined with qualitative assessment of subclinical EEG abnormalities in our newly developed protocol, the first results for EEG-informed prescription of antidepressants show sufficient feasibility in a clinical setting. To our knowledge, this is the first attempt to elevate the treatment of depression through these biomarkers, which not only shows the protocol is non-inferior: patients actually show a modest increase in symptom improvement. The proposed protocol therefore not only makes our methods easily translatable to clinical practice, it bears the promise of a small but much needed achievement of higher treatment standards, in a new form of neuropsychiatric health care for depression.

I hope that this dissertation encourages follow-up research, further focusing on informed treatment decision making by (EEG) biomarkers. Our studies hold promising results that may pave the road to personalized medicine, thereby helping the continuously growing group of people suffering from depression.

REFERENCES

Allen, J. J. B., Urry, H. L., Hitt, S. K., & Coan, J. A. (2004). The stability of resting frontal electroencephalographic asymmetry in depression. *Psychophysiology*, **41**(2), 269-80. doi:10.1111/j.1469-8986.2003.00149.x

Alper, K., Schwartz, K. A., Kolts, R. L., & Khan, A. (2007). Seizure incidence in psychopharmacological clinical trials: An analysis of food and drug administration (FDA) summary basis of approval reports. *Biological Psychiatry*, **62**(4), 345-54. doi:10.1016/j.biopsych.2006.09.023

Arns, M., Bruder, G., Hegerl, U., Spooner, C., Palmer, D. M., Etkin, A., . . . Gordon, E. (2016). EEG alpha asymmetry as a gender-specific predictor of outcome to acute treatment with different antidepressant medications in the randomized iSPOT-D study. *Clinical Neurophysiology*, **127**, 509-19. doi:http://dx.doi.org/10.1016/j.clinph.2015.05.032

Arns, M., Gordon, E., & Boutros, N. N. (2017). EEG abnormalities are associated with poorer depressive symptom outcomes with escitalopram and venlafaxine-xr, but not sertraline: Results from the multicenter randomized iSPOT-D study. *Clinical EEG and Neuroscience*, **48**(1), 33-40. doi:10.1177/1550059415621435

Arns, M., Gunkelman, J., Breteler, M., & Spronk, D. (2008). EEG phenotypes predict treatment outcome to stimulants in children with ADHD. *Journal of Integrative Neuroscience*, **7**(3), 421-38. doi:10.1142/s0219635208001897

Askamp, J., & van Putten, M. (2014). Mobile EEG in epilepsy. *International Journal of Psychophysiology*, **91**, 30-35. doi:10.1016/j.ijpsycho.2013.09.002

Baehr, E., Rosenfeld, J. P., Baehr, R., & Earnest, C. (1998). Comparison of two EEG asymmetry indices in depressed patients vs. Normal controls. *International Journal of Psychophysiology* : Official Journal of the International Organization of Psychophysiology, **31**(1), 89-92.

Baker, M., Kale, R., & Menken, M. (2002). The wall between neurology and psychiatry. *The British Medical Journal*, **324**(7325), 1468-9.

Bares, M., Novak, T., Vlcek, P., Hejzlar, M., & Brunovsky, M. (2019). Early change of prefrontal theta cordance and occipital alpha asymmetry in the prediction of responses to antidepressants. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, **143**, 1-8. doi:10.1016/j.ijpsycho.2019.06.006

Baskaran, A., Milev, R., & McIntyre, R. (2012). The neurobiology of the EEG biomarker as a predictor of treatment response in depression. *Neuropharmacology*, **63**(4), 507-13. doi:10.1016/j.neuropharm.2012.04.021

Bauder, R. A., & Khoshgoftaar, T. M. (2018). The effects of varying class distribution on learner behavior for medicare fraud detection with imbalanced big data. *Health Inf Sci Syst*, **6**(1), 9. doi:10.1007/s13755-018-0051-3

Beck, A., Steer, R., & Brown, G. (1996). Manual for the Beck Depression Inventory-II: *Manual for the beck depression inventory-ii.* San Antonio, TX: Psychological Corporation.

Beeney, J. E., Levy, K. N., Gatzke-Kopp, L. M., & Hallquist, M. N. (2014). EEG asymmetry in borderline personality disorder and depression following rejection. *Personal Disord*, **5**(2), 178-85. doi:10.1037/per0000032

Bell, I. R., Schwartz, G. E., Hardin, E. E., Baldwin, C. M., & Kline, J. P. (1998). Differential

resting quantitative electroencephalographic alpha patterns in women with environmental chemical intolerance, depressives, and normals. *Biological Psychiatry*, **43**(5), 376-88.
Berger, H. (1929). Uber das elektrenkephalogramm des menschen. *Arch Psychiatr Nervenkr*, **87**, 527-570.
Boutros, N. N. (2014). *Standard EEG: A research roadmap for neuropsychiatry*. Cham, Switzerland: Springer.
Boutros, N. N. (2018). The forsaking of the clinical EEG by psychiatry: How justified? *CNS Spectrums*, **23**(3), 196-204. doi:10.1017/S1092852917000505
Boutros, N. N., Bowyer, S., Wang, J., Urfy, M. Z., & Loeb, J. A. (2015). Epilepsy spectrum disorders: A concept in need of validation or refutation. *Med Hypotheses*, **85**(5), 656-63. doi:10.1016/j.mehy.2015.08.004
Bozzi, Y., & Borrelli, E. (2013). The role of dopamine signaling in epileptogenesis. F*ront Cell Neurosci*, **7**, 157. doi:10.3389/fncel.2013.00157
Bragatti, J. A., Torres, C. M., Cherubini, P. A., Leistner-Segal, S., & Bianchin, M. M. (2014). Is interictal EEG activity a biomarker for mood disorders in temporal lobe epilepsy? *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, **125**(10), 1952-8. doi:10.1016/j.clinph.2014.02.013
Brower, V. (2011). Biomarkers: Portents of malignancy. *Nature*, **471**(7339), S19-21. doi:10.1038/471S19a
Bruder, G. E., Sedoruk, J. P., Stewart, J. W., McGrath, P. J., Quitkin, F. M., & Tenke, C. E. (2008). Electroencephalographic alpha measures predict therapeutic response to a selective serotonin reuptake inhibitor antidepressant: Pre- and post-treatment findings. *Biological Psychiatry*, **63**(12), 1171-7. doi:10.1016/j.biopsych.2007.10.009
Bruder, G. E., Stewart, J. W., & McGrath, P. J. (2017). Right brain, left brain in depressive disorders: Clinical and theoretical implications of behavioral, electrophysiological and neuroimaging findings. *Neurosci Biobehav Rev*, **78**, 178-191. doi:10.1016/j.neubiorev.2017.04.021
Bruder, G. E., Stewart, J. W., Tenke, C. E., McGrath, P. J., Leite, P., Bhattacharya, N., & Quitkin, F. M. (2001). Electroencephalographic and perceptual asymmetry differences between responders and nonresponders to an SSRI antidepressant. *Biological Psychiatry*, **49**(5), 416-25.
Brzezicka, A., Kamiński, J., Kamińska, O. K., Wołyńczyk-Gmaj, D., & Sedek, G. (2016). Frontal EEG alpha band asymmetry as a predictor of reasoning deficiency in depressed people. *Cognition & Emotion*, **1-11**. doi:10.1080/02699931.2016.1170669
Burton, R. (1883). *Anatomy of melancholy* (1st ed.). Philadelphia: E. Claxton & Company.
Cantisani, A., Koenig, T., Stegmayer, K., Federspiel, A., Horn, H., Müller, T. J., . . . Walther, S. (2015). EEG marker of inhibitory brain activity correlates with resting-state cerebral blood flow in the reward system in major depression. *European Archives of Psychiatry and Clinical Neuroscience,* **1-10**.
Carceller, H., Perez-Rando, M., Castren, E., Nacher, J., & Guirado, R. (2018). Effects of the antidepressant fluoxetine on the somatostatin interneurons in the basolateral amygdala. *Neuroscience*, **386**, 205-213. doi:10.1016/j.neuroscience.2018.06.041
Carethers, J., Braun, J., & Sands, B. (2015). Genetics, genetic testing, and biomarkers of digestive diseases. *Gastroenterology*, **149**(5), 1131-1133. doi:10.1053/j.gastro.2015.08.049
Carvalho, A., Moraes, H., Silveira, H., Ribeiro, P., Piedade, R. A. M., Deslandes, A. C., . . . Versiani, M. (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), 143-8. doi:10.1016/j.jad.2010.08.023

Chu, D. A., Williams, L. M., Harris, A. W. F., Bryant, R. A., & Gatt, J. M. (2013). Early life trauma predicts self-reported levels of depressive and anxiety symptoms in nonclinical community adults: Relative contributions of early life stressor types and adult trauma exposure. *Journal of Psychiatric Research*, **47**(1), 23-32. doi:10.1016/j.jpsychires.2012.08.006

Cipriani, A., Furukawa, T. A., Salanti, G., Chaimani, A., Atkinson, L. Z., Ogawa, Y., . . . Geddes, J. R. (2018). Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: A systematic review and network meta-analysis. *Lancet* (London, England), **391**(10128), 1357-1366. doi:10.1016/S0140-6736(17)32802-7

Ciumas, C., Wahlin, T. B. R., Espino, C., & Savic, I. (2010). The dopamine system in idiopathic generalized epilepsies: Identification of syndrome-related changes. *Neuroimage*, **51**(2), 606-15. doi:10.1016/j.neuroimage.2010.02.051

Ciumas, C., Wahlin, T. B. R., Jucaite, A., Lindstrom, P., Halldin, C., & Savic, I. (2008). Reduced dopamine transporter binding in patients with juvenile myoclonic epilepsy. *Neurology*, **71**(11), 788-94. doi:10.1212/01.wnl.0000316120.70504.d5

Coan, J. A., Allen, J. J. B., & McKnight, P. E. (2006). A capability model of individual differences in frontal EEG asymmetry. *Biological Psychology*, **72**(2), 198-207. doi:10.1016/j.biopsycho.2005.10.003

Cuijpers, P., Karyotaki, E., Reijnders, M., & Ebert, D. D. (2019). Was Eysenck right after all? A reassessment of the effects of psychotherapy for adult depression. *Epidemiology and Psychiatric Sciences*, **28**(1), 21-30. doi:10.1017/S2045796018000057

Cuijpers, P., Sijbrandij, M., Koole, S. L., Andersson, G., Beekman, A. T., & Reynolds, C. F. (2014). Adding psychotherapy to antidepressant medication in depression and anxiety disorders: A meta-analysis. *World Psychiatry : Official Journal of the World Psychiatric Association (WPA)*, **13**(1), 56-67. doi:10.1002/wps.20089

Cuijpers, P., van Straten, A., van Oppen, P., & Andersson, G. (2008). Are psychological and pharmacologic interventions equally effective in the treatment of adult depressive disorders? A meta-analysis of comparative studies. *The Journal of Clinical Psychiatry,* **69**(11), 1675-85; quiz 1839-41.

Cuijpers, P., van Straten, A., Warmerdam, L., & Andersson, G. (2009). Psychotherapy versus the combination of psychotherapy and pharmacotherapy in the treatment of depression: A meta-analysis. *Depression and Anxiety*, **26**(3), 279-88. doi:10.1002/da.20519

Cuthbert, B. N. (2014). Translating intermediate phenotypes to psychopathology: The NIMH research domain criteria. *Psychophysiology*, **51**(12), 1205-6. doi:10.1111/psyp.12342

Davidson, R. J. (1984). Affect, cognition, and hemispheric specialization. In C. E. Izard, J. Kagan, & R. B. Zajonc (Eds.), *Emotion, cognition, and behavior* (**pp. 320-65**). Cambridge: Cambridge University Press.

Davidson, R. J. (1998). Anterior electrophysiological asymmetries, emotion, and depression: Conceptual and methodological conundrums. *Psychophysiology*, **35**(5), 607-14.

Davidson, R. J., Kabat-Zinn, J., Schumacher, J., Rosenkranz, M., Muller, D., Santorelli, S. F., . . . Sheridan, J. F. (2003). Alterations in brain and immune function produced by mindfulness meditation. *Psychosom Med*, **65**(4), 564-70.

Davies, R. K. (1978). Incest: Some neuropsychiatric findings. *Int J Psychiatry Med*, **9**(2), 117-21.

Debener, S., Beauducel, A., Nessler, D., Brocke, B., Heilemann, H., & Kayser, J. (2000). Is resting anterior EEG alpha asymmetry a trait marker for depression? Findings for healthy adults and clinically depressed patients. *Neuropsychobiology*, **41**(1), 31-7.

Degner, D., Nitsche, M. A., Bias, F., Rüther, E., & Reulbach, U. (2011). EEG alterations during treatment with olanzapine. *European Archives of Psychiatry and Clinical Neuroscience*, **261**(7), 483-8. doi:10.1007/s00406-011-0208-4

Deldin, P. J., & Chiu, P. (2005). Cognitive restructuring and EEG in major depression. *Biological Psychology*, **70**(3), 141-51. doi:10.1016/j.biopsycho.2005.01.003

d'Elia, G., & Perris, C. (1973). Cerebral functional dominance and depression. An analysis of EEG amplitude in depressed patients. *Acta Psychiatrica Scandinavica*, **49**(3), 191.

Deslandes, A. C., de Moraes, H., Pompeu, F. A. M. S., Ribeiro, P., Cagy, M., Capitão, C., . . . Laks, J. (2008). Electroencephalographic frontal asymmetry and depressive symptoms in the elderly. *Biological Psychology*, **79**(3), 317-22. doi:10.1016/j.biopsycho.2008.07.008

Diagnostic and Statistical Manual of Mental Disorders :DSM-5. (2013). Diagnostic and statistical manual of mental disorders :*DSM-5*. Washington, D.C.: American Psychiatric Association.

Douillard-Guilloux, G., Lewis, D., Seney, M. L., & Sibille, E. (2017). Decrease in somatostatin-positive cell density in the amygdala of females with major depression. *Depression and Anxiety*, **34**(1), 68-78. doi:10.1002/da.22549

Favale, E., Audenino, D., Cocito, L., & Albano, C. (2003). The anticonvulsant effect of citalopram as an indirect evidence of serotonergic impairment in human epileptogenesis. *Seizure*, **12**(5), 316-8. doi:10.1016/S1059-1311(02)00315-1

Feldmann, L., Piechaczek, C. E., Grünewald, B. D., Pehl, V., Bartling, J., Frey, M., . . . Greimel, E. (2018). Resting frontal EEG asymmetry in adolescents with major depression: Impact of disease state and comorbid anxiety disorder. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, **129**(12), 2577-2585. doi:10.1016/j.clinph.2018.09.028

Fingelkurts, A. A., & Fingelkurts, A. A. (2015). Altered structure of dynamic electroencephalogram oscillatory pattern in major depression. *Biological Psychiatry*, **77**(12), 1050-60. doi:10.1016/j.biopsych.2014.12.011

Fingelkurts, A. A., Fingelkurts, A. A., Rytsälä, H., Suominen, K., Isometsä, E., & Kähkönen, S. (2006). Composition of brain oscillations in ongoing EEG during major depression disorder. *Neurosci Res*, **56**(2), 133-44. doi:10.1016/j.neures.2006.06.006

Fournier, J. C., DeRubeis, R. J., Hollon, S. D., Dimidjian, S., Amsterdam, J. D., Shelton, R. C., & Fawcett, J. (2010). Antidepressant drug effects and depression severity: A patient-level meta-analysis. *JAMA*, **303**(1), 47-53. doi:10.1001/jama.2009.1943

Fulcher, B. D., Little, M. A., & Jones, N. S. (2013). Highly comparative time-series analysis: the empirical structure of time series and their methods. *Journal of the Royal Society Interface*, **10**(83), 20130048. doi:10.1098/rsif.2013.0048

Geut, I., Knottnerus, I. L. H., & van Putten, M. J. A. M. (2017). Detecting interictal discharges in first seizure patients: Ambulatory EEG or EEG after sleep deprivation? *Seizure*, **51**, 52-54. doi:10.1016/j.seizure.2017.07.019

Gold, C., Fachner, J., & Erkkilä, J. (2013). Validity and reliability of electroencephalographic frontal alpha asymmetry and frontal midline theta as biomarkers for depression. *Scand J Psychol*, **54**(2), 118-26. doi:10.1111/sjop.12022

Gollan, J. K., Hoxha, D., Chihade, D., Pflieger, M. E., Rosebrock, L., & Cacioppo, J. (2014). Frontal alpha EEG asymmetry before and after behavioral activation

treatment for depression. *Biological Psychology*, **99**, 198-208. doi:10.1016/j.biopsycho.2014.03.003
Goodwin, J. E. (1947). The significance of alpha variants in the EEG, and their relation ship to an epileptiform syndrome. *The American Journal of Psychiatry*, **104**(6), 369-79.
Gordon, E., Palmer, D. M., & Cooper, N. (2010). EEG alpha asymmetry in schizophrenia, depression, PTSD, panic disorder, ADHD and conduct disorder. *Clinical EEG and Neuroscience*, **41**(4), 178-183. doi:10.1177/155005941004100404
Gotlib, I. H., Ranganath, C., & Rosenfeld, J. P. (1998). EEG alpha asymmetry, depression, and cognitive functioning. *Cognition & Emotion*, **12**(3), 449-478.
Grant, A. C., Abdel-Baki, S. G., Weedon, J., Arnedo, V., Chari, G., Koziorynska, E., . . . Omurtag, A. (2014). EEG interpretation reliability and interpreter confidence: A large single-center study. *Epilepsy & Behavior : E&B*, **32**, 102-7. doi:10.1016/j.yebeh.2014.01.011
Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalogr Clin Neurophysiol*, **55**(4), 468-84.
Grinspoon, L., & Bakalar, J. B. (1981). Coca and cocaine as medicines: An historical review. *Journal of Ethnopharmacology*, **3**(2-3), 149-59. doi:10.1016/0378-8741(81)90051-9
Grünewald, B. D., Greimel, E., Trinkl, M., Bartling, J., Großheinrich, N., & Schulte-Körne, G. (2018). Resting frontal EEG asymmetry patterns in adolescents with and without major depression. *Biological Psychology*, **132**, 212-216. doi:10.1016/j.biopsycho.2018.01.003
Hagemann, D., Hewig, J., Seifert, J., Naumann, E., & Bartussek, D. (2005). The latent state-trait structure of resting EEG asymmetry: Replication and extension. *Psychophysiology*, **42**(6), 740-52. doi:10.1111/j.1469-8986.2005.00367.x
Hagemann, D., Naumann, E., Becker, G., Maier, S., & Bartussek, D. (1998). Frontal brain asymmetry and affective style: A conceptual replication. *Psychophysiology*, **35**(4), 372-88.
Hagemann, D., Naumann, E., Thayer, J. F., & Bartussek, D. (2002). Does resting electroencephalograph asymmetry reflect a trait? An application of latent state-trait theory. *Journal of Personality and Social Psychology*, **82**(4), 619-41.
Hamid, H., & Kanner, A. M. (2013). Should antidepressant drugs of the selective serotonin reuptake inhibitor family be tested as antiepileptic drugs? *Epilepsy & Behavior : E&B,* **26**(3), 261-5. doi:10.1016/j.yebeh.2012.10.009
Harmon-Jones, E., & Allen, J. J. (1997). Behavioral activation sensitivity and resting frontal EEG asymmetry: Covariation of putative indicators related to risk for mood disorders. *Journal of Abnormal Psychology*, **106**(1), 159-63.
Hashemi, A., Pino, L. J., Moffat, G., Mathewson, K. J., Aimone, C., Bennett, P. J., . . . Sekuler, A. B. (2016). Characterizing population EEG dynamics throughout adulthood. *ENeuro*, **3**(6). doi:10.1523/ENEURO.0275-16.2016
Henriques, J. B., & Davidson, R. J. (1991). Left frontal hypoactivation in depression. *Journal of Abnormal Psychology*, **100**(4), 535-545.
Hieronymus, F., Lisinski, A., Nilsson, S., & Eriksson, E. (2019). Influence of baseline severity on the effects of ssris in depression: An item-based, patient-level post-hoc analysis. *Lancet Psychiatry*, **6**(9), 745-752. doi:10.1016/S2215-0366(19)30216-0
Hippocrates, Withington, E. T., & Jones, W. H. S. (1923). *Works of hippocrates*. Cambridge, MA: Harvard University Press.
Hovorka, J., Herman, E., & Nemcová, I. (2000). Treatment of interictal depression with citalopram in patients with epilepsy. *Epilepsy & Behavior : E&B*, **1**(6), 444-447. doi:10.1006/ebeh.2000.0123

Insel, T., Cuthbert, B., Garvey, M., Heinssen, R., Pine, D. S., Quinn, K., . . . Wang, P. (2010). Research domain criteria (rdoc): Toward a new classification framework for research on mental disorders. *The American Journal of Psychiatry*, **167**(7), 748-51. doi:10.1176/appi.ajp.2010.09091379

Inui, K., Motomura, E., Okushima, R., Kaige, H., Inoue, K., & Nomura, J. (1998). Electroencephalographic findings in patients with DSM-IV mood disorder, schizophrenia, and other psychotic disorders. *Biological Psychiatry*, **43**(1), 69-75. doi:10.1016/S0006-3223(97)00224-2

Ito, Y., Teicher, M. H., Glod, C. A., & Ackerman, E. (1998). Preliminary evidence for aberrant cortical development in abused children: A quantitative EEG study. *J Neuropsychiatry Clin Neurosci*, **10**(3), 298-307. doi:10.1176/jnp.10.3.298

Ito, Y., Teicher, M. H., Glod, C. A., Harper, D., Magnus, E., & Gelbard, H. A. (1993). Increased prevalence of electrophysiological abnormalities in children with psychological, physical, and sexual abuse. *J Neuropsychiatry Clin Neurosci*, **5**(4), 401-8. doi:10.1176/jnp.5.4.401

JASP Team. (2017). *JASP (version 0.8.2)* [Computer Software].

Jaworska, N., Blier, P., Fusee, W., & Knott, V. (2012). α power, α asymmetry and anterior cingulate cortex activity in depressed males and females. *Journal of Psychiatric Research*, **46**(11), 1483-91. doi:10.1016/j.jpsychires.2012.08.003

Jeffreys, H. (1961). Theory of probability (3 ed.). Oxford, UK: Oxford University Press.

Jensen, O., Gelfand, J., Kounios, J., & Lisman, J. E. (2002). Oscillations in the alpha band (9-12 hz) increase with memory load during retention in a short-term memory task. *Cereb Cortex*, **12**(8), 877-82.

Jesulola, E., Sharpley, C. F., & Agnew, L. L. (2017). The effects of gender and depression severity on the association between alpha asymmetry and depression across four brain regions. *Behavioural Brain Research*, **321**, 232-239. doi:10.1016/j.bbr.2016.12.035

Jesulola, E., Sharpley, C. F., Bitsika, V., Agnew, L. L., & Wilson, P. (2015). Frontal alpha asymmetry as a pathway to behavioural withdrawal in depression: Research findings and issues. *Behavioural Brain Research*, **292**, 56-67. doi:10.1016/j.bbr.2015.05.058

Johnson, J. M., & Khoshgoftaar, T. M. (2019). Survey on deep learning with class imbalance. *Journal of Big Data*, **6**(27). doi:doi.org/10.1186/s40537-019-0192-5

Kaiser, A. K., Doppelmayr, M., & Iglseder, B. (2016). Electroencephalogram alpha asym metry in geriatric depression: Valid or vanished? *Zeitschrift Fur Gerontologie Und Geriatrie*. doi:10.1007/s00391-016-1108-z

Kamal, S. M. (2007). Effect of escitalopram and venlafaxine on seizure score and lipid peroxidation in mice receiving carbamazepine antiepileptic therapy. *J Egypt Soc Toxicol*, **36**, 1-6.

Kanekar, S., Sheth, C. S., Ombach, H. J., Olson, P. R., Bogdanova, O. V., Petersen, M., . . . Renshaw, P. F. (2018). Hypobaric hypoxia exposure in rats differentially alters antidepressant efficacy of the selective serotonin reuptake inhibitors fluoxetine, paroxetine, escitalopram and sertraline. *Pharmacol Biochem Behav*, **170**, 25-35. doi:10.1016/j.pbb.2018.05.002

Kanner, A. M. (2016). Most antidepressant drugs are safe for patients with epilepsy at therapeutic doses: A review of the evidence. *Epilepsy & Behavior : E&B*, **61**, 282-286. doi:10.1016/j.yebeh.2016.03.022

Kanner, A. M. (2017a). Psychiatric comorbidities in new onset epilepsy: Should they be always investigated? *Seizure*, **49**, 79-82. doi:10.1016/j.seizure.2017.04.007

Kanner, A. M., Kozak, A. M., & Frey, M. (2000). The use of sertraline in patients

with epilepsy: Is it safe? *Epilepsy & Behavior : E&B*, **1**(2), 100-105. doi:10.1006/ebeh.2000.0050

Kanner, A. M., Scharfman, H., Jette, N., Anagnostou, E., Bernard, C., Camfield, C., . . . Pohlmann-Eden, B. (2017b). Epilepsy as a network disorder (1): What can we learn from other network disorders such as autistic spectrum disorder and mood disorders? *Epilepsy & Behavior : E&B*, **77**, 106-113. doi:10.1016/j.yebeh.2017.09.014

Karch, J. D., Sander, M. C., von Oertzen, T., Brandmaier, A. M., & Werkle-Bergner, M. (2015). Using within-subject pattern classification to understand lifespan age differences in oscillatory mechanisms of working memory selection and maintenance. *Neuroimage*, **118**, 538-52. doi:10.1016/j.neuroimage.2015.04.038

Kaur, A., Chinnadurai, V., & Chaujar, R. (2020). Microstates-based resting frontal alpha asymmetry approach for understanding affect and approach/withdrawal behavior. *Scientific Reports*, **10**(1), 4228. doi:10.1038/s41598-020-61119-7

Kayser, J., & Tenke, C. E. (2015). On the benefits of using surface laplacian (current source density) methodology in electrophysiology. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, **97**(3), 171.

Kazamel, M. (2018). The split between neurology and psychiatry: Historical overview. *Neurology*, **90**(15 supplement), P5.305.

Keeser, D., Karch, S., Davis, J. R., Surmeli, T., Engelbregt, H., Länger, A., . . . Thatcher, R. W. (2013). Changes of resting-state EEG and functional connectivity in the sensor and source space of patients with major depression. *Klinische Neurophysiologie*, **44**(01), P142.

Kelley, N. J., Hortensius, R., Schutter, D. J. L. G., & Harmon-Jones, E. (2017). The relationship of approach/avoidance motivation and asymmetric frontal cortical activity: A review of studies manipulating frontal asymmetry. *International Journal of Psy chophysiology : Official Journal of the International Organization of Psychophysiology*, **119**, 19-30. doi:10.1016/j.ijpsycho.2017.03.001

Kemp, A., Griffiths, K., Felmingham, K., Shankman, S., Drinkenburg, P., Arns, M., . . . Bryant, R. (2010a). Disorder specificity despite comorbidity: Resting EEG alpha asymmetry in major depressive disorder and post-traumatic stress disorder. *Biological Psychology*, **85**(2), 350-354.

Kemp, B., Värri, A., Rosa, A. C., Nielsen, K. D., & Gade, J. (1992). A simple format for exchange of digitized polygraphic recordings. *Electroencephalogr Clin Neurophysiol*, **82**(5), 391-3. doi:10.1016/0013-4694(92)90009-7

Kentgen, L. M., Tenke, C. E., Pine, D. S., Fong, R., Klein, R. G., & Bruder, G. E. (2000). Electroencephalographic asymmetries in adolescents with major depression: Influence of comorbidity with anxiety disorders. *Journal of Abnormal Psychology*, **109**(4), 797-802. doi:10.1037//0021-843X.109.4.797

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Koretz, D., Merikangas, K. R., . . . National Comorbidity Survey Replication. (2003). The epidemiology of major depressive disorder: Results from the national comorbidity survey replication (NCS-R). *JAMA*, **289**(23), 3095-105. doi:10.1001/jama.289.23.3095

Keune, P. M., Bostanov, V., Hautzinger, M., & Kotchoubey, B. (2011). Mindfulness-based cognitive therapy (MBCT), cognitive style, and the temporal dynamics of frontal EEG alpha asymmetry in recurrently depressed patients. *Biological Psychology*, **88**(2-3), 243-52. doi:10.1016/j.biopsycho.2011.08.008

Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory

performance: A review and analysis. *Brain Res Brain Res Rev*, **29**(2-3), 169-95.
Klimesch, W., Doppelmayr, M., & Hanslmayr, S. (2006). Upper alpha ERD and absolute power: Their meaning for memory performance. *Prog Brain Res*, **159**, 151-65. doi:10.1016/S0079-6123(06)59010-7
Knott, V., Mahoney, C., Kennedy, S., & Evans, K. (2001). EEG power, frequency, asymmetry and coherence in male depression. *Psychiatry Research*, **106**(2), 123-40.
de Kovel, C. G. F., Aftanas, L., Aleman, A., Alexander-Bloch, A. F., Baune, B. T., Brack, I., . . . Francks, C. (2019). No alterations of brain structural asymmetry in major depressive disorder: An ENIGMA consortium analysis. *The American Journal of Psychiatry,* appiajp201918101144. doi:10.1176/appi.ajp.2019.18101144
Krauss, G. L., Fisher, R. S., & Kaplan, P. W. (2010). *The Johns Hopkins atlas of digital EEG : An interactive training guide*. Baltimore: Johns Hopkins University Press.
De La Garza, N., John Rush, A., Grannemann, B. D., & Trivedi, M. H. (2017). Toward a very brief self-report to assess the core symptoms of depression (VQIDS-SR5). *Acta Psychiatr Scand*, **135**(6), 548-553. doi:10.1111/acps.12720
Lemere, F. (1936). The significance of individual differences in the Berger rhythm. *Brain: A Journal of Neurology*, **366-375**.
Lennox-Buchtal, M., Buchtal, F., & Rosenfalck, P. (1960). Correlation of electroencephalographic findings with crash rate of military jet pilots. *Epilepsia*, **1**, 366-72.
Liaw, A., & Wiener, M. (2002, December). Classification and regression by randomforest. *The Newsletter of the R Project*, **2/3**.
Liu, H., Sarapas, C., & Shankman, S. A. (2016). Anticipatory reward deficits in melancholia. *Journal of Abnormal Psychology*, **125**(5), 631-40. doi:10.1037/abn0000172
Lodder, S. S., & van Putten, M. J. A. M. (2011). Automated EEG analysis: Characterizing the posterior dominant rhythm. *J Neurosci Methods*, **200**(1), 86-93. doi:10.1016/j.jneumeth.2011.06.008
da Silva Lourenço, C., Tjepkema-Cloostermans, M., Teixeira, L., & van Putten, M. (2020). Deep learning for interictal epileptiform discharge detection from scalp EEG recordings. In XV mediterranean conference on medical and biological engineering and computing – *MEDICON 2019* (**Vol. 76**, pp. 1984-1997).
Maguire, M. J., Weston, J., Singh, J., & Marson, A. G. (2014). Antidepressants for people with epilepsy and depression. *Cochrane Database of Systematic Reviews*, (**12**). doi:10.1002/14651858.CD010682.pub2
Mathers, C. D., & Loncar, D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Med*, **3**(11), e442. doi:10.1371/journal.pmed.0030442
Mathersul, D., Williams, L. M., Hopkinson, P. J., & Kemp, A. H. (2008). Investigating models of affect: Relationships among EEG alpha asymmetry, depression, and anxiety. *Emotion* (Washington, D.C.), **8**(4), 560-72. doi:10.1037/a0012811
McFarlane, A., Clark, C. R., Bryant, R. A., Williams, L. M., Niaura, R., Paul, R. H., . . . Gordon, E. (2005). The impact of early life stress on psychophysiological, personality and behavioral measures in 740 non-clinical subjects. *Journal of Integrative Neuroscience,* **4**(1), 27-40. doi:10.1142/S0219635205000689
Miller, G. (2010). Is pharma running out of brainy ideas? *Science*, **329**(5991), 502.
Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & Group, P. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Med*, **6**(7), e1000097.
Monin, J., Pruvost-Robieux, E., Huiban, N., Marchi, A., Crepon, B., Dubourdieu, D., . . . Gavaret, M. (2018). Prevalence of benign epileptiform variants during initial EEG

examination in french military aircrew. *Neurophysiol Clin*, **48**(3), 171-179. doi:10.1016/j.neucli.2018.04.001

Mora, M. S., Nestoriuc, Y., & Rief, W. (2011). Lessons learned from placebo groups in antidepressant trials. *Philos Trans R Soc Lond B Biol Sci*, **366**(1572), 1879-88. doi:10.1098/rstb.2010.0394

Murphy, S. E., Norbury, R., O'Sullivan, U., Cowen, P. J., & Harmer, C. J. (2009). Effect of a single dose of citalopram on amygdala response to emotional faces. *The British Journal of Psychiatry : The Journal of Mental Science*, **194**(6), 535-40. doi:10.1192/bjp.bp.108.056093

Niedermeyer, E. (2005). Abnormal EEG patterns: Epileptic and paroxysmal. In F. Lopes Da Silva (Ed.), *Electroencephalography: Basic principles, clinical applications, and related fields* (**5th ed**., pp. 255-280). Baltimore: Lippincott Williams & Wilkins.

Noachtar, S., Binnie, C., Ebersole, J., Mauguière, F., Sakamoto, A., & Westmoreland, B. (1999). A glossary of terms most commonly used by clinical electroencephalographers and proposal for the report form for the EEG findings. The international federation of clinical neurophysiology. *Electroencephalogr Clin Neurophysiol Suppl,* **52**, 21-41.

Nusslock, R., Shackman, A. J., McMenamin, B. W., Greischar, L. L., Davidson, R. J., & Kovacs, M. (2018). Comorbid anxiety moderates the relationship between depression history and prefrontal EEG asymmetry. *Psychophysiology*, **55**(1). doi:10.1111/psyp.12953

Nusslock, R., Walden, K., & Harmon-Jones, E. (2015). Asymmetrical frontal cortical activity associated with differential risk for mood and anxiety disorder symptoms: An rdoc perspective. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, **98**(2 Pt 2), 249-61. doi:10.1016/j.ijpsycho.2015.06.004

Oh, D. A., Kim, H., & Bae, E. K. (2018). Electroencephalographic abnormalities in the screening for pilot applicants in korea. *J Clin Neurol*, **14**(2), 212-214. doi:10.3988/jcn.2018.14.2.212

Okazaki, M., Adachi, N., Ito, M., Watanabe, M., Watanabe, Y., Kato, M., & Onuma, T. (2011). One-year seizure prognosis in epilepsy patients treated with antidepressants. *Epilepsy & Behavior : E&B*, **22**(2), 331-5. doi:10.1016/j.yebeh.2011.07.016

Patten, S. B., Williams, J. V. A., Lavorato, D. H., Bulloch, A. G. M., Wiens, K., & Wang, J. (2016). Why is major depression prevalence not changing? *Journal of Affective Disorders*, **190**, 93-97. doi:10.1016/j.jad.2015.09.002

Paul, R. H., Gunstad, J., Cooper, N., Williams, L. M., Clark, C. R., Cohen, R. A., . . . Gordon, E. (2007). Cross-cultural assessment of neuropsychological performance and electrical brain function measures: Additional validation of an international brain database. *Int J Neurosci*, **117**(4), 549-68. doi:10.1080/00207450600773665

Pfurtscheller, G., Stancák, A., & Neuper, C. (1996). Event-related synchronization (ERS) in the alpha band--an electrophysiological correlate of cortical idling: A review. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, **24**(1-2), 39-46.

Pizzagalli, D. A., Nitschke, J. B., Oakes, T. R., Hendrick, A. M., Horras, K. A., Larson, C. L., . . . Davidson, R. J. (2002). Brain electrical tomography in depression: The importance of symptom severity, anxiety, and melancholic features. B*iological Psychiatry*, **52**(2), 73-85.

Price, G. W., Lee, J. W., Garvey, C., & Gibson, N. (2008). Appraisal of sessional EEG

features as a correlate of clinical changes in an rtms treatment of depression. *Clinical EEG and Neuroscience*, **39**(3), 131-8.
Quera Salva, M. A., Vanier, B., Laredo, J., Hartley, S., Chapotot, F., Moulin, C., . . . Guilleminault, C. (2007). Major depressive disorder, sleep EEG and agomelatine: An open-label study. *The International Journal of Neuropsychopharmacology*, **10**(5), 691-6. doi:10.1017/S1461145707007754
Quinn, C., Rennie, C., Harris, A., & Kemp, A. (2014). The impact of melancholia ver sus non-melancholia on resting-state, EEG alpha asymmetry: Electrophysiological evidence for depression heterogeneity. *Psychiatry Research*, **215**, 614-17. doi:10.1016/j.psychres.2013.12.049
Quraan, M. A., Protzner, A. B., Daskalakis, Z. J., Giacobbe, P., Tang, C. W., Kennedy, S. H., . . . McAndrews, M. P. (2014). EEG power asymmetry and functional connectivity as a marker of treatment effectiveness in DBS surgery for depression. *Neuropsychopharmacology*, **39**(5), 1270-81. doi:10.1038/npp.2013.330
R Core Team. (2014). *R: A language and environment for statistical computing* [Computer Software]. Vienna, Austria: R Foundation for Statistical Computing.
Reid, S. A., Duke, L. M., & Allen, J. J. (1998). Resting frontal electroencephalographic asymmetry in depression: Inconsistencies suggest the need to identify mediating factors. *Psychophysiology*, **35**(4), 389-404.
Reilly, T. J. (2015). The neurology-psychiatry divide: A thought experiment. *BJPsych Bulletin*, **39**(3), 134-5. doi:10.1192/pb.bp.113.045740
Reynolds, E. H., & Wilson, J. V. K. (2013). Depression and anxiety in babylon. *Journal of the Royal Society of Medicine*, **106**(12), 478-81. doi:10.1177/0141076813486262
Reznik, S. J., & Allen, J. J. B. (2018). Frontal asymmetry as a mediator and moderator of emotion: An updated review. *Psychophysiology*, **55**(1). doi:10.1111/psyp.12965
Ribot, R., Ouyang, B., & Kanner, A. M. (2017). The impact of antidepressants on seizure frequency and depressive and anxiety disorders of patients with epilepsy: Is it worth investigating? *Epilepsy & Behavior : E&B*, **70**(Pt A), 5-9. doi:10.1016/j.yebeh.2017.02.032
Richter, P. L., Zimmerman, E. A., Raichle, M. E., & Liske, E. (1971). Electroencephalograms of 2,947 united states air force academy cadets (1965-1969). *Aerosp Med*, **42**(9), 1011-4.
Rosell, R., Bivona, T. G., & Karachaliou, N. (2013). Genetics and biomarkers in personalisation of lung cancer treatment. *Lancet* (London, England), **382**(9893), 720-31. doi:10.1016/S0140-6736(13)61715-8
Rosenberg, M. S., Adams, D. C., & Gurevitch, J. (2000). *MetaWin: Statistical software for meta-analysis*, **version 2.1**. Sinauer, Sunderland, MA.
Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review*, **16**, 225-237. doi:10.3758/PBR.16.2.225
Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., & Faubert, J. (2019). Deep learning-based electroencephalography analysis: A systematic review. *Journal of Neural Engineering*, **16**(5), 051001. doi:10.1088/1741-2552/ab260c
Rush, A. J., Trivedi, M. H., Wisniewski, S. R., Nierenberg, A. A., Stewart, J. W., Warden, D., . . . Fava, M. (2006). Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: A STAR*D report. *The American Journal of Psychiatry*, **163**(11), 1905-17. doi:10.1176/appi.ajp.163.11.1905
Ruxton, G. D., & Neuhäuser, M. (2010). When should we use one-tailed hypothesis testing? *Methods in Ecology and Evolution*, **1**, 114-117. doi:doi: 10.1111/j.2041-

210X.2010.00014.x
Saletu, B., Brandstätter, N., Metka, M., Stamenkovic, M., Anderer, P., Semlitsch, H. V., . . . Koll, B. (1996). Hormonal, syndromal and EEG mapping studies in menopausal syndrome patients with and without depression as compared with controls. *Maturitas*, **23**(1), 91-105.
Sanchez, C., Reines, E. H., & Montgomery, S. A. (2014). A comparative review of escitalopram, paroxetine, and sertraline: Are they all alike? *Int Clin Psychopharmacol*, **29**(4), 185-96. doi:10.1097/YIC.0000000000000023
Saveanu, R., Etkin, A., Duchemin, A. M., Goldstein-Piekarski, A., Gyurak, A., Debattista, C., . . . Williams, L. M. (2015). The international study to predict optimized treatment in depression (iSPOT-D): Outcomes from the acute phase of antidepressant treatment. *Journal of Psychiatric Research*, **61**, 1-12. doi:10.1016/j.jpsychires.2014.12.018
Schaffer, C. E., Davidson, R. J., & Saron, C. (1983). Frontal and parietal electroencephalogram asymmetry in depressed and nondepressed subjects. *Biological Psychiatry*, **18**(7), 753-62.
Segrave, R. A., Cooper, N. R., Thomson, R. H., Croft, R. J., Sheppard, D. M., & Fitzgerald, P. B. (2011). Individualized alpha activity and frontal asymmetry in major depression. *Clinical EEG and Neuroscience*, **42**(1), 45-52.
Sharbrough, F. W. (2005). Nonspecific abnormal EEG patterns. In E. Niedermeyer & F. Lopes Da Silva (Eds.), *Electroencephalography: Basic principles, clinical applications, and related fields* (**5th ed**., pp. 235-254). Baltimore: Lippincott Williams & Wilkins.
Sheehan, D. V., Lecrubier, Y., Sheehan, K. H., Amorim, P., Janavs, J., Weiller, E., . . . Dunbar, G. C. (1998). The mini-international neuropsychiatric interview (M.I.N.I.): The development and validation of a structured diagnostic psychiatric interview for DSM-IV and ICD-10. *The Journal of Clinical Psychiatry*, **59**, 22-33.
Sheline, Y. I., Barch, D. M., Donnelly, J. M., Ollinger, J. M., Snyder, A. Z., & Mintun, M. A. (2001). Increased amygdala response to masked emotional faces in depressed subjects resolves with antidepressant treatment: An fmri study. B*iological Psychiatry*, **50**(9), 651-8. doi:10.1016/s0006-3223(01)01263-x
Shelley, B. P., & Trimble, M. R. (2009). "All that spikes is not fits", mistaking the woods for the trees: The interictal spikes-an "EEG chameleon" in the interface disorders of brain and mind: A critical review. *Clinical EEG and Neuroscience*, **40**(4), 245-61. doi:10.1177/155005940904000407
Shelley, B. P., Trimble, M. R., & Boutros, N. N. (2008). Electroencephalographic cerebral dysrhythmic abnormalities in the trinity of nonepileptic general population, neuropsychiatric, and neurobehavioral disorders. *J Neuropsychiatry Clin Neurosci*, **20**(1), 7-22. doi:10.1176/appi.neuropsych.20.1.7
Sitges, M., Aldana, B. I., Gómez, C. D., & Nekrassov, V. (2012). The antidepressant sertraline prevents the behavioral and EEG changes induced in two animal models of seizures. *Epilepsy & Behavior : E&B*, **25**(4), 511-6. doi:10.1016/j.yebeh.2012.09.005
Smith, E. E., Reznik, S. J., Stewart, J. L., & Allen, J. J. B. (2017). Assessing and conceptualizing frontal EEG asymmetry: An updated primer on recording, processing, analyzing, and interpreting frontal alpha asymmetry. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, **111**, 98-114. doi:10.1016/j.ijpsycho.2016.11.005
Specchio, L. M., Iudice, A., Specchio, N., La Neve, A., Spinelli, A., Galli, R., . . . Murri, L. (2004). Citalopram as treatment of depression in patients with epilepsy. *Clin Neuropharmacol*, **27**(3), 133-136.

Spronk, D., Arns, M., Barnett, K., Cooper, N., & Gordon, E. (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. doi:10.1016/j.jad.2010.06.021

Spronk, D., Arns, M., Bootsma, A., van Ruth, R., & Fitzgerald, P. B. (2008). Long-term effects of left frontal rtms on EEG and erps in patients with depression. *Clinical EEG and Neuroscience*, **39**(3), 118-24. doi:10.1177/155005940803900305

Stewart, J. A., Deliyannides, D. A., Hellerstein, D. J., McGrath, P. J., & Stewart, J. W. (2012). Can people with nonsevere major depression benefit from antidepressant medication? *The Journal of Clinical Psychiatry*, **73**(4), 518-25. doi:10.4088/JCP.10m06760

Stewart, J. L., Bismark, A. W., Towers, D. N., Coan, J. A., & Allen, J. J. B. (2010). Resting frontal EEG asymmetry as an endophenotype for depression risk: Sex-specific patterns of frontal brain asymmetry. *Journal of Abnormal Psychology*, **119**(3), 502-12. doi:10.1037/a0019196

Stewart, J. L., Coan, J. A., Towers, D. N., & Allen, J. J. B. (2014). Resting and task-elicited prefrontal EEG alpha asymmetry in depression: Support for the capability model. *Psychophysiology*, **51**(5), 446-55. doi:10.1111/psyp.12191

Struve, F. A., & Boutros, N. N. (2005). Somatic implications of generalized and/or focal EEG slowing in psychiatric patients. *Clinical EEG and Neuroscience*, **36**(3), 171-5.

Sutton, S. K., & Davidson, R. J. (1997). Prefrontal brain asymmetry: A biological substrate of the behavioral approach and inhibition systems. *Psychological Science*, **8**(3), 204-210.

Teicher, M. H., Andersen, S. L., Polcari, A., Anderson, C. M., Navalta, C. P., & Kim, D. M. (2003). The neurobiological consequences of early stress and childhood maltreatment. *Neurosci Biobehav Rev*, **27**(1-2), 33-44.

Teicher, M. H., Ito, Y., Glod, C. A., Andersen, S. L., Dumont, N., & Ackerman, E. (1997). Preliminary evidence for abnormal cortical development in physically and sexually abused children using EEG coherence and MRI. *Ann N Y Acad Sci*, **821**, 160-75.

Telles-Correia, D., & Marques, J. G. (2015). Melancholia before the twentieth century: Fear and sorrow or partial insanity? *Frontiers in Psychology*, **6**, 81. doi:10.3389/fpsyg.2015.00081

Tenke, C. E., Kayser, J., Pechtel, P., Webb, C. A., Dillon, D. G., Goer, F., . . . Bruder, G. E. (2017). Demonstrating test-retest reliability of electrophysiological measures for healthy adults in a multisite study of biomarkers of antidepressant treatment response. *Psychophysiology,* **54**(1), 34-50. doi:10.1111/psyp.12758

The World Bank (2019 revision). *Population, total - World*. Retrieved from https://data.worldbank.org/indicator/SP.POP.TOTL?locations=1W&view=chart

Thibodeau, R., Jorgensen, R. S., & Kim, S. (2006). Depression, anxiety, and resting frontal EEG asymmetry: A meta-analytic review. *Journal of Abnormal Psychology*, **115**(4), 715-29. doi:10.1037/0021-843X.115.4.715

Thomé-Souza, M. S., Kuczynski, E., & Valente, K. D. (2007). Sertraline and fluoxetine: Safe treatments for children and adolescents with epilepsy and depression. *Epilepsy & Behavior : E&B*, **10**(3), 417-25. doi:10.1016/j.yebeh.2007.01.004

Tjepkema-Cloostermans, M. C., da Silva Lourenço, C., Ruijter, B. J., Tromp, S. C., Drost, G., Kornips, F. H. M., . . . van Putten, M. J. A. M. (2019). Outcome prediction in postanoxic coma with deep learning. *Critical Care Medicine*, **47**(10), 1424-1432. doi:10.1097/CCM.0000000000003854

Tjepkema-Cloostermans, M. C., de Carvalho, R. C. V., & van Putten, M. J. A. M. (2018). Deep learning for detection of focal epileptiform discharges from scalp EEG recordings. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, **129**(10), 2191-2196. doi:10.1016/j.clinph.2018.06.024

Tjepkema-Cloostermans, M. C., Hofmeijer, J., Beishuizen, A., Hom, H. W., Blans, M. J., Bosch, F. H., & van Putten, M. J. A. M. (2017). Cerebral recovery index: Reliable help for prediction of neurologic outcome after cardiac arrest. *Critical Care Medicine*, **45**(8), e789-e797. doi:10.1097/CCM.0000000000002412

Tomarken, A. J., Davidson, R. J., Wheeler, R. E., & Kinney, L. (1992). Psychometric properties of resting anterior EEG asymmetry: Temporal stability and internal consistency. *Psychophysiology*, **29**(5), 576-92.

Towers, D. N., & Allen, J. J. B. (2009). A better estimate of the internal consistency reliability of frontal EEG asymmetry scores. *Psychophysiology*, **46**(1), 132-42. doi:10.1111/j.1469-8986.2008.00759.x

van der Does, W. (2002). *Manual of the Dutch version of the Beck Depression Inventory (BDI-II-NL)*. Amsterdam, NL: Harcourt.

van der Vinne, N., Vollebregt, M. A., Boutros, N. N., Fallahpour, K., van Putten, M. J. A. M., & Arns, M. (2019a). Normalization of EEG in depression after antidepressant treatment with sertraline? A preliminary report. *Journal of Affective Disorders*, **259**, 67-72. doi:10.1016/j.jad.2019.08.016

van der Vinne, N., Vollebregt, M., Van Putten, M., & Arns, M. (2019b). Stability of frontal alpha asymmetry in depressed patients during antidepressant treatment. *Neuro Image: Clinical*, (102056). doi:10.1016/j.nicl.2019.102056

van der Vinne, N., Vollebregt, M. A., van Putten, M. J., & Arns, M. (2017). Frontal alpha asymmetry as a diagnostic marker in depression: Fact or fiction? A meta-analysis. *NeuroImage: Clinical*, **16**, 79-87. doi:10.1016/j.nicl.2017.07.006

van Dijk, H., Schoffelen, J. M., Oostenveld, R., & Jensen, O. (2008). Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *The Jour nal of Neuroscience : The Official Journal of the Society for Neuroscience*, **28**(8), 1816-23. doi:10.1523/JNEUROSCI.1853-07.2008

van Leeuwen, K. G., Sun, H., Tabaeizadeh, M., Struck, A. F., van Putten, M. J. A. M., & Westover, M. B. (2019). Detecting abnormal electroencephalograms using deep convolutional networks. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, **130**(1), 77-84. doi:10.1016/j.clinph.2018.10.012

van Putten, M. J. A. M. (2006). Extended BSI for continuous EEG monitoring in carotid endarterectomy. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, **117**(12), 2661-6. doi:10.1016/j.clinph.2006.08.007

van Putten, M. J. A. M. (2007). The revised brain symmetry index. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, **118**(11), 2362-7. doi:10.1016/j.clinph.2007.07.019

van Putten, M. J. A. M. (2008). The colorful brain: Visualization of EEG background patterns. *J Clin Neurophysiol*, **25**(2), 63-8. doi:10.1097/WNP.0b013e31816bdf85

van Putten, M. J. A. M., Peters, J. M., Mulder, S. M., de Haas, J. A. M., Bruijninckx, C. M. A., & Tavy, D. L. J. (2004). A brain symmetry index (BSI) for online EEG monitoring in carotid endarterectomy. *Clinical Neurophysiology : Official Journal of the International Federation of Clinical Neurophysiology*, **115**(5), 1189-94. doi:10.1016/j.clinph.2003.12.002

Vos, T., Allen, C., Arora, M., Barber, R. M., Bhutta, Z. A., Brown, A., . . . Graetz, N. (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. *Lancet* (London, England), **388**(10053), 1545-1602. doi:10.1016/S0140-6736(16)31678-6

Vuga, M., Fox, N. A., Cohn, J. F., George, C. J., Levenstein, R. M., & Kovacs, M. (2006). Long-term stability of frontal electroencephalographic asymmetry in adults with a history of depression and controls. *International Journal of Psychophysiology : Official Journal of the International Organization of Psychophysiology*, **59**(2), 107-15. doi:10.1016/j.ijpsycho.2005.02.008

Widge, A. S., Bilge, M. T., Montana, R., Chang, W., Rodriguez, C. I., Deckersbach, T., . . . Nemeroff, C. B. (2019). Electroencephalographic biomarkers for treatment response prediction in major depressive illness: A meta-analysis. *The American Journal of Psychiatry*, **176**(1), 44-56. doi:10.1176/appi.ajp.2018.17121358

Williams, L. M., Debattista, C., Duchemin, A. M., Schatzberg, A. F., & Nemeroff, C. B. (2016). Childhood trauma predicts antidepressant response in adults with major depression: Data from the randomized international study to predict optimized treatment for depression. *Translational Psychiatry*, **6**, e799. doi:10.1038/tp.2016.61

Williams, L. M., Rush, A. J., Koslow, S. H., Wisniewski, S. R., Cooper, N. J., Nemeroff, C. B., . . . Gordon, E. (2011). International study to predict optimized treatment for depression (iSPOT-d), a randomized clinical trial: Rationale and protocol. *Trials*, **12**, 4. doi:10.1186/1745-6215-12-4

Williams, L. M., Simms, E., Clark, C. R., Paul, R. H., Rowe, D., & Gordon, E. (2005). The test-retest reliability of a standardized neurocognitive and neurophysiological test battery: "Neuromarker". *Int J Neurosci*, **115**(12), 1605-30. doi:10.1080/00207450590958475

Woolley, S. B., Cardoni, A. A., & Goethe, J. W. (2009). Last-observation-carried-forward imputation method in clinical efficacy trials: Review of 352 antidepressant studies. *Pharmacotherapy*, **29**(12), 1408-16. doi:10.1592/phco.29.12.1408

Wu, W., Zhang, Y., Jiang, J., Lucas, M. V., Fonzo, G. A., Rolle, C. E., . . . Etkin, A. (2020). An electroencephalographic signature predicts antidepressant response in major depression. *Nat Biotechnol*. doi:10.1038/s41587-019-0397-3

Yasuhara, A. (2010). Correlation between EEG abnormalities and symptoms of autism spectrum disorder (ASD). *Brain Dev*, **32**(10), 791-8. doi:10.1016/j.braindev.2010.08.010

Zimmerman, M., Martinez, J. H., Young, D., Chelminski, I., & Dalrymple, K. (2013). Severity classification on the hamilton depression rating scale. *Journal of Affective Disorders*, **150**(2), 384-388.

Zotev, V., Yuan, H., Misaki, M., Phillips, R., Young, K. D., Feldner, M. T., & Bodurka, J. (2016). Correlation between amygdala BOLD activity and frontal EEG asymmetry during real-time fmri neurofeedback training in patients with depression. *NeuroImage. Clinical*, **11**, 224-38. doi:10.1016/j.nicl.2016.02.003

ENGLISH SUMMARY

The symptoms of major depressive disorder (MDD) include a depressed mood most of the day (nearly every day), loss of interest in daily activities, weight gain or diet independent weight loss, fatigue, and feelings of worthlessness or guilt (as determined by the Diagnostic and Statistical Manual of Mental Disorders 5). A long-lasting course characterizes this mood disorder, which in many cases turns into a chronic problem. After a long history of describing, explaining and treating this disorder, patients and their doctors are still facing the challenges of beating this disease, despite research efforts.

MDD is commonly treated with antidepressant medication (AD) or psychotherapies like cognitive behavioral therapy. Clinical efficacy rates unfortunately only reach 37% remission after a first AD prescription, with declining remission rates after each consecutive AD trial. The large (and continuously growing) prevalence of MDD in the world is disconcerting: 183 million people had MDD in 2005, up to 216 million in 2015. The development of new AD medication is subjected to suspended research and development budgets for central nervous system drugs, including ADs. Since a growing group of people affected by this disease, new approaches to the treatment of MDD are needed to serve this group.

With the aim to improve treatment outcome in depressed patients, we described our first steps towards the implementation of an EEG biomarker informed protocol. We zoomed in on detailed characteristics of biomarkers that proved to be promising. We attempted to utilize automated processes for fast, professionalized EEG assessments. We developed a protocol in which all knowledge on biomarker informed AD prescription was implemented, and performed a feasibility trial. We also compared protocol outcomes with the results of a control group.

Chapter 2 provides an up to date meta-analysis on the diagnostic value of the biomarker frontal alpha asymmetry (FAA) in MDD, and the evaluations of discrepancies in a large cross-sectional dataset. Sixteen studies were included (MDD: n = 1883, controls: n = 2161). The main result was a non-significant, negligible ES, demonstrating limited diagnostic value of FAA in MDD. The high degree of heterogeneity across studies indicates covariate influence, as was confirmed by cross-sectional analyses.

Chapter 3 explores the stability of biomarker FAA, that was demonstrated in earlier, smaller studies. In patients with MDD, FAA did not change significantly after eight weeks of treatment (n = 453, p = .234), nor did we find associations with age, sex, depression severity, or change in depression severity. We demonstrate that FAA is a stable trait, robust to time, state and pharmacological status. This confirms FAA stability.

Chapter 4 explores whether depressed patients with an abnormal EEG show a normalization of the EEG related to AD treatment and response, and whether such effect is drug specific. In fifty-seven patients with subclinical EEG abnormalities, the EEGs did not normalize significantly more with sertraline compared to the other ADs, escitalopram and venlafaxine. However, response rates in patients with normalized EEG taking sertraline were 5.2 times (significantly) larger than in subjects treated with escitalopram/venlafaxine.

Chapter 5 shows the utilization of computed features, to improve the assessment of abnormal EEGs in the depressed population, and compare them to our previous methods. The computed features CNN probability, the dominant frequency, and the tBSI all successfully showed good performance in identifying the specific and notably "light" abnormalities. A random forest model containing the combined features, trained on predicting treatment outcome per AD drug, did not reliably predict treatment outcome.

Chapter 6 evaluates the results from a first prospective feasibility trial. The EEG biomarkers abnormal EEG activity, alpha peak frequency, and FAA were prospectively used for EEG informed prescription

of ADs. Seventy patients were stratified to AD based on their EEG biomarkers, 52 patients received AD treatment as usual. In general, both professionals and stratified patients were satisfied with the new protocol and practical implementation proved to be feasible, with better symptom improvement in patients who received EEG informed prescribed ADs using EEG biomarkers.

NEDERLANDSE SAMENVATTING

Bij een depressieve stoornis horen een aantal symptomen, waaronder een depressieve stemming gedurende een groot gedeelte van de dag (bijna elke dag), verlies van interesse in alledaagse activiteiten, gewichtstoename of dieet-onafhankelijk gewichtsverlies, vermoeidheid, en het gevoel van waardeloosheid of schuld (zoals beschreven in de Diagnostic and Statistical Manual of Mental Disorders 5). Deze stemmingsstoornis wordt gekenmerkt door een langdurig verloop, dat in veel gevallen uiteindelijk een chronisch probleem wordt. Na een lange geschiedenis van beschrijven, verklaren en behandelen van de stoornis, lopen patiënten en hun behandelaren nog altijd aan tegen de uitdaging om de stoornis aan te pakken, ondanks het vele onderzoek.

Depressie wordt doorgaans behandeld met antidepressiva (AD) of psychotherapieën zoals cognitieve gedragstherapie. Klinische effectiviteit gemeten door remissie, laat echter percentages zien van ten hoogste 37% remissie na een eerste AD, waarna dit percentage zakt na elke opvolgende poging met een AD. De grote (en nog altijd groeiende) wereldwijde depressieprevalentie is verontrustend: 183 miljoen mensen kregen een depressie in 2005, tot aan 216 miljoen in 2015. De ontwikkeling van nieuwe antidepressiva is momenteel grotendeels stilgelegd, budgetten voor dergelijk onderzoek zijn er nauwelijks meer. Omdat een groeiende groep mensen is aangedaan door deze stoornis, zijn er nieuwe vormen van onderzoek nodig om deze groep te bedienen.

Met het doel om de behandeluitkomst voor depressieve patiënten te verbeteren, hebben we in dit proefschrift onze eerste stappen beschreven naar de implementatie van een protocol gebaseerd op EEG-biomarkers. Een biomarker is een meting in het lichaam die informatie geeft over of een persoon een ziekte heeft, of voor het bepalen van de juiste behandeling. Het elektro-encefalogram (ofwel EEG), meet elektrische activiteit in de hersenen door elektroden. In onze studies leg-

den we de focus op de gedetailleerde eigenschappen van veelbelovende biomarkers. We probeerden gebruik te maken van geautomatiseerde processen voor snelle en geprofessionaliseerde EEG-ver-werking. We ontwikkelden een protocol waarin alle kennis over het voorschrijven van AD's gebaseerd op EEG-biomarkers werd geïmplementeerd, en voerden een haalbaarheidsstudie uit. We vergeleken hierbij ook de uitkomst van ons protocol met de resultaten van een controlegroep.

Hoofdstuk 2 geeft een actuele meta-analyse van de diagnostische waarde van de biomarker frontale alfa asymmetrie (FAA) voor depressie. Ook worden de verschillen en afwijkingen in een grote cross-sectionele dataset geëvalueerd. Zestien studies werden geïncludeerd (depressie: $n = 1883$, controles: $n = 2161$). Het hoofdresultaat was een niet-significante, verwaarloosbare effect size, wat aantoont dat FAA weinig diagnostische waarde heeft in depressie. De hoge mate van heterogeniteit in de verschillende studies geeft aan dat er invloed is van andere factoren (covariabiliteit), wat bevestigd werd met cross-sectionele analyses.

Hoofdstuk 3 verkent de stabiliteit van de biomarker FAA, wat al eerder is aangetoond in kleine studies. In depressiepatiënten veranderde FAA niet na acht weken medicatie ($n = 453$, $p = .234$). Ook vonden we geen relatie met leeftijd, geslacht, de ernst van de depressie, of verandering in de ernst van de depressie. We tonen hiermee aan dat FAA een stabiele eigenschap is (trait), die niet veel beïnvloed kan worden door tijd, de toestand van een persoon en medicatie. Dit bevestigt de stabiliteit van FAA.

Hoofdstuk 4 onderzoekt of depressieve patiënten met een licht afwijkend EEG een normalisatie in hun EEG laten zien dat gerelateerd is aan de behandeling met een AD en respons, en of dat effect voorkomt bij één specifiek AD. In 57 patiënten met een afwijkend EEG, normaliseerde het EEG niet méér met sertraline, vergeleken met andere AD's, escitalopram en venlafaxine. Echter, de behandelrespons van patiënten met een genormaliseerd EEG die een behandeling met sertraline kregen was 5,2 keer (significant) hoger dan in patiënten die behandeld werden met escitalopram of venlafaxine.

Hoofdstuk 5 laat zien hoe we verschillende kenmerken in het EEG berekenden met verschillende methodes, voor het verbeteren van het vinden van afwijkingen in EEG's in de depressieve populatie. Deze kenmerken vergeleken we met onze voorgaande methodes gebaseerd op met name visuele beoordeling. Het convolutional neural network (deep learning), de dominante frequentie, en de temporele brain symmetry index registreerden alle drie meer vertraging of afwijkingen in de groep die we visueel al hadden geclassificeerd als het hebben van "lichte" afwijkingen in het EEG, vergeleken met een groep waarin we die afwijkingen visueel niet eerder vonden. Een random forest model dat alle berekende kenmerken bevatte, werd getraind op het voorspellen van behandeluitkomst per AD. Dit model bleek de behandeluitkomst niet te kunnen voorspellen.

Hoofdstuk 6 evalueert de resultaten van een eerste prospectieve haalbaarheidsstudie. De EEG-biomarkers afwijkende EEG activiteit, alfapiek frequentie, en FAA werden prospectief gebruikt voor het voorschrijven van AD's. Zeventig patiënten werden gestratificeerd naar verschillende AD's, gebaseerd op hun biomarkers, 52 patiënten kregen behandeling zoals gebruikelijk. Over het geheel genomen waren zowel de zorgprofessionals als de gestratificeerde patiënten tevreden met het nieuwe protocol. De praktische implementatie bleek voldoende haalbaar te zijn. Er werd significant meer verbetering in symptomen gezien in patiënten die AD's voorschreven kregen op basis van EEG-biomarkers.

CONCLUSIE

De complexiteit van veel verschillen tussen depressieve patiënten (heterogeniteit) maakt het onmogelijk om één behandeling te vinden die voor iedereen werkt. We kunnen deze heterogeniteit echter omarmen door biomarkers te vinden die het mogelijk maken om homogene subgroepen te identificeren, die mogelijk behandeluitkomst kunnen helpen voorspellen. We kunnen meer betrouwbaarheid en kwalitatieve verbetering bereiken door de allernieuwste technieken te gebruiken, alhoewel deze methodes nog verder ontwikkeld moeten worden voordat ze voor patiënten gebruikt kunnen worden. Aan de andere kant zijn er al kwantitatieve methodes die het mogelijk

lijken te maken om behandeluitkomst te voorspellen, gebaseerd op frontale alfa lateralisatie en alfapiek frequentie. Gecombineerd met de kwalitatieve beoordeling van EEG-afwijkingen in ons nieuwe protocol, laten de eerste resultaten zien dat er voldoende haalbaarheid is in een klinische setting. Voor zover wij weten, is dit de eerste poging om de behandeling van depressie te verbeteren via deze biomarkers, die niet alleen aantoonde dat het protocol niet slechter is dan reguliere behandeling: patiënten laten een significante verbetering zien in vergelijking met medicatie zoals door de psychiater voorgeschreven. Ons voorgestelde protocol maakt het daarmee niet alleen mogelijk om onze methodes naar de klinische praktijk te brengen, het draagt ook de belofte van een kleine maar nodige verbetering van onze behandelnormen, in deze nieuwe vorm van neuropsychiatrische gezondheidszorg voor depressie.
Ik hoop dat dit proefschrift nieuw onderzoek bevordert, dat doorgaat met het focussen op het maken van behandelbeslissingen die geïnformeerd zijn door EEG-biomarkers. Onze studies laten veelbelovende resultaten zien, die de weg vrij maken naar gestratificeerde psychiatrie, waarmee een nog altijd groeiende groep mensen met depressie geholpen kan worden.

ABOUT THE AUTHOR

Nikita van der Vinne was born on January 20th 1987 in Deventer. After graduating from her high school Etty Hillesum Lyceum in 2005, she studied psychology at the University of Groningen. In 2010, she traveled to Ethiopia for a half year research internship in social psychology. After getting acquainted with the EEG as a research methodology in psychology during her bachelor's, she decided to focus on neuropsychology for her master's. At Stichting Epilepsie Instellingen Nederland (a hospital for epilepsy patients) she performed her clinical internship, and remained working for a little while, as a neuropsychologist and psychological evaluator.

Always remaining doubtful about whether to work in research or the clinic, and when a new job had to be found, she decided to switch to a research project as a PhD student in 2015. Synaeda Psycho Medisch Centrum offered a position within an outpatient clinic, giving the chance to stay in touch with clinical work, while having the opportunity to receive scientific guidance from Research Institute Brainclinics. In this collaboration, the plan to investigate EEG based prescription of antidepressants was formed with co-promotor dr. Martijn Arns, and was further developed during Nikita's PhD project. Dr. Madelon Vollebregt joined as senior researcher and co-promotor soon hereafter. After a startup year, prof. dr. ir. Michel van Putten from the University of Twente was asked to join the project as Nikita's promotor, as the knowledge from himself and his department of Clinical Neurophysiology would enrich the project. After four and a half years, the project was finished with this dissertation and new projects are being formed. As a psychologist and senior researcher at Synaeda, Nikita is now able to perform work in both fields, co-supervising new PhD students in extending the EEG project and starting a new NCG-TMS project, and treating patients with depression herself.

LIST OF PUBLICATIONS

PUBLISHED

Van der Vinne, N., Vollebregt, M. A., van Putten, M. J. A. M., & Arns, M. (2017). Frontal alpha asymmetry as a diagnostic marker in depression: Fact or fiction? A meta-analysis. *NeuroImage: Clinical*, **16**, 79-87.

Van der Vinne, N., Vollebregt, M. A., Boutros, N. N., Fallahpour, K., van Putten, M. J. A. M., & Arns, M. (2019). Normalization of EEG in depression after antidepressant treatment with sertraline? A preliminary report. *Journal of Affective Disorders*, **259**, 67-72.

Van der Vinne, N., Vollebregt, M. A., van Putten, M. J. A. M., & Arns, M. (2019). Stability of frontal alpha asymmetry in depressed patients during antidepressant treatment. *NeuroImage: Clinical*, **24**.

IN PREPARATION / SUBMITTED

Van der Vinne, N., Vollebregt, M. A., Rush, A. J., M. A., Eebes, M., van Putten, M. J. A. M., & Arns, M. EEG biomarker informed prescription of antidepressants in MDD: a feasibility trial. Under review at *European Neuropsychopharmacology*.

Van der Vinne, N., Tjepkema-Cloostermans, M. C., da Silva Lourenço, C., Vollebregt, M. A., Arns, M., & van Putten, M. J. A. M. Computer assisted EEG abnormality detection and treatment response prediction in major depressive disorder. *In preparation*.

OTHER PUBLICATIONS

Hansen, N., Postmes, T., van der Vinne, N., van Thiel, W. (2012). Information and communication technology and cultural change: How ICT changes self-construal and values. *Social Psychology*, **43**(4), 222-231.

CONFERENCE CONTRIBUTIONS AND OTHER PRESENTATIONS

ORAL PRESENTATIONS

EEG-onderzoek in depressie. De Vijf Wetenschappelijke Zonden, *bijeenkomst Friese psychiaters*, January 2017, Weidum.

Personalized Medicine in Depressiebehandeling: EEG-onderzoek bij Synaeda. Depressiebehandeling net even anders!, *Farmacotherapeutisch Overleg, GGZ Friesland*, June 2018, Leeuwarden.

Temporal Development of Depression Biomarkers From the iSPOT-D Study: State or Trait. *73rd Annual Meeting Society of Biological Psychiatry SOBP*, May 2018, New York. For abstract see: B*iological Psychiatry* **2018**;83(9):136.

Hersenen, antidepressiva en CGT. *Dag van de Inhoud, Synaeda Psycho Medisch Centrum*, October 2018, Eernewoude.

Temporal Development of Depression Biomarkers from the iSPOT- D study: State or Trait. *20th biennial International Pharmaco-EEG Society IPEG Meeting*, November 2018, Zürich. See for abstract: *Neuropsychobiology* 2019;**77**(3):S2-3.

EEG-biomarkers in depressiebehandeling en implementatie in een reguliere gezondheidszorginstelling. Behandeling onder de loep!, *NedKAD-conferentie*, October 2019, Amersfoort.

POSTER PRESENTATIONS

Frontal alpha asymmetry in depression: fact or fiction? A meta-analysis. *19th biennial IPEG Meeting*, October 2016, Nijmegen. For abstract see: *Neuropsychiatric Electrophysiology* 2016;**2**(8):A85.

DANKWOORD

Er zijn ontzettend veel mensen betrokken geweest bij het onderzoekswerk dat ik heb uitgevoerd, die ik allemaal erg graag wil bedanken.

Dit proefschrift was nooit tot stand gekomen als er geen ambitieuze samenwerkingsplannen waren ontstaan tussen Synaeda en Brainclinics. **Martijn Arns**, **Anke Oosterbaan** en **Bianca Klop**, jullie gaan ver buiten de gebaande paden om betere zorg aan patiënten te bieden. Dit was erg inspirerend voor me, en ik dank jullie voor de mogelijkheid dit onderzoek te hebben mogen doen. Martijn, ik heb onder jouw vleugels ontzettend veel geleerd en ik ben ver buiten mijn comfortzone gestapt. Je visie inspireerde me, en ik ben dankbaar dat ik via jou kennis heb gemaakt met de onderzoekswereld. Anke en Bianca, de ruimte die jullie hebben geboden voor wetenschap is zo ontzettend belangrijk, gelukkig komt er een vervolg!

De samenwerking met mijn promotor **Michel van Putten** en de afdeling Clinical Neurophysiology van de Universiteit Twente heeft het mogelijk gemaakt vakgebieden samen te brengen die niet altijd makkelijk samengaan. De moeilijk grijpbare wetenschap van psychiatrische problematiek bij mensen maakt het lastig technische benaderingen geschikt te maken voor betere zorg. Ik dank je voor het openstaan voor de compromissen die er nodig zijn om deze vakgebieden samen te brengen, voor het mogen delen in je kennis en kunde, en voor het terechte wijzen op de enorme ruimte voor verbetering in diverse situaties waarin ik dacht dat ik toch echt al een heel eind was.

Ook alle **proefpersonen**, bedankt voor jullie bereidheid om mee te doen aan mijn onderzoek, wat niet altijd makkelijk was met een depressie. “Een steentje bijdragen aan wetenschappelijk onderzoek” is wat erg veel van jullie aangaven als motivatie, het was inspirerend om zoveel betrokkenheid te voelen.

Prof. Boutros, **prof. van Wingen**, **dr. Bailey**, **prof. Van der Palen**, and **prof. Hermens**, thank you for reviewing my dissertation and taking place in my graduation committee. And Nash, meeting you at the beginning of my PhD project inspired me; your views on both psychiatry and neurology, your vast experience and dedication made me look beyond the numbers and the data, and made me see the people.

I also would like to thank all co-authors on several chapters in my dissertation, in particular **John Rush** and **Kamran Fallahpour**. Some of you I might never meet in person, but I am very grateful for being able to share in your knowledge and expertise. It helped me reach higher grounds.

Madelon Vollebregt, waar was ik geweest zonder jou! Je kennis, inzicht, meedenken, durf, humor, geheugen voor persoonlijke details die ik zelf niet eens onthoud, je oprechte aandacht en wat al niet meer... Op momenten dat ik het zelf niet meer wist, kon jij me precies daarheen leiden waar ik wilde zijn. Je bent op zoveel vlakken belangrijk geweest, ik kan je niet genoeg bedanken. En wat mooi dat je in mijn tijd wel twee nieuwe meisjes erbij hebt gekregen! De skippybal is inmiddels onlosmakelijk verbonden met onze samenwerking, die (en jouzelf!) ga ik missen ;)

Lieve **Michiel**, wat was ik blij toen je aansloot bij mijn tot-dan-toe-eenmansonderzoeksteam. Je kwam als geroepen; je hebt me erg veel werk uit handen genomen en samen hebben we nog veel meer proefpersonen weten te bereiken.

Het directiestokje binnen Synaeda werd overgenomen door **Bob Goeree** en **Vera Veerman**, en daarmee kwamen afgelopen jaar ook nieuwe plannen op tafel. Het succes van de EEG-studie binnen Synaeda is hopelijk de eerste van vele: jullie maakten het mogelijk om het onderzoek een vervolg te geven, en de voorbereidingen zijn al gaande. Ik wil jullie heel erg bedanken voor de mogelijkheid, voor de steun vanaf de zijlijn, en voor het vertrouwen.

Op de Synaedalocatie Fonteinland heb ik kunnen samenwerken met een fantastisch team dat vanaf het begin openstond voor het

EEG-onderzoek. **Ab**, **Agnes**, **Anna**, **Annemarie**, **Annemieke**, **Astrid**, **Bette**, **Chantal**, **Daniël**, **Dennis**, **Dina**, **Erwin**, **Esther**, **Femke Marije**, **Fred**, **Gea**, **Gerard**, **Gonnie**, **Ineke**, **Ingrid**, **Janina**, **Jannie**, **Jeltje**, **Jenny**, **Jet**, **Joëlle**, **Judith**, **Kirsten**, **Lianne**, **Lieuwe**, **Linda**, **Lize**, **Maaike**, **Machtelt**, **Margreet**, **Marianne**, **Marieke**, **Marjolein**, **Maud**, **Maurits**, **Nynke**, **Patricia**, **Rommy**, **Roos**, **Rosiena**, **Sanne**, **Stieneke**, **Trijntje**, **Ymie**; jullie bijdrage aan de planning, het meedenken en rekening houden met mijn nieuwe plannen, het verwijzen van patiënten, maar ook jullie enthousiasme, de (professionele) luisterende oren, jullie (soms heerlijk ongepaste) humor tijdens pauzes; alles heeft geholpen om me uiteindelijk thuis te voelen op Fonteinland en mijn onderzoek uit te voeren. En helemaal met jou, **Emma**, wat heerlijk om gewoon af en toe op alles te kunnen b*tchen, maar ook te toosten! Verder heb ik ook op de locaties **Sixmastraat** en **Drachten** veel gehad aan de interesse die jullie toonden en het luisterend oor voor mijn strubbelingen in de laatste periode. Iedereen, *tige tank*!

De vele treinkilometers die ik heb gemaakt in de afgelopen 4,5 jaar naar Nijmegen waren het meer dan waard om onderdeel te mogen zijn van het onderzoeksteam Brainclinics. Met de tijd groeide het team, en in die tijd heb ik veel gehad aan zowel jullie kennis als jullie warmte. De kerst- en paaslunches, onze congresreizen (ik zal New York nooit vergeten!), het turven van sirenes of het registreren van weersomstandigheden bij verkeerd-aanbellers, de rock van **Tabitha** en **Hanneke** en de Disneymuziek van **Noralie** (maar natuurlijk ook jullie gezelligheid), de humor en het prachtige design van **Mark** (ik heb volgehouden!), de droogkomische nuchterheid van **Tommy** als tegenhanger van al die vrouwen, de korte maar fijne samenwerking met **Lana**, **Charlotte** en **Helena**, het gezellige kletsen met **Nicole** en **neuroCare collega's** tijdens de lunch, en nog belangrijker: de waardevolle vanzelfsprekende hulp aan elkaar. Ik heb jullie allemaal erg hoog zitten en ga jullie missen!

Bij de vakgroep CNPH van de universiteit heb ik in de jaren weinig tijd doorgebracht, waar psychologie niet een alledaags werkveld is waar jullie mee te maken hebben. Toch heb ik me er welkom gevoeld. **Annika**, je bent altijd zo attent, en je interesse in mijn werk heeft me veel goed gedaan. Het gaf me een kleine kickstart voor kennismaking

met de groep. En **Marleen**, jouw betrokkenheid aan Michels zijde gaven mijn project de verbreding die het nodig had. Bedankt voor het delen van je kennis, en dat je me introduceerde in dit werkveld waar ik zo weinig van wist. **Tanja**, door de jaren heen heb je me op de achtergrond heel fijn bijgestaan. Ook je betrokkenheid (meestal via de mail) droeg bij aan het goed regelen en afronden van mijn traject. And **Catarina**, our short but very pleasant collaboration gave that little extra CNPH touch to my work, *obrigado*!

Lieve **papa**, jouw onvoorwaardelijke steun, geloof en vertrouwen hebben me er meer dan eens doorheen gesleept, niet alleen in mijn werk. En met **Jannie** aan je zijde werd dat alleen maar sterker. Het is zo verdrietig dat **mama** er niet kan zijn om daarin te delen. De glimmende trots die ze had bij mijn dansoptredens ga ik ongelooflijk missen op de dag van de ceremonie.
En lieve **Maya**, terwijl ik mijn eigen groei doormaakte de afgelopen 4,5 jaar, deed jij hetzelfde. Wat ben ik ongelooflijk trots op je! Nu is er eindelijk meer tijd om elkaar op te zoeken ;) En kunnen we samen **Anthony** proberen in te maken met Ticket to Ride!

In mijn eentje had ik dit project niet durven aangaan. De aansporing en motivatie die ik van **Liekele** kreeg zijn hierin mede bepalend geweest, en ik wil deze kans aangrijpen om aan te geven dat ik je hier heel dankbaar voor ben.

Daarnaast heb ik door de jaren heen kunnen steunen op, en genieten van zoveel familie en vrienden. Ik stort nogal makkelijk mijn hart bij iedereen uit, en daar hebben jullie allemaal een lief, luisterend oor voor gehad. Hoewel er nog veel meer mensen zijn, wil ik in het bijzonder **Thijs**, **Andrea**, **Oma**, **Michèle**, **Ed**, **Mira**, **Romain**, **Marten**, **Alie**, **Maaike**, **Resi** en **Christina** bedanken. Voor alle borrels, etentjes, het genieten van jullie kleintjes, het theeleuten en de logeerpartijen: dankjulliewel!

Brainclinics®
INSIGHTS

ISBN 978-90-830013-3-3
9 789083 001333