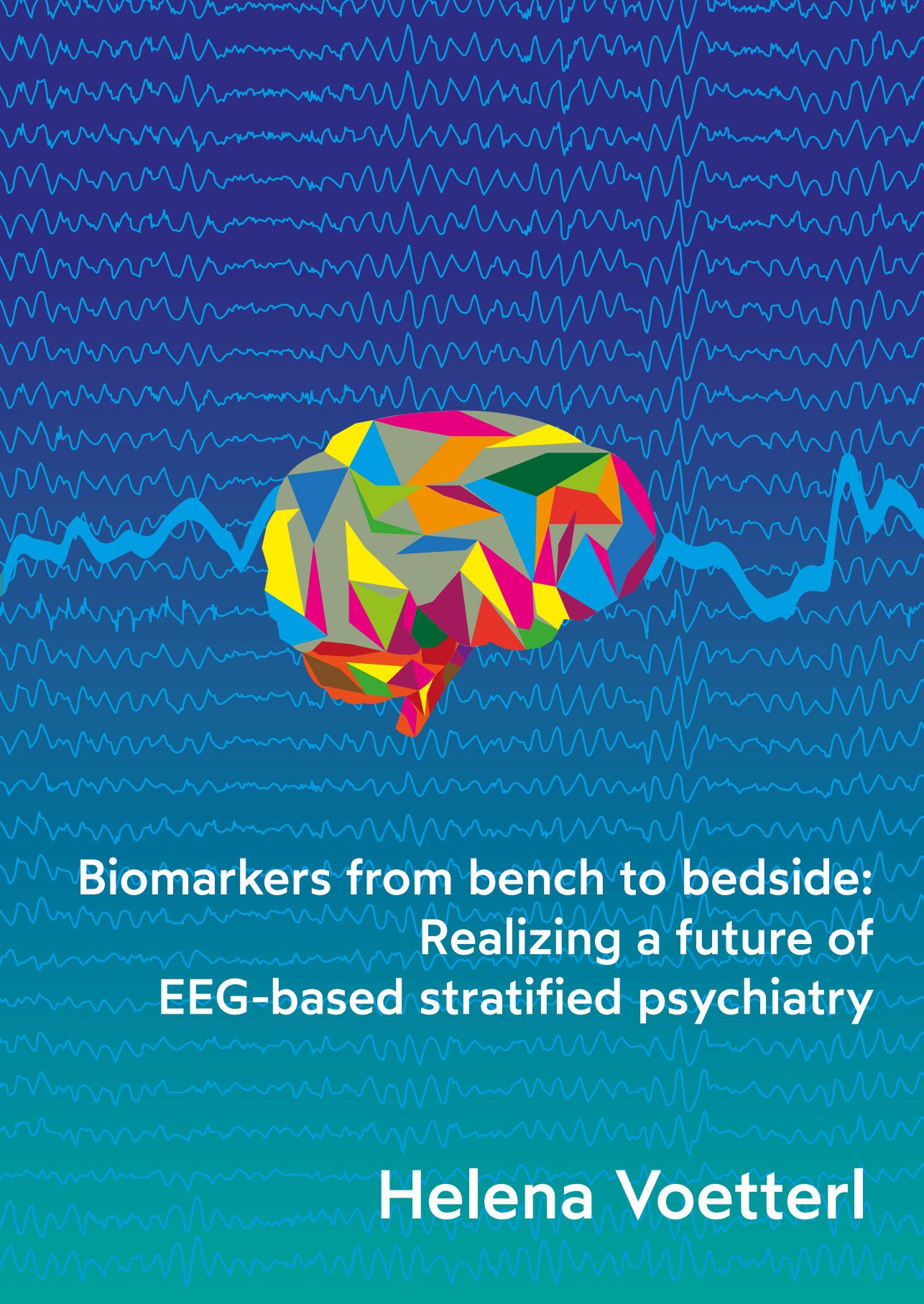


Biomarkers from bench to bedside: Realizing a future of EEG-based stratified psychiatry

Helena Voetterl

BIOMARKERS FROM BENCH TO BEDSIDE:

Realizing a future of EEG-based stratified Psychiatry

Helena Theresa Sofia Voetterl

The work described in this doctoral dissertation was carried out at Stichting Brainclinics Foundation, in collaboration with Maastricht University.

Voetterl, Helena Theresa Sofia
Biomarkers from Bench to Bedside:
Realizing a Future of EEG-Based
Stratified Psychiatry

Published by Brainclinics Insights,
Nijmegen, The Netherlands

Cover and layout: Mark Koppenberg, Velp
ISBN-978-90-830013-5-7

BIOMARKERS FROM BENCH TO BEDSIDE:

REALIZING A FUTURE OF

EEG-BASED STRATIFIED

PSYCHIATRY

DISSERTATION

to obtain
the degree of Doctor at Maastricht University,
on the authority of the Rector Magnificus,
Prof. dr. Pamela Habibović,
in accordance with the decision of the Board of Deans,
to be defended in public on
Tuesday, the 25th of June 2024, at 10:00 hrs.

by

Helena Theresa Sofia Voetterl
born on the 24th of August 1993
in Berlin, Germany

Supervisor
Prof. dr. Alexander Sack,
Maastricht University

Co-supervisors
Dr. Martijn Arns,
Brainclinics Foundation, Nijmegen/
Maastricht University

Dr. Hanneke van Dijk,
Synaeda Research, Drachten/
Maastricht University

Assessment Committee
Prof. dr. Bernadette Jansma (Chair),
Maastricht University
Prof. dr. Flavio Frohlich,
University of North Carolina Neuroscience Center, Chapel Hill, NC, US
Prof. dr. David E.J. Linden,
Maastricht University
Prof. dr. Leanne Williams,
Stanford University School of Medicine , US

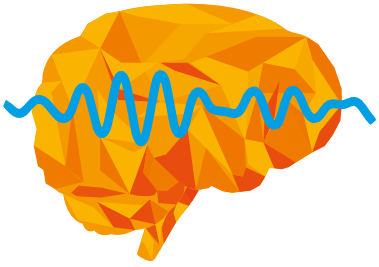


TABLE OF CONTENTS

1

INTRODUCTION

THE BURDEN OF DEPRESSION

Early accounts from the 2nd millennium B.C. in the first Babylonian empire already documented a depression-like state that was thought to be caused by demons or gods (Reynolds & Wilson, 2013). However, it was not long before Hippocrates (460-377 B.C.) ventured a theory of "melancholia" rooted in biology, that presumed its origin in the brain. He described it as "fear or sadness that last a long time" coupled with "aversion to food, despondency, sleeplessness, irritability, [and] restlessness" - a description that resembles modern accounts of depression (DeRubeis & Strunk, 2017). Nowadays it is well known that depression is a mental health condition, distinct from sadness or grief, that is characterized by high symptom heterogeneity. Different symptom profiles can be contradictory, for instance a depressed patient suffering from sleeping problems can either get too little sleep or spend too much time in bed. In the Diagnostic and Statistical Manual of Mental Disorders (DSM) that aids in the diagnosis of mental disorders, this heterogeneity has mostly been disregarded and different depression symptom profiles used to be bundled under one diagnosis (Allsopp et al., 2019). In recent versions of the DSM more ef-

fort has been made to delineate distinct depression subtypes. Although it is generally assumed that various depression subtypes with different etiologies exist (Cai et al., 2020), the evidence for explicit subtypes is still lacking.

Depression in general is both highly prevalent globally and results in immense suffering. Approximately 4.4% of all people worldwide are affected (World Health Organization, 2017), and 10.6% of people will experience a depressive episode at least once throughout their lifetime (Herrman et al., 2022) although this lifetime prevalence might be severely underestimated and could in fact be as high as 30-40% (Moffitt et al., 2010). Prevalence is highest in older adulthood (55-74 years) and women are more affected than men (World Health Organization, 2017). In addition to this high incidence, depression has an immense impact on quality of life, accounting for 50 million years lived with disability (YLD) in 2015, making it the disorder with the highest influence on non-fatal health loss (7.5% of all YLD) (World Health Organization, 2017). When more depressive symptoms are present and symptoms are more severe, a person is considered to have major depressive disorder (MDD) (Fils et al., 2010). MDD is often so incapacitating that many sufferers are unable to maintain employment, sustain relationships and retain daily functioning (Marwaha et al., 2023). In addition to greatly compromising the individual's life, the impact on health systems and social and economic costs are substantial (McLaughlin, 2011).

The longer a depression lasts and the more severe it is, the more likely it becomes that a person ends their own life (Hawton et al., 2013; Lépine & Briley, 2011). Accounting for more than 1 in 100 deaths, suicide is one of the leading causes of death worldwide, with a particularly high ranking in young people (World Health Organization, 2021). Although suicide does not only occur in depression, people with MDD are 20-fold more likely to die of suicide than the general population (Ösby et al., 2001). In addition, the risk of dying of other causes is increased multifold. One reason for that is the higher likelihood to develop a comorbid disease, such as coronary artery disease or obesity. At the same time, having a depression increases the risk of cardiac mortality in those with coronary heart disease, and this risk is elevated with higher depression severity (Kupfer et al., 2012).

When it comes to depression treatments, psychotherapy and pharmacotherapy are considered the gold standard of first-line interventions. Selective serotonin-reuptake inhibitors (SSRIs) work - as their name suggests - by restricting the serotonin transporter protein and thus the transport of serotonin back into the presynaptic cell, thereby enhancing serotonin availability in the synaptic cleft (Sangkuhl et al., 2009). Antidepressants, such as SSRIs and serotonin-norepinephrine reuptake inhibitors (SNRIs), have been proven effective in treating depression (Arroll et al., 2016) although it remains unclear whether serotonin is involved in the etiology of depression or whether the effect is driven by other mechanisms, such as increased expression and signaling of brain-derived neurotrophic factor and other growth factors in hippocampal and cortical areas (Moncrieff et al., 2023; Björkholm, C., & Monteggia, L.M., 2016; Castrén, E., & Monteggia, L. M., 2021). Despite efficacy at the group level, one of the largest and most influential randomized controlled trials (RCT) in depression, the Sequenced Treatment Alternatives to Relieve Depression (STAR*D) trial, concluded that only about a third of MDD patients remitted following an adequately dosed treatment course of Citalopram (Rush et al., 2006). Approximately 50% of patients remitted after two treatment steps, although an increased relapse rate was observed after the second treatment course than among patients who remitted after only one course. Overall, a theoretical cumulative remission rate of 67% after 4 treatment steps was reported although some criticism has been raised, proposing even lower cumulative remission rates (Pigott, 2014; Pigott et al., 2023). Beyond two treatment attempts, the chances to remit drastically decreased (~13%). Rush et al. concluded that about a third of patients do not achieve remission after as many as 4 adequately dosed antidepressant courses. These patients are considered to have a treatment-resistant depression (TRD) (Gaynes et al., 2020), or – more appropriately termed – a difficult-to-treat depression (DTD) (Wilhelm, 2019). The STAR*D investigators hypothesized that these patients might have benefitted from treatment administered earlier in the course of illness, before chronicity had emerged, or from an accelerated transition to a more intensive treatment approach, such as noninvasive brain stimulation (NIBS). According to the critical reanalysis of the STAR*D trial, the prevalence of such DTD patients might be even higher than initially suggested by STAR*D (Pigott et al., 2023).

NONINVASIVELY STIMULATING THE BRAIN

One intervention that has proven effective for relieving symptoms of DTD is repetitive transcranial magnetic stimulation (rTMS). TMS induces a strong, rapidly changing magnetic field delivered noninvasively to the underlying cortex via a magnetic coil placed on the head (Barker et al., 1985; Roth et al., 1991). The magnetic pulses pass the skull unimpededly and create an electric field through electromagnetic induction. This electric field leads to a change in synaptic activity in the stimulated cortical tissue, resulting in excitation or inhibition of neurons (Huerta & Volpe, 2009; Jahanshahi & Rothwell, 2000). Early uses of TMS included the mapping of the motor cortex and investigation of cognitive functions (Jahanshahi & Rothwell, 2000; A Pascual-Leone et al., 1994; Alvaro Pascual-Leone & Hallett, 1994). First case studies and small pilot trials examining TMS for the treatment of depression were published as early as 1993 (Höflich et al., 1993), with promising effects in non-psychotic depression (George et al., 1995; Pascual-Leone et al., 1996). Already in these early studies, the prefrontal cortex (PFC) was the major focus of stimulation and remained it till this day.

Theories claiming a hemispheric imbalance of left hypo- and right hyperactivity in the PFC of depressed people guided application of TMS, applying high frequency rTMS, that was considered to have excitatory effects, over the left PFC, and low frequency, thought to be inhibitory, over the right PFC (Henriques & Davidson, 1991; Soares & Mann, 1997). Although both the imbalance hypothesis and the theory that high frequency stimulation is necessarily excitatory and low frequency stimulation inhibitory have been disproven since (Fitzgerald et al., 2006; van der Vinne et al., 2017), rTMS is still mainly applied over the dorsolateral prefrontal cortex (DLPFC) using the same high and low frequency protocols. One well-founded hypothesis why these protocols have been working in ameliorating depression symptoms despite the original hypotheses being incorrect is related to dysfunctions in brain networks with nodes in the DLPFC (Downar & Daskalakis, 2013; Williams et al., 2023). Although its precise working mechanisms are still unknown, TMS (over the cortex) has been shown to modify activity in subcortical regions as well as cortico-subcortical functional connectivity. It is hypothesized that these changes normalize activity within and between networks,

such as default mode network hyperactivity (Liston et al., 2014; Siebner et al., 2022), while the long-term effects of TMS are presumed to be induced by different plasticity mechanisms (Jannati et al., 2023; Thomson et al., 2020).

Since the early cognitive and clinical studies, research and clinical implementation of rTMS have been on the rise worldwide, with the number of publications increasing steeply and continuously (Matsuda et al., 2021; McLean, 2019; Rossi et al., 2009). Nevertheless, the first approval by the food and drug administration (FDA) was a long time coming; only in 2008 the Neuronetics Neurostar System was approved for the treatment of MDD patients who had failed prior antidepressant treatment (Cohen et al., 2022). Since then, many companies have followed suit, resulting in 24 TMS device approvals with different stimulation protocols and for several psychiatric disorders, such as obsessive-compulsive disorder (OCD). Besides MDD, rTMS has been investigated for the treatment of a multitude of other disorders, such as substance use disorder, post-traumatic stress disorder (PTSD), schizophrenia or Alzheimer's disease with varying levels of success (Cotovio et al., 2023; Lefaucheur et al., 2020).

In MDD, safety and treatment efficacy have been demonstrated in numerous systematic reviews, meta-analyses and even umbrella reviews evaluating multiple meta-analyses (Berlim et al., 2014; Brunoni et al., 2016; Fitzgerald et al., 2022; Razza et al., 2020) with most common side effects being headaches and scalp discomfort at the stimulation site (Milev et al., 2016). Although epileptic seizures are often mentioned in relation to rTMS, these are considered to be related to rTMS protocols exceeding safety guidelines and occur rarely (0.01-0.1%), at a rate comparable to or even lower than that of most psychotropic medication (Rossi et al., 2009, 2021; Steinert & Fröscher, 2017). Treatment efficacy of rTMS consistently outperforms sham with estimated response rates of about 50% and remission rates of 30% in RCTs, while effectiveness in real-world settings is generally observed to be higher (Carpenter et al., 2012; Donse et al. 2018). Moreover, all available and tested rTMS protocols seem to have similar efficacy at the group level (Berlim et al., 2013; Blumberger et al., 2018; J. Chen et al., 2014; Kedzior et al., 2014; Kishi et al., 2023). When

recalling the STAR*D finding that the likelihood to remit to another antidepressant medication after two failed antidepressant treatment attempts was approximately 13%, and taking into account that these rTMS findings are reported in such medication-resistant patients, remission rates of 30% can be considered rather impressive. However, despite this good overall efficacy of rTMS at the group level, 50-70% of patients achieve only partial or no symptom relief from rTMS.

In recent years, much research has focused on enhancing those remission and response rates, utilizing different approaches. Accelerated TMS (aTMS), administering multiple sessions per day, has proven to be effective and feasible while reducing time demands for patients and simultaneously enabling application of more sessions than is customary during a standard treatment course (Cole et al., 2021). Whether aTMS actually speeds up antidepressant effects remains an open question, as evidence is inconclusive (L. Chen et al., 2023). Deep TMS (dTMS) was developed to penetrate deeper brain regions that are implicated in causing depressive symptoms. Enhanced stimulation of white matter tracts might facilitate engagement of subcortical targets thereby improving antidepressant effect (Zibman et al., 2021). DTMS has proven as effective as standard rTMS but not more effective (Levkovitz et al., 2015). Another, relatively novel approach is capable of synchronizing TMS pulses (sTMS) to the individual's brain oscillations. This approach is based on the assumption that the TMS effect is brain-state dependent, meaning for instance that oscillations have phases of differential excitability, making the respective brain region more or less receptive for input. This moreover suggests that stimulating at certain oscillatory phases results in more activation of connected brain regions which could improve TMS effect (Sack et al., 2023). While sTMS has proven valuable for research purposes, more research is needed to evaluate its potential for enhancement of treatment outcome. Finally, a low-tech solution to improving response rates proposes that delivering higher numbers of TMS sessions may benefit a slow response subgroup that does not achieve response or remission after a standard rTMS course of 20-30 sessions (Kaster et al., 2019; McDonald et al., 2011). However, administering ad libitum rTMS sessions is expensive and time-consuming and being able to predict who might benefit from more extensive treatment would be helpful.

Although many promising TMS developments are underway, no solution has been proposed that can reliably enhance remission rates at the group level.

MOVING TOWARDS PRECISION

Findings of increasing treatment resistance with fruitless treatment attempts - as shown by STAR*D for instance - together with dire statistics on increased suicide risk with longer-lasting disease course and the immense burden of depression, demonstrate the need to identify the right intervention and start treatment early on since this could potentially be life-saving (Fogel et al., 2006). Although these statistics are commonly known, treatments are usually still prescribed in a trial-and-error fashion, meaning that individual patient characteristics are not taken into account for treatment decisions. Usually this means following a stepped-care approach (e.g. the stepped care model proposed by the National Institute for Health and Care Excellence (Rivero-Santana et al., 2021) or by the Dutch Multidisciplinary Guideline for Depression (Meeuwissen et al., 2019)), trying out a standardized sequence of multiple milder, evidence-based treatments, such as SSRIs for depression, before escalating to more intensive interventions, such as rTMS or ECT (Arns et al., 2023). To move away from this rather ineffective system, precision psychiatry has become a major focus of research (Williams et al., 2023), aiming at improving overall treatment efficacy by prescribing treatments targeting the individual by means of so-called biomarkers. A biomarker is defined as a biological or behavioral characteristic "that is objectively measured and evaluated as an indicator of normal biological processes, pathogenic processes, or pharmacologic responses to a therapeutic intervention" (Biomarkers Definitions Working Group, 2001).

A neural circuit approach might aid in effectively moving towards a true individualistic treatment model by identifying distinct disease types, based on circuit dysfunction, that may lie outside or overlap the boundaries of traditional mental disorders, reflecting disorder heterogeneity (Goldstein-Piekarski et al., 2020; Williams & Hack, 2020). Important steps have been taken in identifying such imbal-

ances in neural circuits and consequently particular biotypes that might respond to one treatment rather than another (Scangos et al., 2023; Williams, 2016). However, bridging the gap between research and translation of that research to clinical routines has presented a major challenge for precision psychiatry. One reason for this, as has been mentioned before, might be the rather outdated DSM and International Classification of Disease (ICD) that group disorders too broadly to adequately account for disorder heterogeneity (Williams et al., 2023). However, initiatives such as the Research Domain Criteria (RDoC) provide a framework that fosters the use of heterogeneity in psychopathology with malleable definitions of disorders that are able to evolve with new research findings (Williams et al., 2023). Through such neural circuitry approaches, precision psychiatry might eventually be able to match each patient to an individualized intervention, irrespective of whether this treatment is indicated for the patient's disorder. At present, a more practical way of achieving personalization is through treatment stratification, which, based on a shared characteristic, identifies patient subgroups that are more likely to respond to an evidence-based, approved treatment option for the given condition. At the same time, reduced likelihood to respond to a given treatment is taken into account, thereby effectively diminishing the number of available options and, thus, increasing chances to respond (Arns et al., 2022; Grzenda & Widge, 2023). Since stratification options are limited to approved, effective treatments for the disorder, the risk of causing harm is eliminated. Stratification might therefore help to move from the traditional stepped care prescription approach to a more streamlined and, most importantly, effective matched-care concept (Arns et al., 2023).

NEUROIMAGING METHODS FOR BIOMARKER IDENTIFICATION

One method to identify psychiatric biomarkers is through the use of brain imaging techniques, such as electroencephalography (EEG) or magnetic resonance imaging (MRI).

MRI uses a large magnetic field that leads to alignment of atomic nuclei with magnetic properties, such as hydrogen nuclei, which can be measured by a receiver coil (Grover et al., 2015). Due to differences in water content, different types of tissue can be made visi-

ble. Functional MRI (fMRI) uses so called blood-oxygen-level dependent (BOLD) signals, measuring changes in the magnetic property of hemoglobin caused by a change in oxygenation level of the blood (Glover, 2011). BOLD contrast indicates local changes in brain activity, for instance when performing a task. MRI can therefore provide structural images of gray and white matter tissue and brain lesions (structural MRI), as well as information about changes in functional brain activity within different areas (fMRI). Due to the large magnet that is needed to generate the magnetic field, a shielded room around the magnet and large quantities of liquid helium used for cooling in conventional scanners, MRI requires a considerable amount of space and is rather expensive.

EEG, on the other hand, can be measured using small, even portable, equipment, and is thus less cost-intensive and hence better suited for use in smaller practices. Electrodes, embedded in an EEG cap, noninvasively connect to the skin, usually by means of a conductive electrode gel. These electrodes measure postsynaptic potentials of large groups of pyramidal neurons that are synchronized and oriented towards the electrode in the underlying cortex (Beniczky & Schomer, 2020). The resulting signal is the difference in electric potential measured between two electrodes (the recording electrode and a reference electrode), and is depicted as waves on a screen. Due to the nature of the EEG, only cortical activity – i.e., close to the recording electrode- but no subcortical activity can be measured (Beniczky & Schomer, 2020). EEG activity is usually described in terms of its power spectrum at different frequencies. Although there is no consensus on the exact frequency boundaries, the most commonly described ones are: delta (<4 Hz), theta (4–7.5 Hz), alpha (7.5–12.5 Hz), beta (12.5–30 Hz) and gamma (30–40 Hz) (Newson & Thiagarajan, 2019). Frequency power is measured during wakeful rest - called resting-state EEG - or the performance of tasks, or by event-related potentials (ERPs), describing slight changes in the EEG signal following a specific event, such as the presentation of a stimulus (Schomer & Silva, 2011).

One issue in EEG biomarker research has been the variability in processing methods being used to clean the raw EEG signal. These differences make it difficult to compare studies and might result in unsuccessful replication attempts (Widge et al., 2019). Thus, optimization of EEG processing is essential to sound biomarker research. Advantages and disadvantages of different processing parameters and challenges of signal cleaning have been discussed elsewhere (Mumtaz et al., 2021; Yao et al., 2019) and are outside the scope of the present work. However, Chapter 3, to some extent, focuses on the standardization and optimization of signal processing.
Due to its ease of implementation and lower cost compared to other imaging techniques, EEG is a suitable method to identify biomarkers and apply them in clinical practice. Neuroimaging biomarkers, capable of identifying biological substrates and unbalanced neurocircuitry underlying disease, have been a major focus of research with multiple different resting-state, task-dependent and treatment-emergent patterns suggested as potential predictors. However, a meta-analysis by Widge et al., published in 2019, identified profound flaws in EEG biomarker studies, criticizing a lack of out-of-sample validation, insufficient direct replication, differing processing methods, and general publication bias with the conclusion that EEG biomarkers cannot reliably predict depression treatment outcome yet (Widge et al., 2019).

A QUEST FOR BIOMARKERS AND CLOSING THE TRANSLATIONAL GAP – AIMS AND OUTLINE

As the previous pages highlight, the urgency and great personal and societal burden of depression, the high occurrence of a difficult-to-treat patient group and the lower chance of remission with increasing treatment failures necessitate identifying the best treatment option for each person early on during the treatment course, and especially earlier than is presently standard. The current system of trial-and-error treatment prescription is outdated, and the cry for modernization through personalization is getting louder. For this reason, robust, tested and validated predictors of treatment outcome that can advance to the implementation stage without risk of harm

are needed. However, it is well-known that a translational gap exists between biomarker research and treatment personalization in psychiatry that prevents individualized mental health care from making headway (Carvalho et al., 2020; Fernandes et al., 2017). Despite good progress in the past years, full precision psychiatry at the personal level, transcending established disease boundaries, is not quite ready for clinical practice. Treatment stratification might be able to offer an intermediate solution based on patient subgroups instead of the individual, and stratifying each patient to one of multiple approved, effective interventions for a given disorder, thereby ensuring that everyone will receive their individual best treatment option out of the tested alternatives. The current work proposes that, thanks to these characteristics, stratification can bridge the translational gap between biomarker development and application. This can be realized within the framework of a stepped care approach, informing best treatment choice at the individual's disease stage by means of neurophysiology.

In order to reach this goal, it is crucial to first improve on previous criticism of biomarker research to identify promising robust predictors. In the second Chapter, the state-of-the-art of imaging biomarkers, 5 years after the critical meta-analysis of Widge et al. (Widge et al., 2019) will be introduced. It evaluates, based on sample size and replication attempts, which biomarkers are most robust and worthy of bringing to implementation.

One of these is the individual alpha frequency (iAF) which had already been investigated as predictor for both attention-deficit hyperactivity disorder (ADHD) and MDD albeit with inconsistent results (Arns et al., 2008; Arns, Drinkenburg, Fitzgerald, et al., 2012; Arns et al., 2018; Corlier et al., 2019; Krepel et al., 2018, 2020; Roelofs et al., 2020). One potential reason for these discrepancies in findings could be the vast differences in methodology, such as data cleaning, utilized across these studies, which has been another point of critique of previous biomarker studies. Hence, in Chapter 3, 108 EEG data processing permutations were evaluated in a large dataset by means of iAF maturation in childhood and early adolescence to standardize and optimize detection of alpha oscillations and identification of the

alpha peak. The resulting biomarker, Brainmarker-I, was subsequently evaluated in multiple approved ADHD treatments and validated through assigning predicted remitter status in a blinded fashion in independent data, to ensure that any potential predictions did not represent spurious findings.
The same strategy was adopted in Chapter 4 assessing stratification to approved treatments for MDD, such as sertraline, different rTMS protocols and electroconvulsive therapy (ECT), once more comprising multiple, rigorous blinded out-of-sample validations. This MDD prediction was augmented with another higher frequency TMS treatment protocol in Chapter 5.

All studies utilized clinical datasets for biomarker development to ensure that heterogeneity was accounted for, and focused on remission as primary study outcome, since residual symptoms after treatment represent a higher risk for relapse (Paykel, 2008).

Five years after Widge and colleagues published their critique of EEG biomarkers, we show that research on neuroimaging biomarkers can and has improved on former standards, for instance by attempting to replicate findings in independent samples. The following chapters are dedicated to demonstrating that biomarker research has succeeded in identifying some robust imaging predictors that are ready to be implemented in clinical practice.

2

EVALUATING ROBUSTNESS OF BRAIN STIMULATION BIOMARKERS FOR DEPRESSION: A SYSTEMATIC REVIEW OF MAGNETIC RESONANCE IMAGING AND ELECTROENCEPHALOGRAPHY STUDIES.

Based on Klooster, D.*, Voetterl, H.*, Baeken, C., & Arns, M. (2023). Evaluating Robustness of Brain Stimulation Biomarkers for Depression: A Systematic Review of Magnetic Resonance Imaging and Electroencephalography Studies. *Biological psychiatry*, S0006-3223(23)01569-X. Advance online publication. https://doi.org/10.1016/j.biopsych.2023.09.009

*contributed equally to this work as joint first authors.

ABSTRACT

Noninvasive brain stimulation (NIBS) treatments have gained considerable attention as potential therapeutic intervention for psychiatric disorders. The identification of reliable biomarkers for predicting clinical response to NIBS has been a major focus of research in recent years. Neuroimaging techniques, such as electroencephalography (EEG) and functional magnetic resonance imaging (MRI), have been used to identify potential biomarkers that could predict response to NIBS. However, identifying clinically actionable brain biomarkers requires robustness. In this systematic review, we aimed to summarize the current state of brain biomarker research for NIBS in depression, focusing only on well-powered studies ($N \geq 88$) and/or studies that aimed at independently replicating previous findings, either successfully or unsuccessfully. A total of 220 studies were initially identified, of which 18 MRI studies and 18 EEG studies met the inclusion criteria. All focused on repetitive transcranial magnetic stimulation treatment in depression. After reviewing the included studies, we found the following MRI and EEG biomarkers to be most robust: 1) functional MRI-based functional connectivity between the dorsolateral prefrontal cortex and subgenual anterior cingulate cortex, 2) functional MRI-based network connectivity, 3)

task-induced EEG frontal-midline theta, and 4) EEG individual alpha frequency. Future prospective studies should further investigate the clinical actionability of these specific EEG and MRI biomarkers to bring biomarkers closer to clinical reality.

EVALUATING ROBUSTNESS OF BRAIN STIMULATION BIOMARKERS FOR DEPRESSION: A SYSTEMATIC REVIEW OF MAGNETIC RESONANCE IMAGING AND ELECTRO-ENCEPHALOGRAPHY STUDIES.

The search for biomarkers of clinical response to non-invasive brain stimulation (NIBS) treatments has been a major focus of attention over the last decade. Since the introduction of the DSM-5 in 2013 an even stronger focus on biomarker research was ignited by the launch of the National Institute for Mental Health (NIMH) Research Domain Criteria (RDoC) project. A few years later, NIMH made RDoC inclusion mandatory for NIMH funded research, and the term 'personalized medicine' transitioned into the now more frequently used term 'precision psychiatry'. At the same time, some of the largest biomarker studies for major depressive disorder (MDD) emerged, such as the International Study to Predict Optimized Treatment in Depression (iSPOT-D) (Williams et al., 2011), EMBARC (Establishing Moderators and Biosignatures of Antidepressant Response for Clinical Care) (Trivedi et al., 2016), or CAN-BIND (Canadian Biomarker Integration Network in Depression)(Lam et al., 2016). In parallel, a wider adoption of NIBS techniques emerged, such as repetitive transcranial magnetic stimulation (rTMS) for the treatment of MDD and other conditions such as obsessive-compulsive disorder (OCD) or addiction, with currently more than 24 FDA device approvals (Cohen et al., 2022), as well as transcranial electrical stimulation (tES). Many NIBS studies have been complemented by imaging work (Blumberger et al., 2018; Corlier et al., 2019; van Dijk et al., 2022). Since many NIBS applications have built upon neuroscientific knowledge (e.g., frontal asymmetry) and given the focus on interventional psychiatry and brain circuit therapeutics (Siddiqi, Schaper, et al., 2021; Spellman & Liston, 2020), identifying NIBS biomarkers is of great importance,

both to improve clinical outcomes, and to validate hypothesized working mechanisms. We, therefore, aim to systematically review the current state of biomarker-driven precision psychiatry for NIBS. Several prior reviews and meta-analyses have investigated biomarkers for depression focused on EEG (Olbrich et al., 2015) or MRI (Wingen et al., 2020) and a critical meta-analysis questioned the usefulness of EEG biomarkers for guiding antidepressant response (Widge et al., 2019). This latter meta-analysis raised valid concerns about biomarker studies criticizing a lack of, particularly out-of-sample, replications, and demonstrating strong evidence for publication bias, with overrepresentation of studies with large effects and underrepresentation of null findings. This highlights the need for well-powered studies and out-of-sample validations to identify clinically actionable biomarkers. This systematic review, thus, focused on 1) adequately powered imaging studies and 2) studies that attempted to (out-of-sample) replicate earlier findings.

Concretely, biomarkers were considered robust when they were derived from an adequately powered study and/or shown to be replicable. The aim of this systematic review was to systematically extract robust biomarkers of NIBS treatment response.

INCLUSION CRITERIA

One of the main criticisms of Widge et al. (Widge et al., 2019) was that EEG biomarker studies suffered from low sample sizes (median N=25). Therefore, to prevent inclusion of underpowered studies and determine the right minimum sample size for inclusion, we first consulted power calculations from pivotal biomarker studies (see supplement). Given these pivotal trials yielded inconsistent sample-size justifications, we conducted a power calculation in GPower 3.1. (Faul et al., 2007) to determine a minimum sample size to define robust studies. We used a categorical outcome measure reflecting the difference in biomarker presence between responders and non-responders expressed as a medium effect size (Cohen's d=0.5) with an alpha level of $p<0.05$ and power of 0.7, resulting in a sample size of N=88. Furthermore, studies with smaller sample size could be included on the condition that subsequent replication studies were reported in an

independent sample. Studies investigating pre-treatment biomarkers of any NIBS modality and protocol were included. Studies identifying treatment-emergent biomarkers (biomarkers that reflect changes during treatment), were not taken into account in this review since such biomarkers would require high accuracy to justify stopping a treatment course halfway. Ideally, several studies found the same direction of effect in independent samples.

The exact search terms can be found in S2.1. Figure 2.1 visualizes the inclusion/exclusion and final selection of studies for EEG and MRI.

RESULTS

Results are summarized below as well as in Figure 2.1, with biomarkers grouped thematically.

Information on the included MRI and EEG studies are summarized in Table 2.1 and 2.2, respectively. The systematic review only yielded rTMS studies since no studies on other NIBS/TES modalities met our inclusion criteria. rTMS is a technique that can be used to non-invasively modulate brain activity, based on the principles of electromagnetic induction (Baeken et al., 2019). In the specific case of depression treatment, mostly the left DLPFC is stimulated (O'Reardon et al., 2007). When a different stimulation location was used or the biomarker was protocol-specific, this is explicitly stated. Additionally, a detailed description of technical terms used in this section can be found in Figure 2.4.

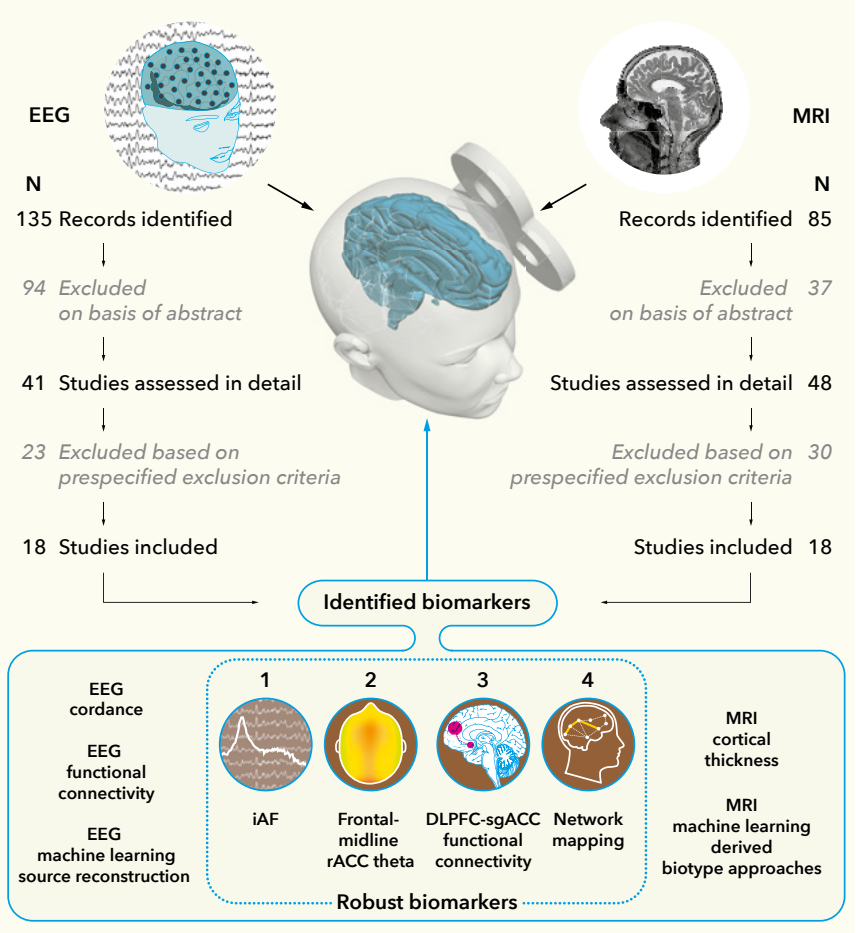


Figure 2.1. *Flow diagram of total studies identified, excluded, and included in the systematic review for electroencephalography (EEG) biomarkers (left) and magnetic resonance imaging (MRI) biomarkers (right), as well as all biomarkers identified and the most robust biomarkers that emerged from this systematic review (1, 2 for EEG and 3, 4 for MRI). Records were excluded on the basis of the abstract if they turned out to be nonhuman studies, were not original research, pertained to a pathology other than major depressive disorder or to a biomarker other than EEG/MRI or a treatment other than noninvasive brain stimulation. Prespecified exclusion criteria were 1) treatment-emergent biomarker and 2) sample size <88 and no replication. DLPFC, dorsolateral prefrontal cortex; iAF, individual alpha frequency; rACC, rostral anterior cingulate cortex; sgACC, subgenual anterior cingulate cortex.*

MRI BIOMARKERS

Anatomical MRI: Cortical thickness

In a first study, Boes et al. reported thinner rostral anterior cingulate (rACC) cortex at baseline to be associated with better clinical improvement (Boes et al., 2018). However, subsequent work failed to replicate this finding (Baeken et al., 2021) albeit here accelerated intermittent theta burst stimulation (iTBS) was used, whereas Boes and colleagues used 10 Hz rTMS.

fMRI: DLPFC-Subgenual ACC functional connectivity

In an influential 2012 study, Fox and colleagues suggested that the DLPFC (as part of the Central Executive Network) should only be seen as an entry point to a network relevant to the pathophysiology of depression (Fox et al., 2012). They demonstrated that clinical benefit of rTMS for depression was related to intrinsic functional connectivity (FC) of the respective DLPFC target to the sgACC (as part of the Default Mode Network), as determined by resting-state functional MRI (rs-fMRI). This functional connectivity was indexed as 'anti-correlation' of the sgACC to prefrontal cortical areas, and suggestive of a way to individualize prefrontal rTMS sites for MDD treatment by selecting the most sgACC-anti-correlated prefrontal site.

Several studies have attempted to replicate this finding, with successful conceptual replications by Weigand and colleagues (Weigand et al., 2018), Cash (Cash et al., 2019), Siddiqi (Siddiqi, Weigand, et al., 2021) and Elbau and colleagues (Elbau et al., 2023). However, studies in which a whole brain FC analysis was performed, using the sgACC as seed-region showed no relationship between functional anti-correlations between the seed and stimulation targets in the left DLPFC and response (Baeken et al., 2014, 2017; Ge et al., 2020; Hopman et al., 2021; Persson et al., 2020; Salomons et al., 2013). These non-replication studies are all based on individual rs-fMRI data. Hopman and colleagues even suggested an inverse relationship, i.e., stronger connectivity between the sgACC and stimulation site was related to improved clinical response (Hopman et al., 2021).

The studies by Fox et al. (Fox et al., 2012) and Weigand et al. (Weigand

et al., 2018) employed a normative functional connectome to derive FC. Cash et al. reasoned that using individual rs-fMRI data instead of a normative functional connectome may potentially improve TMS-personalization (Cash et al., 2019). Besides replication of previous results based on the normative connectome, this study showed that the relation between functional anti-correlation and clinical response was preserved when individual rs-fMRI data were used instead of group connectome data.

In 2021, Cash et al. introduced new insights into the relationship between FC and clinical responses (Cash et al., 2020). Instead of the direct FC between the stimulation site in the left DLPFC and the sgACC, the proximity between the clinically applied stimulation site and the rs-fMRI-personalized target in the left DLPFC was found to be related to clinical response. This relationship was not significant when personalized targets were replaced by a group average target derived from a normative functional connectome, arguing for the first time for the advantages of using individual rs-fMRI data. Siddiqi et al. (Siddiqi, Weigand, et al., 2021) confirmed the importance of distance and even reported a response rate of 100% for patients whose stimulated target was within 25 mm of the personalized target.

Recently, Elbau et al. published the largest study (N=295), focusing on the potential of sgACC connectivity to infer TMS coil positions, as of now (Elbau et al., 2023). Although an association between FC between the sgACC and left DLPFC target and clinical response was observed, this association was much weaker (r=-.16) compared to previous studies (e.g. r=-.355 (Fox et al., 2012)), with low explained variance (3%). Only when subject-specific TMS-induced electric field simulations were performed and a weighted seedmap method introduced by Cash et al. (Cash, Cocchi, et al., 2021) was used to derive the time series of the sgACC, the weak but robust correlation was found. Of note, this relation was stronger in a subgroup of patients with strong global signal fluctuations due to burst breathing patterns (Elbau et al., 2023). It was suggested that this weaker relationship could potentially be attributed to the relatively low-resolution of the rs-fMRI data (voxel size 5x5x5mm) (Siddiqi & Philip, 2023). Indeed, better data quality could lead to better predictions and nowadays more so-

phisticated scanning sequences such as multi-echo and multi-band sequences, are available (Lynch et al., 2020). Moreover, studies that showed stronger relations between anti-correlations and clinical responses based on high(er) resolution rs-fMRI data used strong smoothing parameters, effectively lowering the spatial resolution.
FC between the sgACC and the left DLPFC has been studied extensively in relation to clinical response to rTMS treatment in MDD. This information can be used to define personalized coil-positions and might in the future become a robust MRI-derived biomarker. However, optimal methodology to compute FC needs further investigation, and future prospective studies are warranted to demonstrate utility of this approach on the individual level.

fMRI: Lesion Network Mapping
In addition to using functional connections between specific brain regions as potential biomarkers, connectivity of stimulation sites with brain networks, in line with previous lesion network mapping (Boes et al., 2015) have also been related to clinical response. A general depression network was identified by studying FC profiles from the normative connectome of 14 independent datasets including data on brain lesions, TMS, or deep brain stimulation (DBS), representing different sources of causal effects (Siddiqi, Schaper, et al., 2021). Correlations between the individual connectivity maps of the TMS stimulation site and the depression network predicted the efficacy of the stimulation target. Cash et al. used a comparable approach to derive a network related to aberrant emotional processing in MDD patients, using coordinate network mapping of spatially heterogeneous coordinates (Cash et al., 2023). Of note, this emotional network resembles the depression network by Siddiqi et al. (ρ=0.47, p=0.00)(Siddiqi, Schaper, et al., 2021). Closer proximity between the stimulation target and the emotional-network-derived personalized targets was associated with better clinical response (Cash et al., 2023).
These findings suggest that in the future, effective rTMS stimulation sites could be derived from correlations between individual connectivity profiles and the depression network.

fMRI: ML-derived biotype approach

Using FC as input to machine learning (ML) approaches, Drysdale et al. (Drysdale et al., 2017) identified four clusters, called biotypes, which in a subsequent validation showed differential sensitivity to response to rTMS over the dorsomedial prefrontal cortex (dmPFC). Subtype 1, represented by reduced connectivity in a fronto-amygdalar network and reduced connectivity to anterior cingulate and orbitofrontal areas, showed a high partial response rate of 83% (25%, 61% and 30% for subtypes 2, 3 and 4, respectively). Of note, partial response was defined as a >25% reduction in Hamilton depression rating scale (HDRS), albeit results were similar when using the more traditional >50% cut-off for response but predicted full-response was lower (e.g. ~63% for biotype 1).

Later work by Dinga (Dinga et al., 2019) failed to replicate these findings in a more heterogeneous sample of 187 patients with depression and anxiety. Their analysis led to three instead of four clusters. Neither the canonical correlates nor the clusters were statistically significant. Potential methodological explanations for this non-replication are overfitting of the nonregularized canonical correlation analysis and arbitrary definitions of the subtypes (Dinga et al., 2020). Also, variations in the clinical sample characteristics might explain the non-replication (Grosenick & Liston, 2020).

Figure 2.2.*(page 36) Overview of study details on the included magnetic resonance imaging studies based on sample size ((N ≥ 88; highlighted in green) or on replication work (highlighted in blue). Strength of finding reports the area under the receiver operating characteristic curve (AUC), effect size, correlation coefficient, or another measure of effect size, depending on what was reported in the article. Total N refers to the full sample size used to compute the biomarker while group N is the sample size of the group in which the biomarker was tested for repetitive transcranial magnetic stimulation (rTMS). ACC, anterior cingulate cortex; aiTBS, accelerated intermittent theta burst stimulation; ANOVA, analysis of variance; BDI, Beck Depression Inventory; DBS, deep brain stimulation; DLPFC, dorsolateral prefrontal cortex; dmPFC, dorsomedial prefrontal cortex; HDRS, Hamilton Depression Rating Scale; MADRS, Montgomery–Åsberg Depression Rating Scale; QIDS, Quick Inventory of Depressive Symptomatology; rsFC, resting-state functional connectivity; sgACC, subgenual ACC.*

Biomarker Category	Study	Total N (Group N)	rTMS protocol	Outcome measure	Strength of finding	Positive Finding	Alternative Finding
Cortical thickness	Boes, 2018	48 (48)	10 Hz rTMS	BDI/HDRS-24	r N.S., p<.001	Thinner cortex in the rostral part of the ACC at baseline was associated with better clinical response	
	Baeken, 2021	50 (21)	Accelerated iTBS	HDRS-17	r = 0.51, p = .02		Thicker cortex in the right caudal part of the ACC at baseline was associated with better clinical response 3 days after stimulation
DLPFC-sgACC functional connectivity	Fox, 2012	149 (27)	10 Hz rTMS	MADRS	r = –0.355, p<.05	DLPFC sites with higher clinical efficacy showed higher anti-correlations with the sgACC	
	Baeken, 2014	20 (12)	Accelerated 10 Hz rTMS	HDRS-17	ANOVA F-value = 3.62		Responders showed significantly stronger rs FC anti-correlation between the sgACC and parts of the left superior medial prefrontal cortex
	Salomons, 2014	25 (25)	10 Hz rTMS (dmPFC)	HDRS-17	Peak z-score 3.6		Higher baseline connectivity between sgACC and dlPFC was associated with better clinical response.
	Weigand, 2017	25 (25) 16 (12)	10 Hz rTMS or 20 Hz rTMS 10 Hz rTMS	BDI MADRS	r = –0.55, p < .005 r = –0.52, p < .05	Better clinical response was related to higher anti-correlations between the stimulation site in the DLPFC and the sgACC	
	Baeken, 2017	50 (44)	Accelerated iTBS	HDRS-17	p<.01		FC between sgACC and medial orbitofrontal cortex at baseline could distinguish aiTBS responders from non-responders
	Cash, 2019	47 (24)	10 Hz rTMS	MADRS	r = –0.61, p=.001	Better clinical response was related to higher anti-correlations between the stimulation site in the DLPFC and the sgACC	
	Cash, 2020	26 (26)	10 Hz rTMS	MADRS	r = –0.54, p = .002	Better clinical response was related to higher anti-correlations between the stimulation site in the DLPFC and the sgACC	
	Ge, 2020	50 (50) 50 (32)*	10 Hz rTMS or iTBS	HDRS-17	AUC = 0.87, p<.001, r = –0.62, p<.001 AUC = 0.79, p<.001, r = –0.49, p=.001		Clinical response (post-treatment and 12 weeks after rTMS) was related to lower functional connectivity between the sgACC and the right DLPFC
	Persson, 2020	30 (20)	iTBS	MADRS	p = .021, T = 6.75		Baseline functional connectivity between the sgACC and the precuneus is negatively correlated with clinical response
	Hopman, 2021	70 (61 long-term, 63 short-term)	10 Hz rTMS	MADRS	Effect size: .26 – .30 depending on area cluster (for long-term responders vs non-responders)		Stronger DLPFC-sgACC connectivity was associated with symptom improvement. Long-term responders showed higher connectivity between sgACC and frontal pole, superior parietal lobule, and occipital cortex and between the left DLPFC and the central opercular cortex
	Siddiqi, 2021	25 (25)	10 Hz rTMS	BDI	r = –0.6, p < .005	Better clinical response was related to higher anti-correlations between the stimulation site in the DLPFC and the sgACC	
	Elbau, 2023	414 (295)	10 Hz rTMS or iTBS	QIDS-SR	r = –0.16, p=.006	Clinical response was related to higher anti-correlations between the stimulation site in the DLPFC and the sgACC	
Network Mapping	Siddiqi, 2021	713 (151)	10 Hz or 20 Hz rTMS	BDI/MADRS/ HDRS-24	Weighted mean r = 0.22, p < .001	Circuits derived from lesions, rTMS, and DBS stimulation sites are similar and connectivity to this circuit predicts efficacy of rTMS treatment	
	Cash, 2023	26 (26)	10 Hz rTMS	MADRS	r = –0.41, p = .018	Closer proximity between actual and emotional network-specific TMS targets is associated with better clinical outcome	
Machine Learning	Drysdale, 2017	1188 (154)	10 Hz or iTBS (dmPFC)	HDRS-17	χ^2 = 25.7, p<.001	Four distinct biotypes, characterised by dysfunctional connectivity in limbic and frontostriatal networks predicted clinical response to dmPFC rTMS	
	Dinga et al. 2019	187 (187)	Not applicable	IDS	ns		Biotypes could not be replicated

* *12-week Follow up*
Group N denotes the treatment group tested for effect
ns = not significant; N.S= not specified; IDS = inventory of depressive symptomatology

Sample size ≥88
Replication studies, sample size potentially <88
Robust biomarker found

Biomarker Category	Study	Total N (Group N)	rTMS protocol	Outcome measure	Strength of finding	Positive Finding	Alternative Finding
Theta power & iAF	Arns, 2012	90 (90)	10 Hz or 1 Hz	BDI	.814	Non-response characterised by increased fronto-central theta, slower iAF, larger P300 amplitude in response to high-pitched targets of auditory oddball task, decreased prefrontal delta and beta cordance	
	Krepel, 2018	106 (106)	10 Hz or 1 Hz	BDI	ns		Non-replication of Arns, 2012
iAF	Widge, 2013	180 (86)	10 Hz	HDRS-17	ns		All variables were non-significant (non-replication of iAF)
	Arns, 2014	90 (90)	10 Hz or 1 Hz	BDI	.697 (for alpha), .793 (for iAF)	Decrease in Lempel-Ziv complexity (LZC) from minute 1 to minute 2 in non-responders, increase in responders and controls; predictive accuracy improved when LZC was calculated on iAF range	
	Corlier, 2019	147 (68)	10 Hz, 5 Hz or sequential bilateral (10 Hz and 1 Hz)	IDS-30 SR (response ≥40%)	r = –0.305 (adj p = .045)	A higher iAF and lower iAF distance to 10 Hz were significantly correlated with symptom improvement to 10 Hz but not to 5 Hz or bilateral rTMS	
	Roelofs, 2021	153 (59)	10 Hz or 1 Hz	BDI	r = –0.250 (p = .028)	Significant negative correlation between distance of iAF to 10 Hz and BDI percent change for 10 Hz but not 1 Hz rTMS	
ML & theta cordance	Erguzel, 2014	147 (147)	25 Hz	HDRS-17	.904 (using genetic algorithm)	ML algorithm based on delta and theta cordance can classify responders and nonresponders with high accuracy	
	Erguzel, 2016	147 (147)	25 Hz	HDRS-17	.807 - .918	Erguzel, 2014 was replicated in same sample but with added external validation and assessing different classifiers	
ML & source reconstruction	Wu, 2020	177 (152)	10 Hz or 1 Hz	DASS	rTMS: p = .004; effect size N.S. Sertraline: AUC = 0.67 (taken from Nilsonne and Harrell, 2020)	Values of SELSER algorithm below median predict better outcome to 1 Hz rTMS in anxiety subscale of DASS	
	Meijs, 2022	193 (95)	10 Hz or 1 Hz	BDI	.719 (model with baseline BDI and age)	PRS-informed fICA EEG component, reflecting delta and theta power in left DLPFC, inversely correlated with delta power in right anterior PFC, distinguishes response/non-response	
EEG functional connectivity	Zhang, 2020	179 (179)	10 Hz or 1 Hz	BDI	ns		Identified subtypes based on beta functional connectivity could not distinguish response/nonresponse for rTMS
	Bailey, 2019	71 (42)	10 Hz initially; later unilateral 10 Hz or 1 Hz or sequential bilateral	HDRS-17	Balanced accuracy: 86.6%	Responders showed higher theta connectivity (averaged across EO and EC) than controls; ML model based on 54 alpha and theta power, connectivity, iAF and theta cordance features can classify responders/non-responders with high accuracy	
	Bailey, 2021	193 (193)	10 Hz or 1 Hz or sequential bilateral	BDI	ns (Cohen's d = 0.25241)		Non-replication of Bailey, 2019: no difference between responders/non-responders in all measured variables
Theta cordance	Bares, 2015	50 (25)	1 Hz	MADRS	.82	Baseline cordance and decrease in cordance after week 1 of treatment predictive of response	
	Hunter, 2018	18 (18)	10 Hz (with potential switch to bilateral after session 10)	IDS-SR, CGI-I	Baseline ns (p = .15)		Central cordance change at week-1 but not at baseline was significantly associated with treatment outcome
	Hasanzadeh, 2021	46 (46)	10 Hz initially; last 6 sessions unilateral 10 Hz or 1 Hz or sequential bilateral	BDI, HRSD	Accuracy: 91.3% beta, 76.1% cordance		Cordance features were not significantly different between responders/non-responders
Frontal-midline theta	Li, 2016	36 (24)	10 Hz	HDRS-17	.799	Responders showed significant increase of frontal theta after RECT	
	Li, 2021	105 (70)	10 Hz or iTBS	HDRS-17	.800 (for 10 Hz), .549 (for iTBS)	Replication of Li, 2016: post-RECT frontal theta predictive of 10 Hz rTMS response but not of response to iTBS	

Group N denotes the treatment group tested for effect
ns = not significant
N.S= not specified

Sample size ≥88
Replication studies, sample size potentially <88
Robust biomarker found

Figure 2.3. *(page 37) Overview of study details of included electroencephalography (EEG) studies based on sample size (N ≥ 88; highlighted in green) or on replication work (highlighted in blue). Strength of finding reports the area under the receiver operating characteristic curve (AUC), effect size, correlation coefficient, or another measure of effect size, depending on what was reported in the article. Total N refers to the full sample size used to compute the biomarker while group N is the sample size of the group in which the biomarker was tested for repetitive transcranial magnetic stimulation (rTMS).*

BDI*, Beck Depression Inventory;* CGI-I*, Clinical Global Impressions Scale (Improvement);* DASS*, Depression Anxiety Stress Scale;* DLPFC*, dorsolateral prefrontal cortex;* EO*, eyes open;* EC*, eyes closed;* fICA*, functional independent component analysis;* HDRS*, Hamilton Depression Rating Scale;* iAF*, individual alpha peak frequency;* IDS(SR)*, Inventory of Depressive Symptomatology (self-rated);* iTBS*, intermittent theta burst stimulation;* MADRS*, Montgomery–Åsberg Depression Rating Scale;* ML*, machine learning;* PRS*, polygenic risk score;* RECT*, rostral anterior cingulate cortex engaging cognitive task;* rTMS*, repetitive transcranial magnetic stimulation;* SELSER*, sparse EEG latent space regression.*

EEG BIOMARKERS

EEG frequency band power: Theta power

EEG biomarker studies have traditionally focused on frequency band power (e.g. theta or alpha), however few sufficiently powered biomarkers have been found and replicated for NIBS.

An early study by Arns et al. (Arns, Drinkenburg, Fitzgerald, et al., 2012) in 90 MDD patients reported that high frontocentral theta power, low prefrontal delta and beta cordance and high P300 amplitude at baseline were associated with non-response to 10 Hz rTMS over DLPFC. However, in a replication attempt the findings for theta and P300 could not be replicated by the same group (Krepel et al., 2018).

Frontal-midline theta power and change in frontal theta power, measured after a rostral ACC-engaging cognitive task demonstrated predictive potential in a small pilot study (Li et al., 2016). The findings were replicated in an independent sample and moreover it was shown that the obtained predictor was specific to 10 Hz rTMS since it could not predict response to iTBS treatment (Li et al., 2021). In both studies, response was evaluated after 10 treatment sessions - a low number to assess clinical improvement. The final sample size was small (N=24 in the pilot and N=35 per treatment arm in the replication), however, the concept of independent-sample replication

strengthens the findings, and the differential prediction for iTBS vs 10 Hz rTMS suggests potential for future treatment stratification.

EEG ML and source reconstruction

Wu et al. reported on ML applied to the alpha band, where response to sertraline – but not placebo – could be specifically predicted in the EMBARC dataset (Wu et al., 2020). When this alpha-signature of response to rTMS was prospectively tested, it predicted change on the anxiety subscale of the DASS (Depression, Anxiety and Stress Scale) after 1 Hz rTMS treatment. Notably, the predictive effect was specific to 1 Hz treatment (and not 10 Hz), and opposite that of sertraline, offering potential for stratification. However, since no effects for depressive symptoms were reported (neither BDI nor DASS-depression), this analysis cannot be considered a true out-of-sample validation. Moreover, when another group inferred the data points reported for the sertraline finding, and calculated the ROC curve, model performance was rather weak with an AUC= 0.67 (for a detailed critique about the methodology, see (Nilsonne & Harrell, 2020)).

A novel approach which conceptually resembles the previously mentioned rs-fMRI biotype analyses, applied independent component analysis to source-reconstructed EEG frequency band data. An EEG signature was identified that was associated with the polygenic risk scores for antidepressant response (Meijs et al., 2022). Subsequent application of this signature to new samples yielded an association with response to both antidepressants and rTMS in men, but not women. As selecting EEG biomarkers using genetic data is a novel technique, this study should rather be viewed as a proof-of-concept that could aid in future biomarker development but requires further replication and comparison of the obtained networks with other known rs-fMRI or EEG networks.

Individual alpha peak frequency

One of the most heritable and reproducible aspects of the EEG is the individual alpha peak frequency (iAF) - the exact frequency of the alpha oscillations (Van Beijsterveldt & van Baal, 2002; Posthuma et

al., 2001; Smit et al., 2005). Initial findings for iAF were mixed. Some studies reported an association between slow alpha and non-response to DLPFC rTMS (Arns et al., 2010; Arns, Drinkenburg, Fitzgerald, et al., 2012) which could not be replicated by the same group (Krepel et al., 2018) or by Widge et al. (Widge et al., 2013). Adding iAF to a predictive model of non-linear EEG features of the alpha band, on the other hand, improved model prediction albeit in a rather small group of non-responders (N=20)(Arns, Cerquera, et al., 2014). More recent work shed light on these contradictory results by showing a predictive effect of iAF that was specific to 10 Hz rTMS treatment outcome (with no such effect for 1 Hz R-DLPFC rTMS) and could only be found using an average reference (indexing more focal activity than the linked-ears montage used in the studies mentioned above) (Roelofs et al., 2020). Furthermore, the association between iAF and symptom improvement turned out to be a quadratic instead of the previously assumed linear effect, demonstrating that the distance of iAF to 10 Hz was negatively correlated with symptom improvement after 10 Hz rTMS (Corlier et al., 2019). These results were successfully replicated (Roelofs et al., 2020) in the same sample by Krepel et al. (Krepel et al., 2018), where previous findings (using linked ears reference) could initially not be replicated. This emphasizes the importance of exact methodological replications and a uniform way to preprocess and analyze EEG data.

EEG Cordance

A study investigating prefrontal theta cordance, originally developed by Leuchter et al (Leuchter et al., 1994) found that baseline cordance could predict response to 1 Hz rTMS with high accuracy (Bares et al., 2015) although this could not be replicated in another study where only 1-week change in theta cordance at central electrode sites predicted differences in response but not baseline or prefrontal cordance (Hunter et al., 2017).

Two ML studies investigated pretreatment frontal cordance to predict outcome to 25Hz rTMS in the same dataset of 147 subjects, using artificial neural networks (Erguzel et al., 2014; Erguzel & Tarhan, 2016). High classification accuracies were obtained, albeit in the first

study only a 6-fold cross validation was conducted but models were not tested in an external validation set which is considered necessary to prevent over-fitting (Ho et al., 2020). The second study in 2016 included a separate sample of 36 subjects for external validation, achieving high accuracy (AUC=.807-918). However, another ML study that used minimal-redundancy-maximal-relevance feature selection to test response prediction with frontal and prefrontal baseline cordance found no differences between responders and non-responders (Hasanzadeh et al., 2019). Thus, no conclusions can be drawn about the predictive value of baseline cordance.

EEG functional connectivity

Zhang et al. used ML to identify differences in beta connectivity in frontal and posterior regions during eyes-open recordings which could distinguish two clinical subtypes that responded differentially to psychotherapy in posttraumatic stress disorder and SSRI treatment in MDD (Zhang et al., 2020). However, no such differences between subtypes were found for rTMS, suggesting little relevance for rTMS prediction, but possible relevance for stratification between SSRI and rTMS treatment. Another ML model, built on 54 EEG features, such as baseline and week-1 alpha and theta connectivity (and other features such as power, iAF and cordance), demonstrated high predictive accuracy of response (86.6%) (Bailey et al., 2019), which could not be replicated in an independent sample (Bailey et al., 2021). The discovery analysis was based on only 12 responders compared to 128 responders in the replication sample. One important caveat of the replication analysis was the strong differences in EEG processing that can lead to different results (Roelofs et al., 2020).

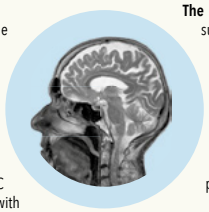
Findings regarding FC are, thus, inconclusive with different processing and modelling approaches hampering robust findings.

Figure 2.4. *(Following page) Glossary of terms used throughout the article. 3D, 3-dimensional; DLPFC, dorsolateral prefrontal cortex; EEG, electroencephalography; FC, functional connectivity; MRI, magnetic resonance imaging; rs-fMRI, resting-state functional MRI; sgACC, subgenual anterior cingulate cortex; TMS, transcranial magnetic stimulation.*

MRI

A normative functional connectome is an averaged connectivity map derived from rs-fMRI scans from many individuals, also called functional group connectome or human connectome. This connectome represents the average wiring diagram of the brain's functional connections. The advantage of a normative functional connectome is that the signal-to-noise ratio is higher compared to individual rs-fMRI data. However, inter-individual differences in functional connectivity are discarded.

TMS-induced electric field simulations can provide insight in the distribution of the TMS effects within the brain. When a TMS pulse is applied to the brain, a secondary electric field is induced in the superficial layers of the cortex. The exact distribution of this TMS-induced electric field depends on the shape of the TMS coil used as well as on the individual's gyral folding pattern.



The weighted seedmap method, introduced by Cash et al (1), is an alternative method to compute the time-series in the sgACC combining knowledge from the normative functional connectome with the individual rs-fMRI data. According to the weighted seedmap approach the time-series of the sgACC is computed as the weighted spatial average of the time-series in the gray matter voxels of the individual rs-fMRI data, excluding the DLPFC region. The weights are derived from the connectivity strength between the sgACC and the gray matter voxels in the normative functional connectome.

Global signal is the mean of the voxel time-series within the brain. Particularly in the work of Elbau et al. (2), the global signal is relevant since it was shown to reflect burst breathing patterns. Especially the subset of patients showing global signal patterns related to burst breathing showed strong negative correlations between sgACC-stim-site FC and clinical response.

Network mapping is an analysis technique that does not solely consider focal brain regions but is also sensitive to networks connected to those regions. At first, network mapping used lesions to seek convergence for symptoms caused by lesions in different non-overlapping brain regions (3). Network mapping has since been expanded to contain other (causal) sources of information such as TMS stimulation sites (TMS network mapping) (4) or coordinates related to abnormal brain functioning (coordinate network mapping) (5).

The emotional network, identified by Cash et al. (5), involves the subgenual cingulate cortex, pregenual anterior cingulate cortex, left DLPFC, cingulum, and superior frontal gyrus including the pre-supplementary motor area.

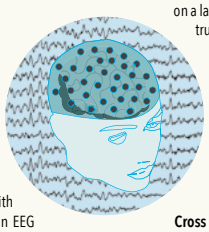
The depression network, derived by Siddiqi et al. (8), contains positive peaks in the DLPFC, frontal eye fields, inferior frontal gyrus, intraparietal sulcus and extrastriate visual cortex and negative peaks in the subgenual cingulate cortex and ventromedial prefrontal cortex. .

Canonical correlation analysis (CCA) is a well-established method used to identify the association between two sets of variables. Drysdale et al. (6) used CCA to select a low-dimensional representation of FC features that were related to weighted combinations of clinical symptoms. **Regularized CCA** is based on a subset of features. This prevents overfitting of CCA as might be the case in **nonregularized CCA.**

EEG

Frequency band power is most commonly calculated for the 5 standard frequency bands delta (1-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz) and gamma (30-100 Hz) although the frequency ranges are not standardized and often differ between studies. The power spectrum within a frequency band is usually calculated by Fast-Fourier Transform (FFT), an algorithm that transforms a signal from a time or space domain to a frequency domain.

The P300 is an event-related potential (ERP), which can be observed in the EEG in response to an infrequent tone in a row of frequent tones. It denotes a positive deflection approximately 300ms following the stimulus and is assumed to be involved in attention and memory processes.



Independent component analysis (ICA) is a computational method to filter a multivariate signal into its distinct subcomponents. ICA was here applied to data which had been source reconstructed with **LORETA (Low Resolution Brain Electromagnetic Tomography)**, an EEG method for 3D imaging brain activity to estimate where signals come from in the brain.

Polygenic risk scores (PRS) estimate a person's genetic predisposition to develop certain traits or disorders, based on their genetic profile and genome-wide association study data.

The individual alpha peak frequency (iAF) is the frequency at which an individual's alpha oscillations are most pronounced. It is calculated by determining the power spectrum within the alpha frequency band (see above) and identifying the highest (modal) peak in that spectrum.

Brainmarker-I is an iAF-based biomarker which has been age- and sex-normalized on a large dataset (>4000 individuals) by employing the biological ground truth that the iAF matures (speeds up) during childhood and adolescence (7).

Cordance is an EEG measure, originally developed by Leuchter and colleagues (8) that combines both absolute and relative power within a specific frequency band, with negative values reflecting increased slow-wave and decreased fast activity. This pattern was termed discordance and is assumed to reflect low perfusion and metabolism.

Cross validation is a statistical method used in machine learning to evaluate model performance. Ideally, an **external validation dataset** is used to test model predictions. Often, cross validation is done on a segment basis, meaning all data segments from all participants are merged and some segments are kept for later validation. This can lead to high prediction accuracy, so-called overfitting, since the model is predicting the participant instead of the signal.

1. Cash RFH, Cocchi L, Lv J, Wu Y, Fitzgerald PB, Zalesky A (2021): Personalized connectivity-guided DLPFC-TMS for depression: Advancing computational feasibility, precision and reproducibility. *Hum Brain Mapp* 42: 4155–4172.
2. Elbau IG, Lynch CJ, Downar J, Vila-Rodriguez F, Power JD, Solomonov N, *et al.* (2023): Functional Connectivity Mapping for rTMS Target Selection in Depression. *Am J Psychiat* 180: 230–240.
3. Boes AD, Prasad S, Liu H, Liu Q, Pascual-Leone A, Caviness VS, Fox MD (2015): Network localization of neurological symptoms from focal brain lesions. *Brain* 138: 3061–3075.
4. Siddiqi SH, Schaper FLWVJ, Horn A, Hsu J, Padmanabhan JL, Brodtmann A, *et al.* (2021): Brain stimulation and brain lesions converge on common causal circuits in neuropsychiatric disease. *Nat Hum Behav* 1–10.
5. Cash RFH, Müller VI, Fitzgerald PB, Eickhoff SB, Zalesky A (2023): Altered brain activity in unipolar depression unveiled using connectomics. *Nat Mental Health* 174–185.
6. Drysdale AT, Grosenick L, Downar J, Dunlop K, Mansouri F, Meng Y, *et al.* (2017): Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat Med* 23: 28–38.
7. Voetterl H, Wingen G van, Michelini G, Griffiths KR, Gordon E, DeBeus R, *et al.* (2022): Brainmarker-I differentially predicts remission to various attention-deficit/hyperactivity disorder treatments: a blinded discovery, transfer and validation study. *Biological Psychiatry Cognitive Neurosci Neuroimaging.* https://doi.org/10.1016/j.bpsc.2022.02.007
8. Leuchter AF, Cook IA, Lufkin RB, Dunkin J, Newton TF, Cummings JL, *et al.* (1994): Cordance: A New Method for Assessment of Cerebral Perfusion and Metabolism Using Quantitative Electroencephalography. *Neuroimage* 1: 208–219.

DISCUSSION

The aim of this systematic review was to assess the progress regarding EEG- and MRI-biomarkers for NIBS techniques. To improve upon previous criticisms, particularly the lack of replications as highlighted by Widge et al. (Widge et al., 2019), we focused on robustness in this review. To achieve this, we included only studies with a sample size of N ≥ 88 or those that attempted to replicate biomarkers in independent samples, in order to identify robust biomarkers that can be used clinically to predict response to NIBS techniques.

Eighteen MRI and 18 EEG biomarker studies were included (visualized in figure 2.1). All studies focused on rTMS while no relevant imaging biomarker studies were found for TES.

MRI BIOMARKERS

The most robust rs-fMRI based metric predicting clinical response supported by a large sample (N=295) (Elbau et al., 2023) as well as several independent replications, is the anti-correlation between the stimulation target (within the left DLPFC) and sgACC (Fox et al., 2012). This anti-correlation was shown to be related to response to various rTMS protocols, such as iTBS and 10 Hz rTMS. However, replication in the largest sample yielded only weak effects (Elbau et al., 2023), potentially suggesting reduced utility in clinical practice. Thus, prospective studies targeting the personalized location in the DLPFC with the highest anti-correlation with the sgACC should demonstrate if this connection has true biomarker potential.

A newer method based on network mapping also demonstrated biomarker potential of the connection between the stimulation site and a general depression network or an emotional network. Even though these findings are based on a large study using data from 151 rTMS stimulation sites (four merged rTMS datasets) (Siddiqi, Schaper, et al., 2021) and was independently replicated (Cash et al., 2023), more and prospective research is warranted to demonstrate clinical value.

EEG BIOMARKERS

For EEG biomarkers, frontal-midline-elicited theta power after an rACC-activating task and iAF emerged as the most promising and robust EEG biomarkers. Frontal-midline theta power has been extensively described in the literature as a biomarker for treatment prediction and is thought to reflect rACC theta (for review see Pizzagalli (Pizzagalli, 2011)), supported by the finding that an rACC-engaging task can elicit this frequency (Li et al., 2016, 2021). Interestingly, rACC activation was found to be predictive across imaging modalities, including EEG and fMRI (Pizzagalli, 2011). However, this was true for both sertraline and placebo response (Pizzagalli et al., 2018). Thus, despite successful replication, future studies should further investigate whether this finding is specific to 10 Hz rTMS vs. iTBS or should rather be considered a non-specific predictor of response, including placebo.

The iAF finding emerged from two well-powered studies (N=143; N=153) by two independent groups. Interestingly, this result was specific to 10 Hz rTMS (proximity of iAF to 10 Hz was associated with better clinical response, suggesting synchronization effects of rTMS to the endogenous iAF rhythm). Recent work has indicated promise for the iAF-based Brainmarker-I to stratify between 10 Hz left DLPFC and 1 Hz right DLPFC rTMS to enhance clinical outcomes (54), providing additional clinical merit for this biomarker.

LESSONS LEARNED: THE DEVIL IS IN THE DETAILS

There are many methods to derive seed regions and compute prefrontal-sgACC FC. Even though earlier work used circles or weighted cone models to derive seed region time-series, currently more advanced methods such as individual TMS-induced electric field simulations and weighted seedmap methods are used. These methodological details have shown to be highly influential since Elbau et al. (Elbau et al., 2023) only found a relation between the FC between stimulation site and sgACC and clinical response when the stimulated area was derived from the simulated electric field distribution and the sgACC time-series were derived using a seedmap approach.

Four of the papers included in this review demonstrate clinical value of using individual rs-fMRI data compared to group connectome data (Cash, Weigand, et al., 2021; Cash et al., 2023; Elbau et al., 2023; Siddiqi, Weigand, et al., 2021). Future research needs to compare biomarkers derived from these different connectomes and answer the question whether baseline individualized rs-fMRI data collection should be added to treatment protocols.

In the case of the iAF measured with EEG, initial findings were mixed, even though several well-powered studies were used to examine the effect (e.g. N=180 (Widge et al., 2013) or N=90-106 (Arns, Drinkenburg, Fitzgerald, et al., 2012; Krepel et al., 2018), and replication analyses were conducted. Later work actually led to consistent and robust findings (Corlier et al., 2019; Roelofs et al., 2020), showing that the crucial factors were: 1) Use of the correct EEG montage: initial studies used the less focal linked-ears reference, while Roelofs (Roelofs et al., 2020) demonstrated that the main result critically depended on the average reference montage; 2) protocol-specific effects for 10 Hz TMS and no such effect for 1 Hz TMS, meaning effects could average out when combined, and 3) a quadratic association between TMS response and iAF as opposed to the presumed linear association (i.e. lower iAF predicts worse TMS response).

The actual predictive value of clinical response of these MRI- and EEG-derived metrics, thus, depends on the preprocessing pipelines used. Future research is necessary to investigate if the content of these metrics is related to core brain mechanisms or reflect other sources of signal fluctuations such as respiration or cardiac patterns.

ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING

The present review reveals a limited biomarker potential for AI and ML-techniques in both EEG and rs-fMRI studies.

Although large – and often multiple - samples were used and results seemed promising at first glance, some studies lacked external validation samples (Erguzel et al., 2014) which are needed to prevent overfitting (Ho et al., 2020); some out-of-sample validation results were

only significant for different measures than the discovery measure (e.g. anxiety instead of depression) (Wu et al., 2020); and some could not be replicated, possibly due to overfitting (Bailey et al., 2021). Finally, it is important to use consistent definitions for response and remission (e.g. not 'partial response'), in order to keep outcomes comparable.

FUTURE DIRECTIONS

It remains to be investigated whether the biomarkers described in this review generalize to multiple rTMS protocols. If not, this might at least partly explain some of the unsuccessful replication attempts. Moreover, especially for biomarkers with weaker effects, the cost/benefit ratio needs to be assessed.
The present manuscript only discusses robust biomarkers for MDD. Future research is needed to determine if these are also predictive of treatment response in other disorders.

Finally, prospective studies, similar to van der Vinne et al (van der Vinne et al., 2021), will be necessary to test treatment individualization in daily clinical practice.

CONCLUSION

This systematic review has identified four robust neuroimaging biomarkers that have reached a sufficient level for testing in prospective trials to evaluate their feasibility and clinical actionability. Some of those biomarkers show promise for treatment stratification which might be a more realistic and feasible approach for clinical practice compared to precision psychiatry (Arns et al., 2022).

Overall, a limited number of studies met our inclusion criteria, highlighting the need for improvements in the quality of imaging biomarker research for rTMS. Nevertheless, the identification of four robust biomarkers over the past decade presents a promising outlook and justifies large trials, similar to iSPOT-D and EMBARC for antidepressant medication, but then aimed at rTMS and NIBS.

ACKNOWLEDGMENTS AND DISCLOSURES

This work was supported by the Queen Elisabeth Medical Foundation for Neurosciences (to CB), by an Applied Biomedical grant of the Agency for Innovation through Science and Technology (to CB), part of the Research Foundation - Flanders (FWO) PrevenD Project 2.0 (Grant No. T000720N [to CB]), and FWO project G01118N [to CB]). This work was also supported by a junior postdoc grant from FWO (to DK).
We thank Mark Koppenberg for designing the figures.

MA is unpaid chairman of the Brainclinics Foundation, holds equity/stock in neurocare, and served as consultant to neurocare and Roche. Brainclinics Foundation received research and consultancy support from neurocare and equipment support from Deymed and MagVenture. All other authors report no biomedical financial interests or potential conflicts of interest.

SUPPLEMENTARY MATERIAL

Supplementary material S2.1. Search Strategy and Article Selection

For EMBARC, no power calculation was performed and N=121 patients were treated with sertraline (1). For iSPOT-D, a rather strict power calculation was conducted powered on an odds ratio=1.3 with N=217, N=234 and N=204 treated with escitalopram, sertraline and venlafaxine, respectively (2). For CAN-BIND, a sample size of 98 per group (response/non-response) was estimated (3).

We conducted a PubMed search for articles published before 17th of March 2023 whose title/abstract contained keywords, in line with Widge, 2019, matching the following query:

(electroencephalogram OR electroencephalography OR EEG OR QEEG OR resting-state OR event-related potential OR ERP OR coherence OR spectral OR spectrum OR alpha OR beta OR theta OR delta OR gam-

ma OR N1 OR P2 OR P300 OR N200 OR SSVEP OR VEP OR AEP OR evoked potential OR oscillation OR electrical activity)

OR

(MRI OR magnetic resonance imaging OR connectivity OR functional connectivity OR FC OR resting-state OR connections OR structural connectivity OR diffusion OR perfusion OR network mapping OR graph measures)

AND

(depression OR depressive OR major depression OR major depressive disorder OR major depressive episode OR depressed OR antidepressant OR mood disorder)

AND

(differential OR predictor OR prediction OR predict OR predicted[Title/Abstract] OR predictive[Title/Abstract] OR biomarker OR marker OR phenotype OR response index OR subtype)

AND

(response OR respond[Title/Abstract] OR responds[Title/Abstract] OR responded[Title/Abstract] OR remission OR remit[Title/Abstract] OR remits[Title/Abstract] OR remitted[Title/Abstract] OR treatment response OR responsiveness OR nonresponse OR non-response OR responder OR non-responder OR remitter OR non-remitter OR therapeutic OR outcome OR treatment resistance OR comparative effectiveness OR effectiveness OR treatment selection OR efficacy)

AND

(rTMS OR transcranial magnetic stimulation OR iTBS OR theta burst stimulation OR non-invasive brain stimulation OR noninvasive brain stimulation OR NIBS OR tES OR transcranial electric stimulation OR tDCS OR transcranial direct current stimulation)

In addition, articles referenced in recent relevant reviews and relevant articles known to one of the authors were included.
Articles in a language other than English, non-human studies and unpublished data were excluded first.
All abstracts that appeared to involve EEG or MRI, treatment prediction, (non-psychotic) depressive illness and rTMS or tECS treatment were retained for further review.
We only included original research, and no articles published as book chapters.
We excluded articles which did not attempt to predict treatment response in depression or which used alternative methods for prediction or predicted treatment outcome for a treatment other than noninvasive brain stimulation.
Lastly, we only included adequately powered studies ($N \geq 88$) or studies that had been replicated in an independent sample. Unsuccessful replications were also included.
This protocol was not pre-registered.

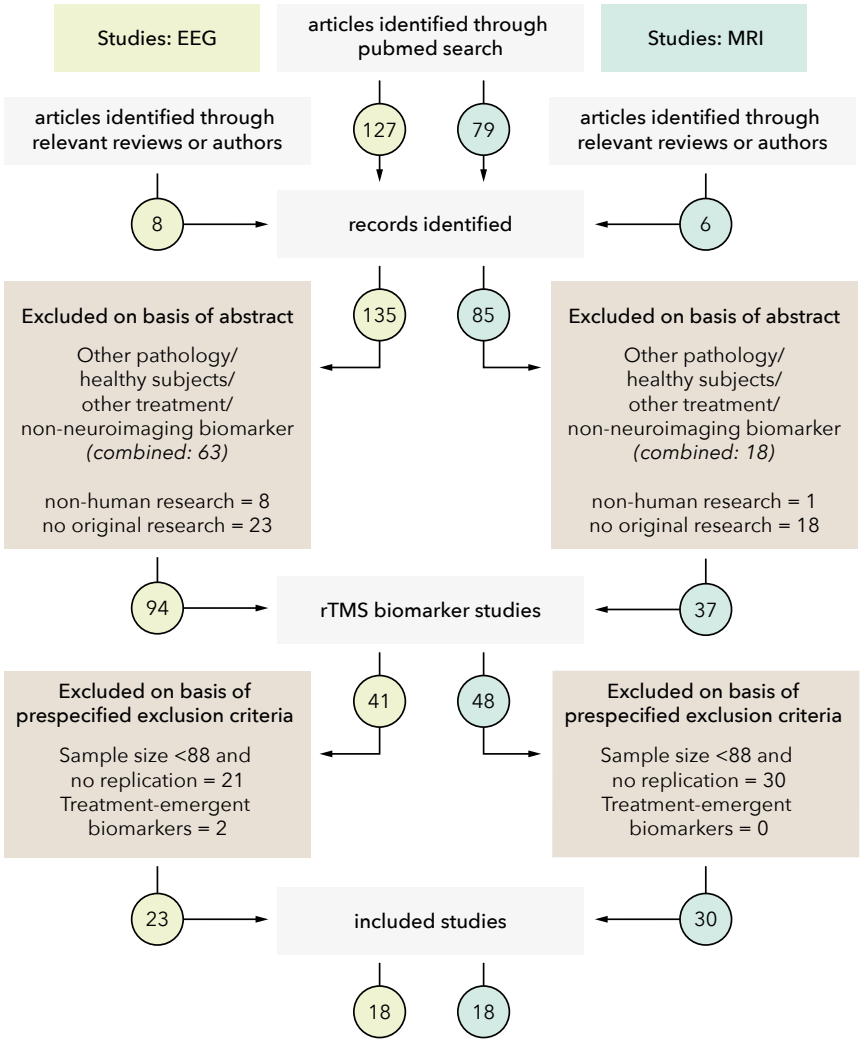


Figure S2.1. *Full flow-diagram study inclusion*
Flow-diagram of total studies identified, excluded and included in the systematic review for EEG-biomarkers (left) and MRI-biomarkers (right).
Records were excluded on basis of the abstract if they turned out to be non-human studies, no original research, pertain to another pathology than MDD, or another biomarker than EEG/MRI, or another treatment than NIBS.
Prespecified exclusion criteria were: 1. Treatment-emergent biomarker, and 2. Sample size <88 and no replication (EEG= electroencephalography; MRI=magnetic resonance imaging; rTMS= repetitive transcranial magnetic stimulation)

SUPPLEMENTARY REFERENCES

Pizzagalli DA, Webb CA, Dillon DG, Tenke CE, Kayser J, Goer F, et al. (2018): Pretreatment Rostral Anterior Cingulate Cortex Theta Activity in Relation to Symptom Improvement in Depression: A Randomized Clinical Trial. *Jama Psychiat* 75: 547.

Williams LM, Rush AJohn, Koslow SH, Wisniewski SR, Cooper NJ, Nemeroff CB, et al. (2011): International Study to Predict Optimized Treatment for Depression (iSPOT-D), a randomized clinical trial: Rationale and protocol. *Trials* 12: 1–17.

Lam RW, Milev R, Rotzinger S, Andreazza AC, Blier P, Brenner C, et al. (2016): Discovering biomarkers for antidepressant response: protocol from the Canadian biomarker integration network in depression (CAN-BIND) and clinical characteristics of the first patient cohort. *Bmc Psychiatry* 16: 105.

3

BRAINMARKER-I DIFFERENTIALLY PREDICTS REMISSION TO VARIOUS ATTENTION-DEFICIT/ HYPERACTIVITY DISORDER TREATMENTS: A DISCOVERY, TRANSFER, AND BLINDED VALIDATION STUDY.

Based on Voetterl, H., van Wingen, G., Michelini, G., Griffiths, K. R., Gordon, E., DeBeus, R., Korgaonkar, M. S., Loo, S. K., Palmer, D., Breteler, R., Denys, D., Arnold, L. E., du Jour, P., van Ruth, R., Jansen, J., van Dijk, H., & Arns, M. (2023). Brainmarker-I Differentially Predicts Remission to Various Attention-Deficit/Hyperactivity Disorder Treatments: A Discovery, Transfer, and Blinded Validation Study. *Biological psychiatry. Cognitive neuroscience and neuroimaging*, 8(1), 52–60. https://doi.org/10.1016/j.bpsc.2022.02.007

ABSTRACT

Background: Attention-deficit/hyperactivity disorder is characterized by neurobiological heterogeneity, possibly explaining why not all patients benefit from a given treatment. As a means to select the right treatment (stratification), biomarkers may aid in personalizing treatment prescription, thereby increasing remission rates. Introduction

Methods: The biomarker in this study was developed in a heterogeneous clinical sample (N=4249), and first applied to two large transfer datasets, a priori stratifying young males (<18 years) with a higher individual alpha peak frequency (iAPF) to methylphenidate (N=336) and those with a lower iAPF to multimodal Neurofeedback, complemented with sleep coaching (N=136). Blinded, out-of-sample validations were conducted in two independent samples. In addition, the association between iAPF and response to Guanfacine and Atomoxetine was explored.

Results: Retrospective stratification in the transfer datasets resulted in a predicted gain in normalized remission of 17-30%. Blinded out-of-sample validations for methylphenidate (N=41) and multimodal

Neurofeedback (N=71) corroborated these findings, yielding a predicted gain in stratified normalized remission of 36% and 29%, respectively.

Conclusion: The present study introduces a clinically interpretable and actionable biomarker based on the iAPF assessed during resting-state electroencephalography. Our findings suggest that acknowledging neurobiological heterogeneity can inform stratification of patients to their individual best treatment and enhance remission rates.

BRAINMARKER-I DIFFERENTIALLY PREDICTS REMISSION TO VARIOUS ATTENTION-DEFICIT/HYPERACTIVITY DISORDER TREATMENTS: A DISCOVERY, TRANSFER, AND BLINDED VALIDATION STUDY.

INTRODUCTION

Attention-deficit/hyperactivity-disorder (ADHD) is arguably the most common neurodevelopmental disorder and is characterized by highly heterogeneous impairment profiles and etiology (Banaschewski et al., 2017; Luo et al., 2019). Due to this heterogeneity and differential modes of treatment action (e.g., psychostimulant vs non-stimulant medication vs non-pharmacological treatments), even the most common interventions, although generally effective in the treatment of ADHD, only work in part of the ADHD population as shown by a large meta-analysis by Cortese et al (Cortese et al., 2018) on efficacy of various commonly prescribed ADHD medications (Cortese et al., 2018; Molina et al., 2009), with real-life remission rates of 31-57%, (reflecting effectiveness of treatments in the clinical setting rather than treatment efficacy as assessed in RCT's) (Pimenta et al., 2021). Therefore, individualized treatment recommendation based on biomarkers that predict clinical response to specific therapeutic interventions is desirable, one example being specific activity patterns measured by electroencephalography (EEG) (Atkinson et al., 2001).
Ideally, treatment should be individually adapted to a given patient

as envisioned in precision psychiatry. However, the multidimensionality of psychiatric disorders, in contrast to such clearly delineated problems as tumor-tissue, complicates tailoring treatment to a single person (Olbrich et al., 2016). An implementable intermediate step is treatment stratification, which aims to select a treatment from a range of effective treatments for a given disorder, informed by a biomarker (for review see Arns et al. (Arns et al., 2022)).

As an example, EEG biomarker studies for treatment prediction in major depressive disorder (MDD) have shown that specific EEG patterns or abnormalities are differentially associated with drug-specific or drug-class specific antidepressant treatment effects, as well as rTMS outcome (Arns, Bruder, et al., 2016; Arns et al., 2017; Olbrich & Arns, 2013; Roelofs et al., 2020; Wu et al., 2020). Many such studies yielded sex-specific EEG predictors of MDD treatment response (Arns, Bruder, et al., 2016; Van Dinteren et al., 2015; Iseger et al., 2017), as well as of methylphenidate (MPH) response in ADHD (Arns et al., 2018). Treatment stratification has already been implemented in the treatment of different cancer types (Deng et al., 2020; Kato & Manabe, 2018; Orr & McHugh, 2019) and recently also MDD, where stratification to different antidepressant medications was informed by pre-treatment EEG biomarkers, resulting in improved remission rates relative to treatment-as-usual (van der Vinne et al., 2021).

EEG is one of the most cost-effective and easily deployable methods to measure brain activity and is, thus, suitable for broad usage in clinical practice. Although several EEG patterns have been proposed for predicting treatment success in different mental disorders (Keizer, 2019; Olbrich et al., 2016), in ADHD most biomarker studies have focused on diagnostic biomarkers, while studies investigating prognostic ADHD biomarkers are still scarce (Clark et al., 2004; Pahor & Jaušovec, 2016).

The individual alpha peak frequency (iAPF) is the modal frequency at which an individual's alpha activity oscillates and is known to index brain maturation (Lindsley, 1936; Smith, 1938). This EEG pattern has been extensively investigated and shows promise in predicting outcome to various treatments across different disorders (Arns, 2012;

Olbrich & Arns, 2013). A higher mean frequency or a faster alpha peak is often associated with better cognitive performance, possibly reflective of faster information processing in thalamocortical pathways (Clark et al., 2004; Grandy et al., 2013; Klimesch, 1999; Pahor & Jaušovec, 2016) Conversely, many mental disorders such as Alzheimer's disease, mild cognitive impairment (Rodriguez et al., 1999), psychosis/schizophrenia (Murphy & Öngür, 2019; Yeum & Kang, 2018) and ADHD (Bazanova et al., 2018) are characterized by a slowed iAPF, potentially reflective of reduced or slowed information flow between the thalamus and the cortex (Clark et al., 2004). Furthermore, slow iAPF has been associated with worse clinical outcome to different treatments such as psychostimulants in ADHD (Arns et al., 2008, 2018) and antidepressant medication in MDD (Ulrich et al., 1984), whereas it was found to be related to better clinical outcome to multimodal EEG neurofeedback (NFB) treatment in ADHD (Krepel et al., 2020) and sertraline in MDD (Arns et al., 2017).

The current study therefore investigated whether iAPF is able to differentially predict clinical outcome to two disparate ADHD treatments, MPH and a multimodal behavioral intervention including NFB, sleep hygiene and coaching (MM-NFB).

Given the opposite implications reported for these treatments, we hypothesized that iAPF can help subdivide a heterogeneous population into more homogeneous subpopulations with relevance to clinical outcome and thus serve as a biomarker informing treatment stratification between medications (e.g. MPH) and MM-NFB. While there has been controversy regarding the specificity of EEG-NFB in the treatment of ADHD (Cortese et al., 2016; The Neurofeedback Collaborative Group et al., 2021), this manuscript focuses on EEG-NFB as part of a broader multimodal treatment including sleep hygiene and coaching, for which remission rates of 32-57% have been reported (Pimenta et al., 2021) and lasting clinical benefit has been demonstrated (Doren et al., 2018; The Neurofeedback Collaborative Group et al., 2021), although this is likely not solely attributable to the EEG-NFB alone. Given the stratification approach investigated here, being able to prescribe MM-NFB to people for whom psychostimulants are unlikely to work would nonetheless be advantageous.

Across the EEG literature, EEG (pre-) processing, EEG montages and frequency-band definitions vary considerably, which diminishes comparability and reproducibility that might at worst result in different findings (Yao et al., 2019) (see supplement S3.1 for more details). We therefore first developed Brainmarker-1 in a Biomarker Discovery Phase, where the most precise iAPF algorithm, i.e. the algorithm yielding the most biologically plausible iAPF, was determined. This algorithm was validated against a ground truth scenario, in this case relying on the well-established finding that iAPF indexes brain-maturation (Lindsley, 1936; Smith, 1938). The resulting biomarker was subsequently used to predict treatment outcome in the previously mentioned MPH and MM-NFB datasets based on previous findings (Arns et al., 2018; Krepel et al., 2020). These predictions were then corroborated in blinded, out-of-sample validations in a MPH and a MM-NFB dataset, which – to our knowledge – had not been attempted in EEG-biomarker studies before. We, furthermore, tested the biomarker's capacity to predict remission to two other pharmacological treatments, Guanfacine (GUAN) and Atomoxetine (ATX).

In order to maximize clinical utility of this stratification biomarker, we focused on remission as primary outcome, it representing the most clinically relevant measure (Steele et al., 2006; Swanson et al., 2001).

METHODS AND MATERIALS

DATASETS – BIOMARKER DISCOVERY PHASE

Since the goal was to explain variance in clinical data, the large TD-BRAIN+ dataset (see Table 3.1. for overview), comprising patients with various psychiatric disorders was utilized to determine the optimal parameters of iAPF calculation. The resulting optimized iAPF EEG processing pipeline was used to develop an age-standardized biomarker for males and females separately in accordance with previous reports of sex differences (Arns, Bruder, et al., 2016; Hermens et al., 2005), which was subsequently divided into deciles, for enhanced interpretability.
The open access TD-BRAIN dataset (N=1274), a subset of the data

used for the discovery phase, is freely available at www.brainclinics.com/resources (van Dijk et al., 2022), with all data recorded at Research Institute Brainclinics (Brainclinics Foundation, Nijmegen). In the TD-BRAIN+ dataset this was complemented with data from additional clinics (EPI-PIT clinics (Eindhoven & Tilburg; author JJ), EEG resource (Nijmegen; author RB), Neuroscan (Dordrecht; author PdJ), neuroCare clinics (Hengelo; Groningen; Munich; Sydney, author RvR)), for which the lab setup including EEG caps, amplifiers, instructions and other details were identical to the iSPOT-A trial (Arns et al., 2018).

DATASETS - BIOMARKER TRANSFER PHASE

The biomarker determined in the discovery phase was utilized to find the best way to stratify patients to Methylphenidate (iSPOT-A: N=257 (Arns et al., 2018)) and MM-NFB (N=50) (Krepel et al., 2020) according to the previously demonstrated directionality of effects (Arns et al., 2018; Krepel et al., 2020). Neurofeedback protocols comprised standard protocols such as Sensory-Motor-Rhythm (SMR), Theta-Beta-Ratio (TBR) and Slow Cortical Potential (SCP) neurofeedback.
This step focused on boys only, due to limited sample size of girls and no robust a priori knowledge regarding directionality of effect (e.g. Arns et al (Arns et al., 2018) only found effects for boys).

DATASETS - BIOMARKER VALIDATION PHASE

For independent out-of-sample replication analysis, we conducted a blinded prediction of remission in the MPH/GUAN dataset (Loo et al., 2016) (Table 3.1) and the International Collaborative ADHD Neurofeedback (ICAN) study (The Neurofeedback Collaborative Group et al., 2021), with accuracy verified by a third person not involved in the EEG analysis.
In the former trial, subjects were blindly randomized to either MPH (N=58) or GUAN (N=55) treatment. In the ICAN study (N=96), subjects were blindly randomized to a multimodal treatment of sleep and nutrition counselling and either theta/beta ratio NFB (MM-NFB)

or a control treatment ("NFB" administered based on a pre-recorded EEG to facilitate blinding of all).

Full Datasets	TD-BRAIN+	iSPOT-A	NFB	MPH/GUAN *Dataset*	ICAN	ACTION
Sample size (N)	4249	336	136	141	142	56
Age range, years	6-88	6-18	6-68	7-15	7-11	6-16
Included in analysis						
Sample size (N)	4126	184	41	76	71	39
Males (%)	2528 (60)	184 (100)	41 (100)	76 (100)	71 (100)	39 (100)
Mean Age (SD), years	29.3 (18.3)	11.8 (3.1)	11.1 (3.1)	10 (2.0)	8.6 (1.2)	11.5 (2.5)
Treatment	NA	MPH	NFB multimodal treatment*	MPH / GUAN	NFB / multimodal treatment	ATX

Table 3.1. Baseline demographics
Full datasets sample size reflects N of people who were enrolled. Sample size included in analysis reflects N of people with complete baseline data who finished treatment (except for TD-BRAIN+, where only baseline data but no clinical data was used). In the TD-BRAIN+ dataset the full age range was used for age-standardization while an age range of 6-18 years was used for the correlation analyses. Sample size of this age range was 1715 (1253 male); mean age was 11.8 (SD: 3.1). MPH = Methylphenidate, MM-NFB = Multimodal Neurofeedback, GUAN = Guanfacine, ATX = Atomoxetine, SD = Standard Deviation, NA = not applicable, since no treatment effects were assessed in the discovery dataset
** NFB treatment augmented with advice on sleep hygiene & coaching*

DATASETS - BIOMARKER EXPLORATION PHASE

In the exploratory phase to test performance of the biomarker to another commonly prescribed form of pharmacotherapy for ADHD (i.e., noradrenergic medications), the predictive value of the biomarker to ATX (n=47) and GUAN (n=55) was examined in the ACTION dataset (Griffiths et al., 2018) and in the MPH/GUAN dataset that had already been used in the validation phase for MPH replication (Loo et al., 2016).

All participants (or their parents or care-takers) gave written informed consent prior to testing.

EEG DATA COLLECTION AND PREPROCESSING

All EEGs were recorded in a standardized manner as developed by Brain Resource Ltd. (for more details see (Arns, Bruder, et al., 2016)) apart from the independent MPH/GUAN validation dataset (Loo et al., 2016).

In short, EEGs were recorded from 26 channels according to the 10-20 electrode international system (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,O2; Quikcap, NuAmps). Measurements consisted of 2-minute Eyes-Open (EO) and 2-minute Eyes-Closed (EC) recordings. During EO recordings, participants were asked to fixate a dot in the middle of the computer screen.

Data was recorded with the ground at AFz, and a sampling rate of 500 Hz and a low-pass filter with an attenuation of 40 dB per decade above 100 Hz was employed prior to digitization. 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 <10 kΩ for all electrodes.

Automatic artifact detection and removal were performed using a custom-built Python package (Harris et al., 2020; Hunter, 2007; The Pandas development team, 2020; Virtanen et al., 2020) and were in accordance with deartifacting as described in (Arns, Bruder, et al., 2016) and van Dijk et al. (van Dijk et al., 2022), with full code available online (www.brainclinics.com/resources).

For the MPH/GUAN validation dataset (Loo et al., 2016), eyes-closed EEGs were recorded from 40 channels (AF3,AF4,AFz,C3,C4,CPz,Cz,F10,F3,F4,F7,F8,F9,FCz,FP1,FP2,FPz,FT10,FT7,FT8,FT9,Fz,Iz,O1,O2,Oz,P10,P3,P4,P7,P8,P9,POz,Pz,T7,T8,TP10,TP7,TP8,TP9) for 5 minutes with a sampling rate of 256 Hz and referenced to linked ears (for further details, see (Loo et al., 2016; McCracken et al., 2016)). Recordings were subsequently matched to the other data, i.e., the 40 channels were reduced to 22 channels matching TD-BRAIN+ set-up (with FC3,

FC4, CP3 and CP4 missing). Artifact rejection for the independent validation dataset was performed in BrainVision Analyzer Version 2.2.0 (Brain Products GmbH, Gilching, Germany) by semi-automatic removal of epochs with signal amplitudes >150mV.

IAPF DETERMINATION

The individual alpha peak frequency was determined by computing the FFT of the preprocessed, artefact-free data. Subsequently each individual's iAPF was determined by identifying the highest peak within the frequency range of 7 to 13 Hz.

BIOMARKER DISCOVERY PHASE

Biomarker discovery a priori focused on males and females separately due to previously reported qualitative sex differences (Arns, Bruder, et al., 2016; Hermens et al., 2005).
In short, data with LVA were identified and excluded from further analysis (for more details, see supplement S3.2). In order to optimize EEG processing, iAPFs determined with different processing parameters (e.g. segment length, reference montage) were correlated with age (<18 years). The parameter combination with the highest correlation and retention of subjects was used for further prospective testing. Subsequently, the data was age- and sex-standardized and resulting values divided into 10 equal-sized bins (deciles) to improve interpretability.
For more information, see supplement S3.3.

BIOMARKER TRANSFER PHASE

We first aimed to align previous findings which differed with regard to primary outcome measure (response vs remission) and subsample (boys aged 6-18 vs boys aged 12-18) (Arns et al., 2018; Krepel et al., 2020). To increase comparability and clinical impact, we focused our analyses on males in the age range of 6-18 years and on remission – defined as an item mean of ≤1.00 on the ADHD-RS-IV - as primary clinical outcome (Steele et al., 2006).

BIOMARKER VALIDATION PHASE

Finally, the biomarker was prospectively validated on the same subsample (boys aged 6-18 years) for MPH and MM-NFB treatment by a blinded prediction of remission status, solely based on age, sex and baseline EEG in two independent datasets.

BIOMARKER EXPLORATION PHASE

Analyses for the exploration phase were similar to those in the transfer phase but without a guided hypothesis.

STATISTICS

First, Spearman correlations between the various iAPFs resulting from different EEG processing combinations and age in subjects below 18 years (N=1671) were calculated. To determine standardized iAPF values independent of age, we derived non-linear regression models based on the full TD-BRAIN+ dataset that most closely fit the given data for each electrode (Fz, Pz, Oz). Different mathematical models following the developmental trajectory of the iAPF (such as a Log gaussian model, in line with van Dinteren et al., 2014) were contrasted against a linear model (null hypothesis) and individually adjusted for females and males and for each site (channel). Divergence values representing where the individual's iAPF lies in relation to other people's iAPFs, were calculated from the resulting models by subtracting the model-derived average iAPF for each subject's age from the person's actual iAPF. Correlations between divergence values and age were conducted to confirm that the age effect had been eliminated from the data. The resulting divergence values were ranked from low to high and divided into 10 equal-sized bins (deciles) to improve interpretability by clinicians.

The final stratification outcome for the transfer phase and stratification decision for the exploration phase were based on the positive predictive values (PPVs) at different decile cut-off points, indicating remission rate within the subsample of patients that the biomarker would have stratified to the respective treatment. Since PPVs are

dependent on prevalence (here: observed remission) and remission rates differed between treatment datasets, we normalized PPVs for better comparability across datasets by dividing the PPV by the observed remission and subtracting 1.
Curve fitting models were developed in GraphPad Prism version 8.4.0 for MacOS. Spearman correlations were conducted with Python modules scipy, and numpy.
All other statistical analyses were performed in IBM SPSS Statistics for Macintosh, Version 27.0.

RESULTS

DATASETS

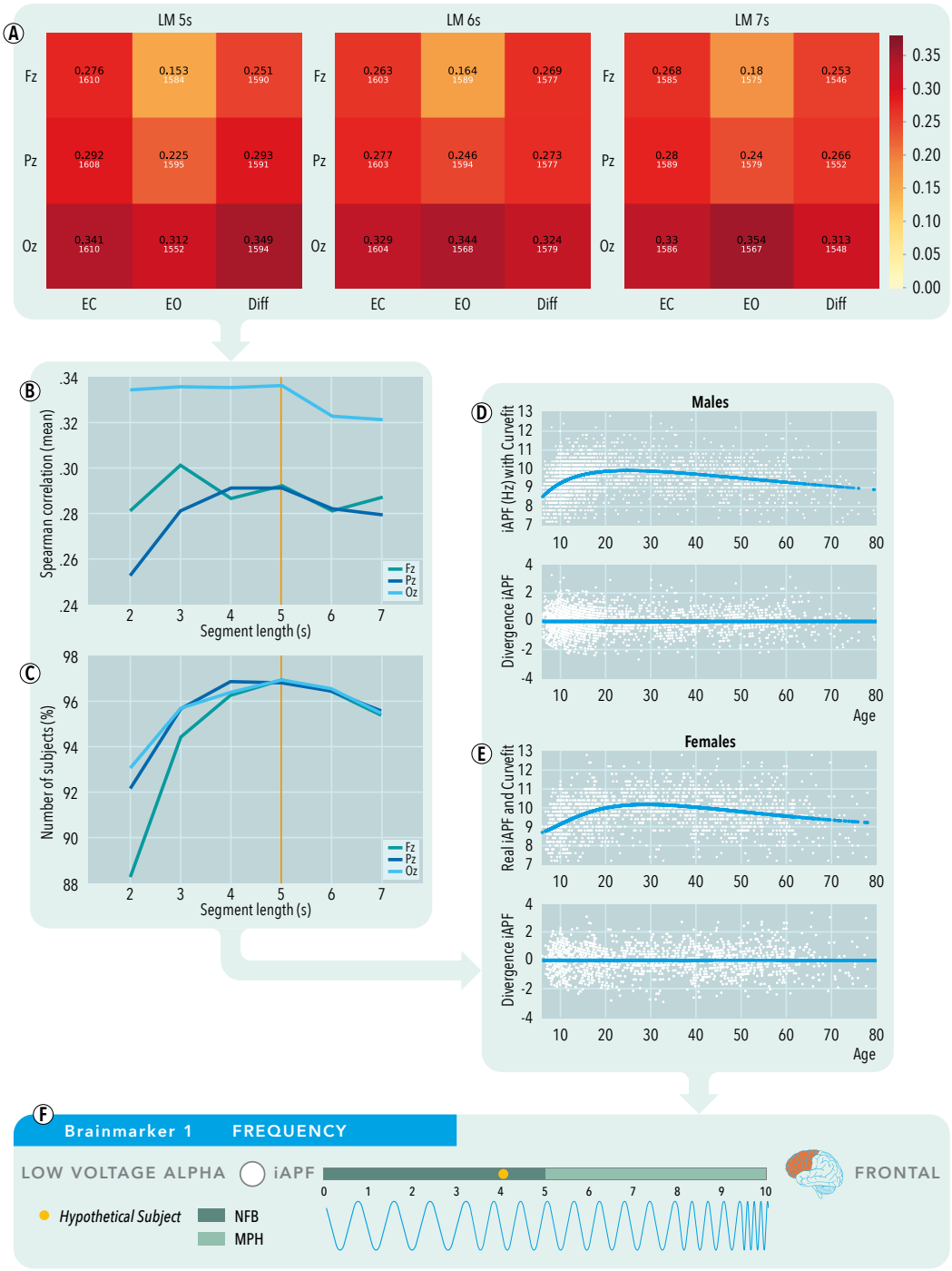
Table 3.1 provides a summary of the basic demographic information of all datasets.

BIOMARKER DISCOVERY PHASE

Figure 3.1 visualizes the individual steps of the biomarker development.

In short, a total of 108 algorithm permutations were tested (Figure 3.1.A). The resulting best permutation (linked-mastoid reference/eyes closed/5s segments) was selected for further prospective testing of the biomarker (Figure 3.1.B and C). A linear regression of the resulting age-standardized divergence values and age yielded a model with a slope of 0 (β =.000), demonstrating that the curve fitting procedure successfully removed the age effect seen before (e.g. Fz: R^2=.000).

For an overview of all correlation and secondary analyses, see supplementary material S3.3.


A
LM 5s
LM 6s
LM 7s
Fz
Pz
Oz
EC
EO
Diff
0.276 1610
0.153 1584
0.251 1590
0.292 1608
0.225 1595
0.293 1591
0.341 1610
0.312 1552
0.349 1594
0.263 1603
0.164 1589
0.269 1577
0.277 1603
0.246 1594
0.273 1577
0.329 1604
0.344 1568
0.324 1579
0.268 1585
0.18 1575
0.253 1546
0.28 1589
0.24 1579
0.266 1552
0.33 1586
0.354 1567
0.313 1548
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00
B
Spearman correlation (mean)
.34
.32
.30
.28
.26
.24
2 3 4 5 6 7
Segment length (s)
Fz
Pz
Oz
C
Number of subjects (%)
98
96
94
92
90
88
2 3 4 5 6 7
Segment length (s)
Fz
Pz
Oz
D
Males
iAPF (Hz) with Curvefit
13 12 11 10 9 8 7
10 20 30 40 50 60 70 80
Divergence iAPF
4 2 0 -2 -4
10 20 30 40 50 60 70 80
Age
E
Females
Real iAPF and Curvefit
13 12 11 10 9 8 7
10 20 30 40 50 60 70 80
Divergence iAPF
4 2 0 -2 -4
10 20 30 40 50 60 70 80
Age
F
Brainmarker 1 FREQUENCY
LOW VOLTAGE ALPHA
iAPF
FRONTAL
0 1 2 3 4 5 6 7 8 9 10
Hypothetical Subject
NFB
MPH

BIOMARKER TRANSFER PHASE: STRATIFICATION WITH BIOMARKER RESULTS IN HIGHER LIKELIHOOD OF REMISSION

To account for possible confounding effects of symptom severity, we first conducted a partial correlation between baseline ADHD-RS scores and iAPF, controlling for age, which was not significant (r=-.064, n=253, p=.311).

Figure 3.2 summarizes the outcome of the transfer phase. The direction of stratification was informed by the previously reported directionality of effects (higher iAPF indicating stratification to MPH (Arns et al., 2018), lower iAPF indicating stratification to MM-NFB (Krepel et al., 2020)) and was based on the Fz electrode as primary site based on prior literature (Arns et al., 2018) (see supplement S3.4 for a post-hoc analysis examining stratification based on Fz and Oz). A decile cut-off point of 1-5 for MM-NFB and 6-10 for MPH was chosen a priori, stratifying approximately half of the patients to each treatment. To test this a priori decision, positive predictive values (PPVs), indicating remission rates in the patient subsample that would have been stratified according to our biomarker were determined for different decile cut-off points. The chosen cut-off point of decile 5 indeed led to the highest combined PPV (supplementary table S3.1). Therefore, the presented biomarker (Brainmarker-1) was based on this cut-off point, recommending MM-NFB treatment to boys with a relatively lower iAPF in the decile range 1-5 and MPH to boys with a relatively higher iAPF in deciles 6-10 (see supplementary table S3.2 for additional accuracy measures).

Fig. 3.1. (page 66) Biomarker Discovery Phase. (A) Excerpt from heatmaps of the total of 108 algorithm permutations (27 depicted) that were tested and selected based on the highest correlation between age and iAPF in subjects <18 years (Spearman Correlation ρ; black digits) and the highest retention of data (number of subjects N; white digits). (B, C) Spearman Correlation ρ between age (6-18 yrs.) and iAPF (B) and number of subjects (C) for each electrode and segment length (2-7 seconds) for condition eyes-closed (EC) averaged across reference montages (N=1715). (D, E) Flattening the iAPF-age curve for males (D) and females (E) separately at electrode location Oz. Upper subplots depict non-standardized iAPFs and the optimized Log Gaussian model fit. Lower subplots depict the age-standardized divergence values and a linear fit through the data. (F) Example of the derived biomarker (Brainmarker-1) based on the final age- and sex-standardized scores, with deciles 1-5 yielding a recommendation for neurofeedback (NFB) treatment and deciles 6-10 yielding a recommendation for methylpheniddate (MPH). EO, eyes-open; LM, linked-mastoid

The normalized PPV indicated a predicted increase in remission rate of 17% compared to the observed remission rate if patients had received MPH (PPV=41%) and of 30% if patients had received MM-NFB (PPV=62%) as treatment recommendation based on Brainmarker-1. In a post-hoc analysis predicting remission with Brainmarker-1 calculated at the occipital site (Oz), no improvement could be seen for MPH (normalized PPV=+1.7%), however, for MM-NFB the PPV increased to 71.4% (normalized PPV=+51% as compared to +30% in Fz). Despite this improvement for MM-NFB treatment, Fz remained the primary stratification site, as prediction for MPH was only possible with the iAPF recorded at this location. For the results of stratification based on both Fz and Oz locations, we direct the reader to supplement S3.4.

OUT-OF-SAMPLE VALIDATION PHASE: STRATIFICATION BIOMARKER PREDICTS REMISSION IN PROSPECTIVE VALIDATION ANALYSIS

Next, the biomarker was validated by predicting remission to MPH and multimodal treatment including MM-NFB (ICAN) in two independent datasets (The Neurofeedback Collaborative Group et al., 2021; Loo et al., 2016), blinded to clinical outcome, and based solely on the subjects' age, sex and baseline iAPF. Accuracy was verified by a third person not involved in the EEG analysis (for MPH: authors GM and SKL; and for MM-NFB: author MA). Results are visualized in Figure 3.2.

In line with the previous analyses, we normalized PPVs to improve comparability with the transfer datasets. The normalized PPV predicted an increase in remission rate of 36% (PPV=50%) compared to the observed remission rate if patients had received MPH and of 29% (PPV=29%) if patients had received the multimodal treatment based on Brainmarker-1.

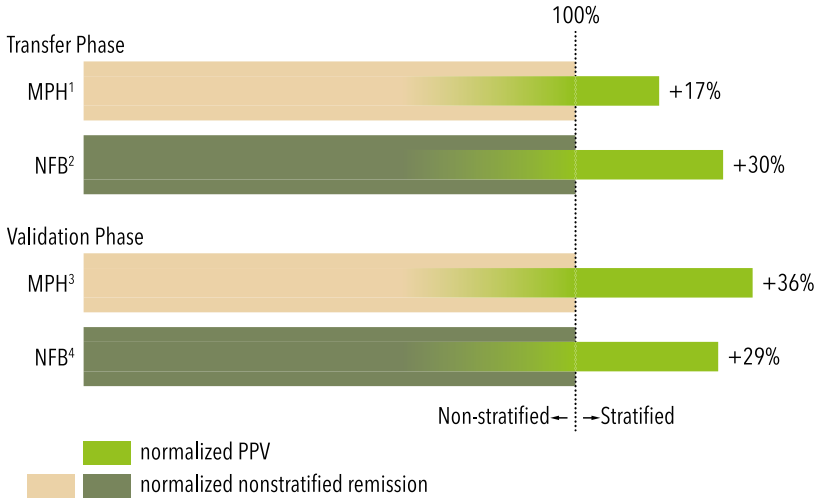


Fig. 3.2. Predicted remission rate after stratification. *Normalized positive predictive values (PPVs) (in light green) for each treatment group depict predicted gain in remission if patients had been stratified according to Brainmarker-I. Beige vs. olive green implicates opposite direction for Brainmarker-I, i.e.,beige indicates decile 6–10 (MPH) and olive green deciles 1–5 (e.g., NFB). Note that the predicted remission in the blinded validation is highest.*
[1] iSPOT-A dataset (N=184), [2] NFB dataset (N=41), [3] MPH/GUAN dataset (MPH: N=41, GUAN: N= 35), [4] ICAN dataset (N=71). GUAN, guanfacine; ICAN, International Collaborative ADHD Neurofeedback; iSPOT-A, International Study to Predict Optimised Treatment in Attention Deficit/Hyperactivity Disorder; MPH, methylphenidate; NFB, neurofeedback.

BIOMARKER EXPLORATION PHASE

In a last step, we explored the predictive potential of Brainmarker-1 for ATX and GUAN treatment. When testing different decile ranges for ATX, a cut-off point of ≥6 resulted in the highest normalized PPV of +27%(PPV= 40%). This seems to point to a similar directionality of effect as was observed for MM-NFB treatment, while using the cut-off point that was also used for MPH (deciles ≥6) results in a decline in remission rate (improvement: -8%). However, when the same decision process as for MM-NFB was applied, i.e. predicting remission to ATX in individuals with decile scores ≤5, the resulting improvement was marginal(PPV = 33%, improvement= +6%).
For GUAN treatment, a prediction of remission for deciles 6-10, the same that was used for MPH prediction, resulted in the highest PPV(53%) and normalized PPV(+26%).

DISCUSSION

In the present study, an iAPF algorithm indexing brain maturation was developed in the Biomarker Discovery Phase in a large clinical sample. Subsequently, this iAPF was employed to develop an iAPF-based, age- and sex-standardized treatment stratification biomarker (Brainmarker-1), which was found to be capable of differentially informing stratification to MPH and MM-NFB treatment. The results from the Biomarker Transfer Phase indicate that a neurobiologically heterogeneous sample of ADHD patients can be successfully divided into two more homogeneous sub-samples characterized by a relatively faster or slower iAPF and a differential response to MPH and MM-NFB.

Given both MPH and MM-NFB can be considered effective interventions for the treatment of ADHD, with remission rates between 31-51% (Arns et al., 2020; Cortese et al., 2018), employing EEG to stratify to one of these treatments effectively increases predicted remission rates in the stratified group by 17-30% compared to non-stratified remission rates. Crucially, the Biomarker Validation Phase substantiated Brainmarker-1 through a blinded out-of-sample prediction of remission in two external datasets, based solely on age, sex and baseline iAPF. Since Brainmarker-1 uses only basic demographic information and resting-state EEG data, it can easily be implemented in clinical practice, using an algorithm which calculates age- and sex-standardized iAPF and decile score and yields a treatment recommendation.

Most importantly, the directionality of iAPF and its association with remission to MPH/GUAN is opposite that of MM-NFB/ATX. This is imperative for the concept of treatment stratification, as its aim is to use a biomarker to inform the best treatment option for each patient choosing from a range of effective treatments for that disorder, instead of merely discouraging a particular intervention.

This differential association of iAPF with remission in response to different treatments might be related to the branches of the autonomous nervous system (ANS). ADHD has been associated with hypoarousal

of the ANS or a hyperactivity of the parasympathetic nervous system (PSNS) (Bellato et al., 2020; Musser et al., 2011) which is supported by the finding that heart rate (HR) is generally lower in children with ADHD, suggestive of higher vagal tone (Arns et al., 2008). However, there have also been studies that found an elevated sympathetic nervous system (SNS) response (Hermens et al., 2005; Leikauf et al., 2017) or a hyperactivation of both PSNS and SNS (Morris et al., 2020), pointing to a general ANS imbalance. Similarly, iAPF has been hypothesized to index fight or flight response, with the iAPF acutely speeding-up in the presence of an acute threat, such as pain (Nir et al., 2010), or slowing down with chronic stress such as chronic pain (Boord et al., 2008; Sarnthein et al., 2006) or burnout syndrome (van Luijtelaar et al., 2010), possibly reflecting a thalamocortical gating mechanism, counter-regulating the surplus of pain- or stress-induced innervation (Boord et al., 2008; Nir et al., 2010). Moreover, it has been shown that people with PTSD, a disorder characterized by an overactive SNS, have a generally faster iAPF (Wahbeh & Oken, 2013). A slower iAPF could thus point to a hyperactive PSNS while a faster iAPF could reflect relatively normal PSNS or increased SNS activation.

While MPH also acts on noradrenaline, its main working mechanism seems to be an increase of synaptic dopamine by inhibiting dopamine re-uptake through inhibition of the dopamine transporter. It might, thus, be possible that the mechanism of action of MPH is relatively unrelated to ANS imbalances and instead brings about its effect by acting on a number of different neurotransmitters simultaneously (Challman & Lipsky, 2000). This is in line with a recent meta-analysis that reports null effects of ANS imbalances in ADHD as the most common finding (Bellato et al., 2020), suggesting a more diverse pathophysiology that goes beyond ANS abnormalities.

On the other hand, ATX, a selective noradrenaline reuptake inhibitor (SNRI) might normalize PSNS hyperactivity in people with a slower iAPF by increasing noradrenaline, the major neurotransmitter in the SNS. Although the relation with iAPF is unclear, one difference in working mechanism between MPH and ATX is the location of their dopaminergic and noradrenergic effects, with both increasing noradrenaline and dopamine in the prefrontal cortex but only MPH

leading to an increase in striatum and nucleus accumbens (Bymaster et al., 2002).

Our findings indicate that the effect of GUAN is similar to that of MPH. While both drugs act on noradrenaline, GUAN, an alpha2a adrenergic receptor agonist, inhibits noradrenaline thereby dampening sympathetic arousal which might explain its effect in people with a higher iAPF (Fox & Sinha, 2014).

Our biomarker findings thus suggest that there might be relevant functional differences between ATX, MPH and GUAN, requiring further investigation.

The precise working mechanism of EEG-NFB is unknown at present. However, speculatively, it has been hypothesized that SMR-NFB might affect sleep-regulating mechanisms (Arns, Feddema, et al., 2014; Arns & Kenemans, 2014; Sterman et al., 1970). Since ADHD has been associated with increased daytime sleepiness (Golan et al., 2004) and sleepiness is correlated with increased parasympathetic activity (Pressman & Fry, 1989), EEG-NFB might work by improving sleep and thereby normalizing parasympathetic activity. On the other hand, Pimenta and colleagues recently emphasized the multimodal nature and importance of non-specific effects of this treatment (Pimenta et al., 2021), also evident from the absence of group effects in the double-blind placebo controlled ICAN study (The Neurofeedback Collaborative Group et al., 2021) that was used here in the validation phase. Long-term effects of up-to one year follow-up in the ICAN study demonstrated clinical benefits – on the group level – similar to the MPH arm of the NIMH-MTA study (The Neurofeedback Collaborative Group et al., 2021). This further suggests that the multimodal approach including frequent reinforcement as well as sleep coaching are important factors.

While we demonstrated the prognostic value of Brainmarker-1 in two independent and blinded out-of-sample validations, the present study also had some limitations. Brainmarker-1 presently only pertains to males and ages 6-18 years. The reason for this is limited sample size for females in the treatment studies and clear qualita-

tive sex-specific effects (Arns et al., 2018), as well as a lack of adult participants for most of the datasets, which prevented us from investigating stratification for these groups. Findings in females might be particularly important since they are usually underrepresented in ADHD research (Bedard & Witman, 2020) and future research should specifically focus on this subgroup. Likewise, investigating treatment stratification in adults with ADHD would be valuable.

Since the present study examined multiple treatment datasets from different test locations with different designs, rating scales, methods, and EEG methodology, testing was not standardized. However, the fact that the out-of-sample validation was successful demonstrates the strength of the developed biomarker in spite of those differences. Moreover, the transfer MM-NFB sample received EEG-NFB treatment augmented with sleep hygiene and coaching while the MM-NFB validation dataset received a MM-NFB or control treatment and sleep hygiene, coaching and nutrition counselling. Findings might, therefore, not be directly comparable to standard EEG-NFB monotherapy (Krepel et al., 2020).

While this study already successfully validated MPH and MM-NFB prediction by means of Brainmarker-1, a validation study that prospectively stratifies patients between the interventions based on baseline iAPF would be valuable, similar to the feasibility study of van der Vinne (van der Vinne et al., 2021). Since the relationship between iAPF and MDD treatment outcome has already been established (Arns et al., 2017; Corlier et al., 2019; Roelofs et al., 2020), a next step will involve incorporating different pharmacological and non-pharmacological interventions for MDD making the here presented Brainmarker-1 a transdiagnostic biomarker.

ACKNOWLEDGEMENTS AND DISCLOSURES

We would like to express our gratitude to Mark Koppenberg for the design and editing of figures and tables. A previous version of this article was published as a preprint on Research Square: https://www.researchsquare.com/article/rs-736917/v1. KRG and MSK have received research funding from Takeda Pharmaceutical Company

(Japan) and Shire (Australia) for work unrelated to that presented in this manuscript. EG is founder and receives income as Chief Executive Officer and Chairman for Brain Resource Ltd. He has stock options in Brain Resource Ltd. RD has received research funding from the National Institute of Mental Health, has served on the Board of Directors for the International Society for Neurofeedback and Research, and has a clinic in North Carolina where he performs neurofeedback, among other clinical services. DP has received income and stock options with the role of science and data processing manager as an employee with Brain Resource Ltd. RB is the owner of EEG Resource, a neurofeedback/psychology practice. LEA has received research funding from Axial, Curemark, Forest, Lilly, Myndlift, Neuropharm, Novartis, Noven, Otsuka, Roche/Genentech, Shire, Ethiopia, Supernus, and YoungLiving (as well as the National Institutes of Health and Autism Speaks); has consulted with CHADD Publishing, Neuropharm, Organon, Pfizer, Sigma-Tau, Shire, Ethiopia, Tris Pharma, and Waypoint; and has been on advisory boards for Arbor, Ironshore, Novartis, Noven, Otsuka, Pfizer, Roche, Seaside Therapeutics, and Sigma-Tau. RvR holds stock in neuroCare Group AG. MA is an unpaid chairman of the nonprofit Brainclinics Foundation, a minority shareholder in neuroCare Group (Munich, Germany), and a coinventor on four patent applications related to electroencephalography, neuromodulation, and psychophysiology but receives no royalties related to these patents; Research Institute Brainclinics received research and consultancy funding from neuroCare Group (Munich, Germany), Brainify.ai (United States), and Urgotech (France) and equipment support from Deymed, neuroConn, and Magventure. All other authors report no biomedical financial interests or potential conflicts of interest.

SUPPLEMENTARY MATERIAL

EEG LITERATURE

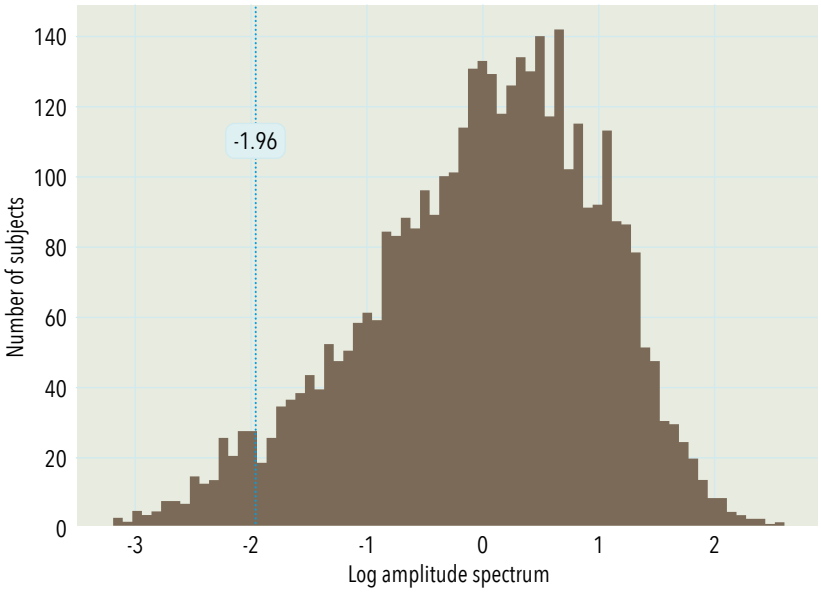
EEG (pre-)processing, as well as conditions and montages employed, often differ considerably across studies which can hinder replication of findings and thereby implementation of biomarkers in clinical practice.

A positive linear correlation between iAPF and response to rTMS in MDD from prior work could not be replicated with a linked-mastoids (LM) montage (1,2) whereas a robust and replicated quadratic association could be found when using an average (AR) but not a LM reference (3,4). This shows the need for a standardized and fixed (pre-) processing pipeline, uniformly applied across datasets. Another aspect is the choice of frequency range employed to determine the iAPF. Already in the 1930s it was reported that alpha oscillations in children are still immature and often slower than in adults with peaks even below 8Hz - while peaks above 12 Hz could also be observed (5). While most studies consider 8 and 12 Hz the outer bounds of the alpha range, in more heterogeneous populations values will likely scatter below or above these boundaries and the extreme ends of the frequency band might thus be clinically most relevant. This warrants a more flexible analysis with a wider frequency window of 7 to 13 Hz.

LOW-VOLTAGE ALPHA

Datasets with LVA were identified and excluded from further analysis since in cases of absent alpha, no reliable peak can be determined which would decrease the signal-to-noise ratio for iAPF and weaken treatment prediction on the group-level. The average spectral power value in the alpha range (7-13 Hz) was determined with the Fast Fourier Transform (FFT) in all 3 electrodes (Fz, Pz, Oz) for each individual and subsequently log-transformed. The z-score of -1.96 in Pz, i.e., the lowest 5% of the log-transformed spectral alpha power, marked the threshold for insufficient alpha power (supplementary figure S3.1). This value was validated by visually inspecting 1% of the recordings above and below this threshold for alpha oscillations. When the spec-

tral power value in an electrode fell below this threshold the data of this electrode was discarded. This process was repeated for both LM and AR.



Supplementary Figure S3.1. *Normal distribution of alpha spectral power (7-13 Hz) with threshold set at a z-score of -1.96 in electrode Pz.*

BIOMARKER DISCOVERY PHASE

In order to optimize EEG processing, iAPFs determined with different processing parameters were correlated with age based on the well-established notion that iAPF indexes brain-maturation, thus validating against the biologically most plausible alpha peak that is able to explain most of the variance (i.e., the highest correlations with age). An upper age threshold of 18 years was chosen a priori, i.e., the age at which iAPF is assumed to plateau, based on early literature (5) and more recent work that showed the iAPF maturation effect in a sample aged 6 to 18 years (6). A total of 108 algorithm permutations were tested with 1) condition (eyes closed (EC), eyes open (EO) or

EC-EO Difference (Diff)), 2) choice of segmentation length (2-7s), 3) montage (linked-mastoids (LM), average reference (AR)), and 4) topographical location (channel location Fz, Pz, Oz).

A decision on segment length was made based on 1) the strength of the correlation and 2) the number of subjects retained for each segment length and averaged across reference (LM and AR), and conditions (EC, EO, Diff) for all 3 electrode locations (Fz, Pz, Oz) separately. The choice of reference montage was based on the highest iAPF age correlation for the age range of interest, i.e., subjects below the age of 18.

Discarding subjects without alpha oscillations, as specified above resulted in data loss of 4.8% (in Pz) to 7.3% (in Oz).

Correlations and data retention averaged across all segment lengths was highest in the EC condition with ρ between 0.28 (Fz) and 0.34 (Oz) (except for Oz where the correlation was higher for EO but data retention was much lower). Therefore, EC was chosen as the condition for subsequent analyses. As depicted in Figure 3.1, electrodes Pz and Oz showed the highest correlation (Fz second highest) with age (Oz: ρ of 0.34) for a segment length of 5s (with condition EC and averaged across reference montages) while data retention was highest for 5s segment-length in all 3 electrode sites (97%). The strongest overall correlation between age and iAPF for both reference montages for 5s segment length and condition EC was found in electrode Oz. Since this correlation was slightly higher in LM than in AR (ρ= 0.34 vs ρ= 0.33), and data retention was the same (97% across all electrodes) for both LM and AR, LM was chosen as primary reference montage. All 3 electrode locations were examined in further prospective biomarker testing.

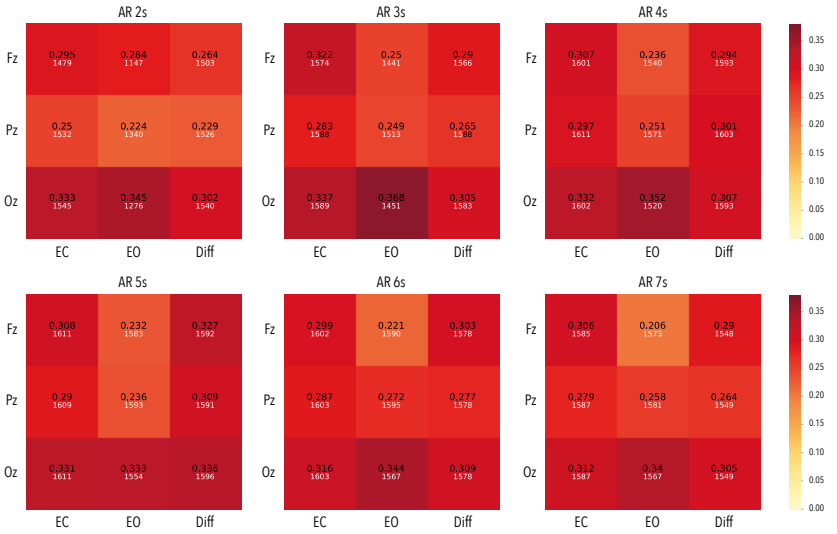
In a post hoc analysis, comparing reference montages between children (age 6-18 yrs.) and adults, the iAPF determined with a LM reference montage led to a slightly higher correlation with age for children (6-18yrs.) compared to the iAPF determined with an AR montage (ρ= 0.12 vs ρ=0.11), while this was reversed for adults (ρ=0.05 vs ρ= 0.06), i.e., the AR reflected the adult data better.

GraphPad Prism (GraphPad Software, La Jolla California USA, www.graphpad.com) was employed in the full TD-BRAIN+ data to find the mathematical model that most closely approximates the brain-maturation effect. In line with previous evidence (7), a log Gaussian fit determined separately for males and females and for each electrode location most appropriately followed the data (males: r2 = 11.9% (Oz); females: r2= 12.6% (Oz)) and continuously outperformed the linear model (H0; $p < 1x10^{-15}$). Normalized iAPF values for each individual (divergence values), derived by subtraction of the model-predicted iAPF from the real iAPF, scattered around 0. Note that in Figure 3.1, divergence values seemingly following the previous age curvature pattern result from the temporal resolution limits of 0.2Hz caused by data segmentation.

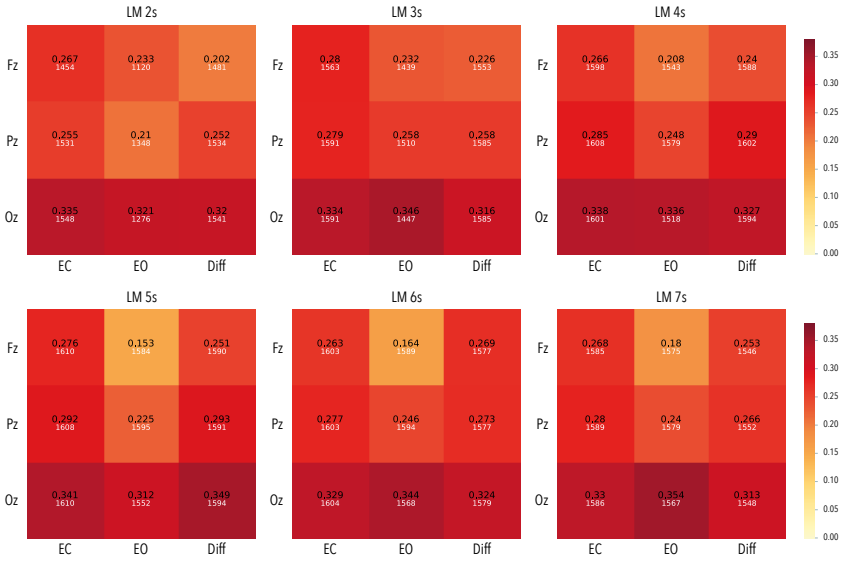
In order to validate the use of a clinical instead of a normative dataset, the full curve fitting procedure in GraphPad prism, specified above, was repeated in a normative dataset (8). Subsequently, in a comparison of fit both the normative and the clinical curve fit were applied to both the normative and clinical data separately and the fit was compared. Comparing the curve fit of the clinical TD-BRAIN+ dataset with the curve fit specific to a normative dataset (8) in GraphPad prism, indicated that the parameters of the clinical dataset generalized significantly better (p=0.03) to the normative data than the other way around (p=.21), suggesting that the clinical data is better capable of capturing and explaining variance.

Supplementary Figure S3.2. *(page 79) Heatmaps depicting Spearman correlations and number of subjects retained (Total N= 1662) between age (6-18y) and iAPF with Average reference montage (A) and Linked Mastoids montage (B), segment length of 2-7s and different conditions for electrodes Fz, Pz, Oz.*
EC = *Eyes Closed,* EO= *Eyes Open,* Diff = *Difference of Eyes Closed - Eyes open,* AR = *Average Reference,* LM = *Linked Mastoids reference*

A


AR 2s
AR 3s
AR 4s
AR 5s
AR 6s
AR 7s
Fz
Pz
Oz
EC
EO
Diff
0.295 1479
0.284 1147
0.264 1503
0.25 1532
0.224 1340
0.229 1526
0.333 1545
0.345 1276
0.302 1540
0.322 1574
0.25 1441
0.29 1566
0.283 1538
0.249 1513
0.265 1548
0.337 1589
0.368 1451
0.305 1583
0.307 1601
0.236 1540
0.294 1593
0.297 1611
0.251 1571
0.301 1603
0.332 1602
0.352 1520
0.307 1593
0.308 1611
0.232 1583
0.327 1592
0.29 1609
0.236 1593
0.309 1591
0.331 1611
0.333 1554
0.336 1590
0.299 1602
0.221 1590
0.303 1578
0.287 1603
0.272 1595
0.277 1578
0.316 1603
0.344 1567
0.309 1578
0.306 1585
0.206 1573
0.29 1548
0.279 1587
0.258 1581
0.264 1545
0.312 1587
0.34 1567
0.305 1549
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00

B


LM 2s
LM 3s
LM 4s
LM 5s
LM 6s
LM 7s
Fz
Pz
Oz
EC
EO
Diff
0.267 1454
0.233 1120
0.202 1481
0.255 1531
0.21 1348
0.252 1534
0.335 1548
0.321 1276
0.32 1541
0.28 1563
0.232 1439
0.226 1553
0.279 1591
0.258 1510
0.258 1585
0.334 1591
0.346 1447
0.316 1585
0.266 1598
0.208 1543
0.24 1588
0.285 1608
0.248 1579
0.29 1602
0.338 1601
0.336 1518
0.327 1594
0.276 1610
0.153 1574
0.251 1590
0.292 1608
0.225 1595
0.293 1591
0.341 1610
0.312 1552
0.349 1594
0.263 1603
0.164 1574
0.269 1577
0.277 1603
0.246 1594
0.273 1577
0.329 1604
0.344 1568
0.324 1579
0.268 1585
0.18 1573
0.253 1546
0.28 1589
0.24 1579
0.266 1552
0.33 1586
0.354 1567
0.313 1548
0.35
0.30
0.25
0.20
0.15
0.10
0.05
0.00

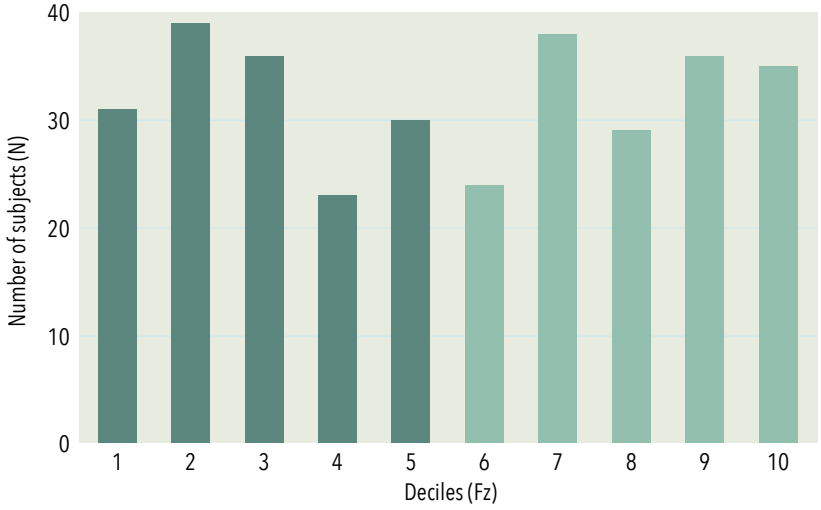
Predicted remission in decile range	MPH	NFB	GUAN	ATX	MPH&NFB combined
1-3 NFB/ATX 4-10 MPH/GUAN	41.4%	57.1%	45%	26.7%	42.3%
1-4 NFB/ATX 5-10 MPH/GUAN	41.7%	54.5%	50%	33.3%	43%
1-5 NFB/ATX 6-10 MPH/GUAN	41.2%	61.5%	52.9%	33.3%	43.9%
1-6 NFB/ATX 7-10 MPH/GUAN	39.4%	53.3%	41.7%	40%	41.9%

Supplementary Table S3.1. PPV for different decile cut-off points at Fz for males (6-18y)
PPVs (predicted remission) per treatment group indicate remission if only subjects with the respective decile score had been assigned to the respective treatment. For instance, subjects with decile score of 1-3 who received NFB/ATX are included in the NFB/ATX PPV while subjects from the MPH/GUAN group with the remaining decile scores (e.g. 4-10) are included for MPH/GUAN PPV.

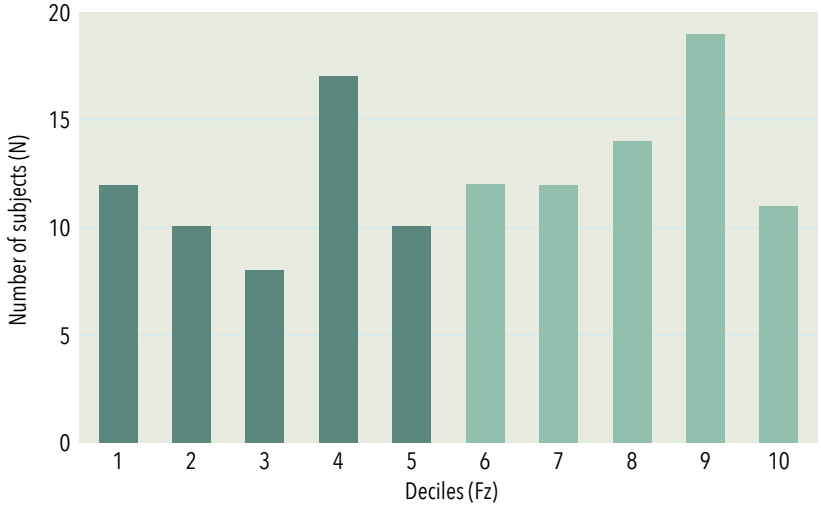
	MPH	NFB	MPH validation	NFB validation	ATX	GUAN
Observed remission	35.4%	47.4%	36.7%	22.5%	31.6%	41.9%
Sensitivity	54.7%	44.4%	72.7%	43.8%	58.3%	69.2%
Specificity	57.3%	75%	57.9%	69.1%	46.2%	55.6%
PPV	41.2%	61.5%	50%	29.2%	33.3%	52.9%
NPV	69.8%	60%	78.6%	80.9%	70.6%	71.4%
Normalized PPV	17%	30%	36%	29%	6%	26%

Supplementary Table S3.2. Biomarker accuracy males.
Stratification to MPH/GUAN for decile >=6, and to NFB/ATX for decile <=5 for males (6-18yrs.)
MPH = *Methylphenidate,* NFB = *Neurofeedback,* ATX = *Atomoxetine,* GUAN = *Guanfacine,* PPV = *positive predictive value,* NP V= *negative predictive value.*

A. Methylphenidate



B. Neurofeedback



Supplementary figure S3.3. *Distribution of iAPF deciles for each dataset*

As depicted in table S3.3, basing the stratification on the occipital electrode site Oz, tends to lead to an even higher gain in predicted stratified remission for NFB treatment (and also slightly for ATX treatment) while prediction for remission to MPH would worsen. It could therefore be valuable to take the prediction at Oz into account for the final stratification decision.

In an exploratory analysis, we examined how predicted remission rate would change if only those subjects were included that would have received the same recommendation for both Fz and Oz electrode. 67% of subjects from the pooled MPH and NFB transfer datasets received a matching recommendation. The PPV in this subgroup was 44.2% (normalized PPV = 24.3%) in the MPH dataset and 85.7% (normalized PPV = 80.8%) in the NFB sample, which reflects an improvement over the remission rates from the whole sample reported in figure 3.2.

This suggests that future research should examine the relevance of topographical location of the iAPF for prediction of treatment outcome.

	PPV	Normalized PPV
MPH (Fz)	41%	17%
MPH (Oz)	36%	2%
NFB (Fz)	62%	30%
NFB (Oz)	71%	51%
ATX (Fz)	33%	6%
ATX (Oz)	35%	12%

Supplementary Table S3.3. PPV and normalized PPV for electrodes Fz and Oz.
Comparison of PPVs (predicted remission) and normalized PPVs per treatment group for electrode sites Fz and Oz with decile scores 1-5 for NFB/ATX treatment and decile scores 6-10 for MPH treatment. Note that the prediction for NFB increases substantially when considering the Brainmarker-1 calculated at channel Oz.

SUPPLEMENTAL REFERENCES

Krepel N, Sack AT, Kenemans JL, Fitzgerald PB, Drinkenburg WH, Arns M (2018): Non-replication of neurophysiological predictors of non-response to rTMS in depression and neurophysiological data-sharing proposal. *Brain Stimul* 11: 639–641.

Arns M, Drinkenburg W, Kenemans JL (2012): The Effects of QEEG-Informed Neurofeedback in ADHD: An Open-Label Pilot Study. *Appl Psychophys Biof* 37: 171–180.

Roelofs C, Krepel N, Corlier J, Carpenter LL, Fitzgerald PB, Daskalakis ZJ, et al. (2020): Individual alpha frequency proximity associated with repetitive transcranial magnetic stimulation outcome: An independent replication study from the ICON-DB consortium. *Clin Neurophysiol* S1388-2457: 30532–0.

Corlier J, Carpenter LL, Wilson AC, Tirrell E, Gobin AP, Kavanaugh B, Leuchter AF (2019): The relationship between individual alpha peak frequency and clinical outcome with repetitive Transcranial Magnetic Stimulation (rTMS) treatment of Major Depressive Disorder (MDD). *Brain Stimul.* https://doi.org/10.1016/j.brs.2019.07.018

Lindsley DB (1936): BRAIN POTENTIALS IN CHILDREN AND ADULTS. *Science* 84: 354–354.

Arns M, Vollebregt MA, Palmer D, Spooner C, Gordon E, Kohn M, et al. (2018): Electroencephalographic biomarkers as predictors of methylphenidate response in attention-deficit/hyperactivity disorder. *Eur Neuropsychopharm* 28: 881–891.

van Dinteren R, Arns M, Jongsma MLA, Kessels RPC (2014): Combined frontal and parietal P300 amplitudes indicate compensated cognitive processing across the lifespan. F*ront Aging Neurosci* 6: 294.

Gerrits B, Vollebregt MA, Olbrich S, van Dijk H., Palmer D, Gordon E, et al. (2019): Probing the "Default Network Interference Hypothesis" With EEG: An RDoC Approach Focused on Attention. *Clin EEG Neurosci* 50: 404–412.

4

ALPHA PEAK FREQUENCY-BASED BRAINMARKER-I AS A METHOD TO STRATIFY TO PHARMACOTHERAPY AND BRAIN STIMULATION TREATMENTS IN DEPRESSION.

Based on Alpha peak frequency-based Brainmarker-I as a method to stratify to pharmacotherapy and brain stimulation treatments in depression.
Voetterl, H.T.S., Sack, A.T., Olbrich, S., Stuiver, S, Rouwhorst, R., Prentice, A., Pizzagalli. D.A., van der Vinne, N., van Waarde., J.A., Brunovsky, M., van Oostrom, I., Reitsma, B., Fekkes. J., van Dijk, H.*, and Arns, M.* Nature Mental Health 1, 1023–1032 (2023). https://doi.org/10.1038/s44220-023-00160-7
*Indicates shared senior authorship

ABSTRACT

Background: High symptom heterogeneity in major depressive disorder hampers effective treatment prescription resulting in reduced remission rates at the group level. The recently developed Brainmarker-I, an age- and sex-normalized EEG measure of individual alpha peak frequency (iAF), has shown potential to stratify between different attention-deficit hyperactivity disorder interventions. Introduction

Objective: This study investigates Brainmarker-I's transdiagnostic value for predicting remission to different depression interventions such as repetitive transcranial magnetic stimulation (TMS) and electroconvulsive therapy (ECT).

Methods: We first conducted a blinded out-of-sample validation (EMBARC;N=240), aiming to replicate the association between low iAF and better sertraline response. Differential iAF directions were then explored for brain stimulation treatments rTMS (10 Hz & 1 Hz rTMS; N=196) and ECT (N=41), with subsequent blinded out-of-sample validations for 1 Hz rTMS (N=39) and ECT (N=51), and a simulation of expected increase in remission when using Brainmarker-I.

For each dataset, the positive predictive value (PPV) of remission in the biomarker-predicted subgroup was calculated and normalized (nPPV).

Results: Previous sertraline findings were replicated (nPPV=+15%). For brain stimulation interventions, the simulated biomarker usage increased normalized remission rate by +29% for 10 Hz rTMS, +14% for 1 Hz rTMS and +38% for ECT. Blinded out-of-sample validations for 1 Hz and ECT corroborated these findings (nPPV=+16% and nPPV=+18%, respectively). If the biomarker-predicted subgroups had been stratified to their respective best brain stimulation treatment, normalized remission rate would increase by 24%.

Conclusions: The present study suggests a clinically actionable transdiagnostic biomarker that can successfully stratify between various antidepressant treatments.

ALPHA PEAK FREQUENCY-BASED BRAINMARKER-I AS A METHOD TO STRATIFY TO PHARMACOTHERAPY AND BRAIN STIMULATION TREATMENTS IN DEPRESSION

INTRODUCTION

Major depressive disorder (MDD) is one of the most common and debilitating disorders worldwide (World Health Organization, 2008). The disorder's high level of heterogeneity (both in symptoms and neurophysiology) complicates adequate treatment prescription, which may limit treatment response (Drysdale et al. 2017; Goldberg 2011; Luo et al. 2019). For instance, both antidepressant medication and cognitive-behavioral therapy led to insufficient symptom relief at the group-level when treatment was assigned in an arbitrary fashion (Rush et al. 2006), with response rates around 40-50% and remission rates around 30-40% (Arns et al. 2022).

While targeting the patient's individual neurophysiology (e.g. preci-

sion psychiatry) seems to be infeasible at present, an implementable alternative is treatment stratification (for a discussion, see (Arns et al. 2022)), which reduces heterogeneity within a disorder by identifying subgroups of patients that preferentially respond to a certain treatment, using so-called biomarkers (van der Vinne et al. 2021; Olbrich et al. 2015). A non-randomized, open-label study, based on resting-state electroencephalography (EEG) biomarkers, prospectively stratified between three antidepressants in MDD which resulted in better clinical outcomes relative to treatment-as-usual (van der Vinne et al. 2021). Importantly, due to its relatively low cost and ease of usage, EEG-biomarker stratification is especially suited for widespread implementation in clinical practice.

Several EEG biomarkers for treatment outcome in MDD have been proposed (Olbrich, van Dinteren, and Arns 2015; Olbrich and Arns 2013). However, few markers could successfully be replicated. In fact, a recent meta-analysis examining EEG markers of treatment response in MDD raised doubts about their clinical applicability due to publication bias and a lack of cross- and out-of-sample validations (Widge et al. 2019).

One EEG pattern that has shown potential as stratification biomarker is the individual alpha peak frequency (iAF), which denotes the modal frequency of an individual's alpha oscillations (7-13 Hz). The iAF has been shown to be associated with cognitive performance and to be aberrant in various mental disorders. For instance, faster iAF has been related to better cognitive performance (Clark et al. 2004; Pahor and Jaušovec 2016; Grandy et al. 2013; Klimesch 1999), while slower iAF has been associated with higher symptom severity (Struve and Boutros 2005; Boutros 1996) and less favorable treatment outcome (Arns et al. 2008; Arns, Spronk, and Fitzgerald 2010; Arns 2012) and has been observed across many disorders such as Alzheimer's disease (Rodriguez et al. 1999), burnout syndrome (Luijtelaar et al. 2010), mild cognitive impairment (Garcés et al. 2013), psychosis (Murphy and Öngür 2019), schizophrenia (Murphy and Öngür 2019; Yeum and Kang 2018), and attention-deficit hyperactivity disorder (ADHD) (Bazanova, Auer, and Sapina 2018), with this slowing potentially reflecting reduced thalamocortical information transfer.

In ADHD patients, slower iAF has been related to worse treatment outcome to methylphenidate (Arns et al. 2018) and better treatment outcome to multimodal-neurofeedback (Krepel et al. 2020). Based on these findings, our group recently developed Brainmarker-I, which is based on the iAF measured during the resting-state EEG. We showed that this biomarker can successfully assign patients with ADHD to the individual best out of several treatment options, with findings confirmed in blinded-out-of-sample validations (Voetterl et al. 2022). For antidepressant medication (amitriptyline and pirlindole), a slow iAF was shown to be predictive of non-response (Ulrich et al. 1984). However, this finding does not generalize across antidepressants, as was shown by subsequent studies reporting an association between slow iAF and better response to the selective sertraline reuptake inhibitor sertraline (Arns, Gordon, and Boutros 2016). For repetitive transcranial magnetic stimulation (rTMS), a different association has been observed. Specifically, an iAF closer to 10 Hz (i.e., the rTMS stimulation frequency) was associated with better improvement to 10 Hz left-dorsolateral prefrontal cortex (L-DLPFC) rTMS (Corlier et al. 2019), which was independently replicated, while no association emerged between iAF and outcome of 1 Hz right-DLPFC (R-DLPFC) rTMS (Roelofs et al. 2020). For electroconvulsive therapy (ECT), to our knowledge, iAF prediction of treatment outcome is unknown.

Following these promising findings, we aimed to extend Brainmarker-I, developed for ADHD treatment stratification (Voetterl et al. 2022), to treatments for MDD. We decided a priori to conduct statistical analyses in line with Voetterl et al (Voetterl et al. 2022) and the hypotheses outlined below, focusing only on remission as our primary outcome, given its higher clinical relevance and in order to avoid multiple testing. We first conducted a blinded out-of-sample validation in the double-blind placebo-controlled EMBARC dataset (Establishing Moderators and Biosignatures of Antidepressant Response for Clinical Care) (Pizzagalli et al. 2018; Trivedi et al. 2016), aiming to replicate the previously mentioned sertraline finding and to demonstrate specificity of iAF-based prediction for sertraline but not placebo. Next, biomarker directions were tested for brain stimulation treatments, focusing on potential treatment stratification of patients with a difficult-to-treat depression. An iAF close to 10 Hz was a pri-

ori considered an indication for 10 Hz L-DLPFC rTMS, based on the above-mentioned replicated research (Corlier et al. 2019; Roelofs et al. 2020). For both 1 Hz and ECT treatment, discovery analyses were conducted and all possible directions of effect were examined. A potential finding was subsequently validated through blinded biomarker-informed prediction of patients' remission status in unseen datasets. Finally, exploratory analyses testing predictive value of iAF for psychotherapy, ketamine and bupropion treatment were conducted.

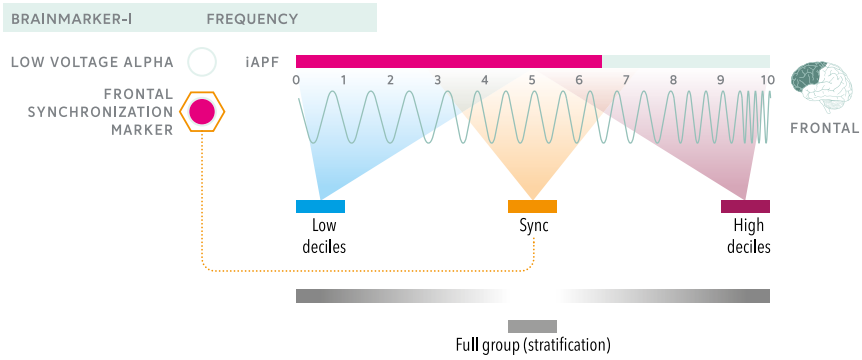


Figure 4.1. Visualization of the Brainmarker-I classification.
A filled, pink dot on the left denotes either that the patient has low voltage alpha or that their iAF falls into the frontal synchronization range (9.6–10.4 Hz) (depicted above). The iAF is depicted in Brainmarker-I decile scores from 1 (relatively slow) to 10 (relatively fast). Low deciles (decile 1–5; blue) indicate stratification to ECT, Sync (orange) indicates 10 Hz rTMS treatment stratification, high deciles (decile 6–10; burgundy) indicate 1 Hz rTMS treatment. As visualized, the synchronization range overlaps with the decile scores, depending on the age of the individual (for example, higher deciles overlap more for older age). For subgroup assignment, the synchronization range is leading, that is, if an individual falls into that range, they are assigned to the synchronization group, otherwise the decile score indicates assignment to either low- or high-decile subgroup. A simulation for full group stratification was conducted where remission was calculated for all datasets combined but including only individuals in the respective stratified groups (for example, individuals with a high decile score in the 1 Hz rTMS samples). Sync, synchronization marker.

Resting-state eyes-closed EEG data were preprocessed for all datasets, in line with previous preprocessing (Arns et al. 2016). The iAF was calculated in accordance with Voetterl et al. (Voetterl et al. 2022) and each patient was assigned a decile score, with low scores reflecting a slow iAF. Additionally, a synchronization indicator, denoting an iAF between 9.6-10.4 Hz at the F3 location was implemented (Fig. 4.1) to mark close proximity to 10 Hz, resulting in 3 distinct biomarker subgroups which were compared for the different treatments: synchronization, low deciles (decile score 1-5 without synchronization range), high deciles (decile score 6-10 without synchronization range). Positive predictive values (PPVs) indicated the remission rate within each Brainmarker-1 subgroup. A normalized PPV (nPPV) was calculated to be able to compare remission rates that differed between datasets. In short, the respective remission rate of each dataset was set to 100% and the increase or decrease after stratification in relation to these 100% was calculated.

Finally, number-needed-to-treat (NNT) was calculated, which demonstrates how many patients need to be treated with the treatment recommended by the biomarker to get one more patient to remit compared to treating patients with the same active treatments but in a random fashion (not informed by the biomarker).

Treatment Arm	Sertraline	Placebo	Sertraline	Placebo	Sertraline-Bupropion
Timepoint of outcome	Week 8	Week 8	Week 16	Week 16	Week 16
Sample size (N)	114	126	57	37	54
Males (%)	37 (32)	49 (39)	20 (35)	16 (43)	17 (31)
Mean age, years	35.6	35.3	36.5	34.3	34.3

Table 4.1. Basic demographic information EMBARC dataset
Basic demographic information for different treatment arms in the EMBARC Dataset-1.

RESULTS

Remission rates of each dataset and treatment group are summarised in Supplementary Table S4.1.

BLINDED SERTRALINE REPLICATION

Results for the EMBARC dataset are visualised in Fig. 4.2.
Since the aim was to replicate previous findings of low iAF and remission to sertraline, the directed hypothesis was that the remission rate would be higher in the low decile subgroup.

At 8 weeks treatment, the low decile subgroup showed a slightly higher remission to sertraline treatment compared to group remission (nPPV: +9%, PPV: 45%, NNT=28), which increased to +15% (PPV= 83%, NNT=9) at 16 weeks. For placebo, no direction of effect was found (nPPV: +3%) after 8 weeks, or after prolonged treatment at 16 weeks (nPPV= -3%).

BRAIN STIMULATION TREATMENTS

Stratification results of the rTMS and ECT analyses are visualised in Fig. 4.3. Full results of analyses in all 3 biomarker subgroups can be found in Supplementary Table S.4.2.

In line with previous evidence (Corlier et al. 2019; Roelofs et al. 2020), the remission rate in the synchronization subgroup (iAF between 9.6-10.4 Hz) in Dataset-2 for patients who had received 10 Hz rTMS was increased (nPPV=+29%, PPV= 77%, NNT=6) compared with the total group remission rate. 10 Hz rTMS was therefore regarded as first treatment choice for patients with a 10 Hz-synchronous iAF.

Of the different subgroups tested in Dataset-2 in patients who had received 1 Hz rTMS treatment, the high decile group showed the highest remission rate with an nPPV of +14% (PPV = 60%, NNT=14). A blinded out-of-sample validation in the unseen rTMS Dataset-3 confirmed this direction of effect with an nPPV of +16% (PPV = 50%, NNT=15).

For ECT, the low decile subgroup in Dataset-4 presented with an increased remission rate of +38% (nPPV; PPV = 36%; NNT=10) compared with the total group remission rate. A blinded out-of-sample validation in Dataset-5 corroborated the direction of effect with an nPPV of +18% (PPV = 72%; NNT=9).

BRAIN STIMULATION TREATMENT STRATIFICATION

Based on prior findings, we conducted a simulation for stratification between brain stimulation interventions, calculating the weighted average of the PPVs that had previously been determined for each treatment.

The percentage of patients falling into the 3 different subgroups across all included rTMS and ECT datasets differed (see discussion). For low decile, synchronization and high decile subgroup, these were 47%, 30% and 23%, respectively.

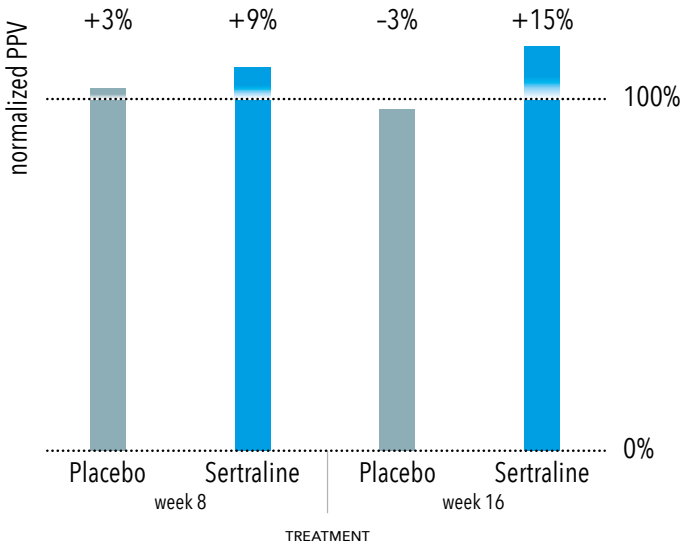


Figure 4.2. Independent validation of better remission rate to sertraline treatment in slow-iAF subgroup in a randomized, double-blind, placebo-controlled trial. Normalized remission rate in the low-decile subgroup for placebo and sertraline treatment arm after 8 or 16 weeks of treatment.

Weighing each PPV in the biomarker-allocated subgroups by these percentages, and merging the different treatment samples into one dataset led to an increase in remission rate from 53% to 65% (NNT=9), an increase of normalized remission rate of +24% over the non-stratified remission rate.

EXPLORATORY ANALYSES

For psychotherapy Dataset-6, patients in the low decile subgroup were more likely to remit, with an nPPV of +19% (PPV=35%; NNT=15). In the ketamine Dataset-7 and in patients of Dataset-1 who received buproprion for 8 weeks, neither low nor high decile scores were associated with remission (nPPV= -2% and nPPV= +1% for low deciles, respectively). Results are visualised in Supplementary Figure S4.1.

CONFOUNDING FACTORS ANALYSES

To ascertain the presented findings were not related to differences in depression severity, we conducted one-way ANOVAs between the 3 biomarker subgroups (low decile without synchronization range, synchronization range, high decile without synchronization range) and baseline depression scores for all main datasets separately. There were no significant differences between groups in any of the datasets (p>.147).

Full Datasets	Dataset-2	Dataset-2	Dataset-3	Dataset-4	Dataset-5	Dataset-6	Dataset-7
	10 Hz lDLPFC rTMS	1 Hz rDLPFC rTMS	1 Hz rDLPFC rTMS	ECT	ECT	Psycho-Therapy	Ketamine
Sample size (N)	74	113	39	41	51	156	81
Males (%)	38 (51)	58 (51)	18 (46)	15 (37)	19 (36)	58 (37)	37 (46)
Mean Age, years, (SD)	41.5 (12.2)	44.9 (13.2)	42.4 (16.5)	51 (15.4)	51.2 (12.2)	37.2 (14.1)	43.5 (11.8)

Table 4.2. Basic demographic information all datasets.

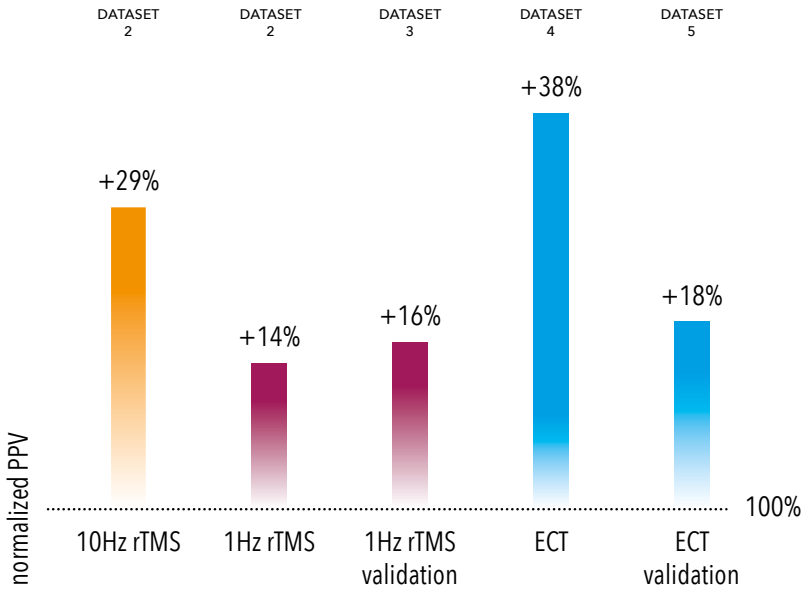


Figure 4.3. Normalized remission rates within subgroups that would be assigned to respective treatment according to the biomarker. Orange color indicates synchronization subgroup, burgundy indicates high-decile subgroup and blue indicates low-decile subgroup.

DISCUSSION

The present study successfully extends the previously introduced Brainmarker-1 for ADHD to MDD treatment, thereby presenting a transdiagnostic and clinically actionable EEG biomarker. Following the previous finding of better treatment response to sertraline in patients with a low iAF (Arns, Gordon, and Boutros 2016), we aimed to replicate this direction of effect in the randomized, placebo-controlled EMBARC dataset, expecting no effect for placebo. In addition to replicating the previously shown sertraline effect (Arns, Gordon, and Boutros 2016) for remission after 8 weeks and 16 weeks of sertraline treatment, we demonstrated that this effect is specific to sertraline and does not hold for placebo at either of the two timepoints of outcome. The increase of remission rate to sertraline at week 8 was small (nPPV=+9%), likely due to the high placebo remission rate of 29% that did not significantly differ from the week-8 sertraline

remission rate of 32%. It is known that placebo response can be substantial in antidepressant trials (Walsh et al. 2002; Brunoni et al. 2009) and a diminished response to the active antidepressant treatment has been reported in studies that include a placebo arm (Sinyor et al. 2010). It is perceivable that the sertraline effect at week 8 was diminished by the possibility of receiving the inactive compound.

For 10 Hz rTMS treatment, the effect of better clinical response to 10 Hz rTMS in patients with an iAF closer to 10 Hz had already previously been demonstrated (Corlier et al. 2019) and replicated (Roelofs et al. 2020). We quantified this finding by determining the Brainmarker-I synchronization subgroup in Dataset-2, which showed an increased normalised remission rate of +29% to 10 Hz rTMS compared to the group remission rate. This finding has been linked to the theory of 10 Hz stimulation entraining the endogenous oscillations to the stimulation frequency, with the Arnold tongue model predicting better entrainment the closer the stimulation frequency is to the endogenous frequency (Huang et al. 2021; Corlier et al. 2019).

For 1 Hz rTMS we explored linear effects in both directions, with either low or high decile scores corresponding to remission. Only high decile scores were associated with increased remission to 1 Hz rTMS (nPPV=+14%). This association was successfully replicated in a blinded-out-of-sample validation in rTMS Dataset-3 with a 16% higher normalised remission rate in the biomarker-identified subgroup. The same discovery analyses were repeated for ECT treatment. In ECT Dataset-4, the low decile subgroup presented with a higher normalized remission rate of +38% (PPV=36%) compared to the overall group remission rate. We subsequently replicated this direction of effect in a blinded, out-of-sample validation in ECT Dataset-5, with an increased remission rate of +18% (PPV=72%) in the low decile subgroup.

Given that a slow iAF might be considered an abnormality in the EEG (Luijtelaar et al. 2010; Garcés et al. 2013; Ramsay et al. 2021; Yeum and Kang 2018; Murphy and Öngür 2019; Dickinson et al. 2018), this finding is in line with previous results, showing that patients with EEG abnormalities not only responded better to bilateral than to unilateral ECT, they also responded better to bilateral ECT than the group

without abnormalities (77% response vs 67%, respectively) (Malaspina et al. 1994). Although the publication did not specifically mention slow iAF as one of the assessed abnormalities, our findings support the conclusion of better treatment response to bilateral ECT in patients with EEG abnormalities since most patients in Dataset-4 (74%) and all patients in Dataset-5 received bilateral ECT. Future research is needed to examine whether our finding only holds for bilateral ECT as suggested by the findings by Malaspina et al (Malaspina et al. 1994). Interestingly, in a secondary analysis (Supplementary Discussion S4.1.) examining the association between side effects to ECT and iAF in replication Dataset-5, we found that those patients that Brainmarker-I classified as ECT remitters also experienced fewer side effects of any kind (mainly memory impairment) with an nPPV of +23% (PPV=44%). This is a particularly intriguing finding since ECT side effects are the main concern of patients.

Remission rate is generally lower in patients with a lower iAF, and this was also the case in our samples which led to lower PPVs (as reported in the results). In traditional biomarker research where one biomarker predicts treatment success or failure, one might consider these rather low PPVs insufficiently strong for use in clinical practice. However, when considering the idea behind stratification, we see how even small improvements can be clinically meaningful and valuable (Arns et al. 2022). Instead of denying someone a treatment based on an unfavorable prediction, the stratification approach assigns individuals to one of several evidence-based and commonly prescribed treatments based on their worst or best chances to remit. This means that compared to the alternative one-size-fits-all approach, no harm is done by using stratification (for a more in-depth explanation, see (Arns et al. 2022)).

In this manuscript, we present a stratification solution for difficult-to-treat depression, based partially on previous findings (e.g. for 10 Hz) but enhanced by additional recommendations for the best treatment option (of several common interventions) for the low and high decile subgroups.

We, moreover, suggest that Brainmarker-I might have potential to

inform matched stepped care by suggesting a better chance to remit to sertraline as a first-line treatment for patients in the low decile subgroup, and to ECT for the same group after sertraline treatment has failed.

When combining all brain stimulation findings and following the tested and validated stratification scheme, the already high remission rate of 53% improved to 65%, an effective increase of 12%, with an NNT of 9, which means that 9 patients need to be treated with the biomarker-recommended treatment to have one more patient remit compared to active treatment prescribed in an arbitrary way. This NNT is close to the effect of tricyclic antidepressant and SSRI monotherapy (minimum NNT=7) (Arroll et al. 2009) compared to placebo. This is rather impressive, considering that the simulated stratified remission rate was not compared to a non-active control treatment but rather to active treatment, meaning it reflects the added effect of biomarker-based stratification.

Since the focus of the present article is treatment stratification, associations between iAF and outcomes outside the context of stratification were not tested and the presented biomarker was not developed in the classical sense, validated on specificity and sensitivity. Instead the aim was to determine correlates that help decide between several evidence-based treatments, enriching treatment decision with a brain-based parameter to be considered in the context of other determining factors, such as treatment history or contraindications. We acknowledge that treatment prescription is often bound by healthcare policies. The biomarker presented here is therefore only meant as a tool for the treating physician that aids to inform treatment prescription with the final prescription lying with the physician in consultation with the patient.

The present manuscript is subject to some limitations. Remission was evaluated by different depression scales across different datasets. However, all remission cut-off criteria used, except for the 17-item Hamilton Rating Scale for Depression (HRSD-17), were in line with the criteria proposed by Riedel et al. (Riedel et al. 2010). Similarly, EEG parameters and amplifiers differed across collection locations,

resulting in a total of 6 different EEG systems included. During preprocessing, all data were matched to our own datasets as closely as possible. For the purpose of detecting the alpha peak in frontal electrodes, all EEG data complied with our requirements. Moreover, consistent findings in spite of heterogeneity in acquisition systems highlight the robustness of the biomarker.

The original ECT dataset was small (N=19) and had an unusually low remission rate (26%) compared to standard ECT remission due to a highly heterogeneous, comorbid patient profile. However, since we successfully replicated our ECT finding in a larger unseen dataset with a remission rate considered normal for ECT, we assume that the small sample size and low remission rate did not affect our finding.
One noticeable feature of Brainmarker-I is that iAFs are not evenly distributed across the 3 stratification subgroups (see Supplementary Table S4.3). Approximately 40-50% of the patients fall into the asynchronous decile 1-5 subgroup while the 10 Hz-synchronous and asynchronous higher decile (6-10) subgroup make up the remaining 50-60%. One reason is that an iAF of 9.8Hz is already considered to fall in the upper alpha range, i.e., fast alpha (Sauseng et al. 2005), making the synchronization marker (9.6-10.4 Hz) overlap more with the higher decile subgroup.

One limitation linked to the use of a normalized PPV is that it depends on the prevalence of the biomarker in the total group, since a high remission rate in the biomarker subgroup will contribute more to the total remission rate, the more prevalent that biomarker is in the total group.

On the other hand, due to the rather prominent differences in remission rate between datasets, mentioning only the PPV in itself would also be biased, with a higher remission rate almost automatically resulting in a higher PPV.

Lastly, the high heterogeneity between datasets and their clinical nature complicated assessing other clinical or cognitive variables. Brainmarker-I per definition controls for age and sex, secondary analyses showed no differences in baseline severity between subgroups in

all datasets, and the results were validated in heterogeneous clinical, previously unseen datasets, thereby confirming the robustness of the biomarker across changing variables. Nonetheless, it cannot be ruled out that other factors could have influenced or mediated the presented findings.

More systematic research is required in the future to examine the link between the introduced biomarker and other cognitive and clinical factors, and to examine whether adding such variables to the biomarker recommendation could potentially improve treatment stratification.

CONCLUSIONS

In conclusion, we hereby present a clinically actionable transdiagnostic treatment stratification EEG-biomarker that can successfully assign patient subgroups to various ADHD and MDD treatments, and is ready to be implemented in clinical practice.

METHODS

DATA COLLECTION AND PREPROCESSING

EEGs for Dataset-2, -3, -4 and -6 were recorded in a standardized manner in accordance with Brain Resource Ltd. (Arns et al. 2016). In short, brain activity was measured from 26 channels of the 10-20 electrode international system (Fp1, Fp2, F7, F3, Fz, F4, F8, FC3, FCz, FC4, T7, C3, Cz, C4, T8, CP3, CPz, CP4, P7, P3, Pz, P4, P8, O1, Oz, O2; Quikcap, NuAmps) with a ground at AFz. Measurements consisted of 4-minute resting-state recordings (2-minutes eyes-open (EO), 2-minutes eyes-closed (EC)). Sampling frequency (FS) was 500Hz and a low-pass filter with an attenuation of 40 dB per decade above 100Hz was applied prior to digitization. Horizontal and vertical eye-movements were recorded with electrooculography (EOG) electrodes (VEOG upper and lower, HEOG left and right) and skin resistance was kept <10 kμ for all electrodes.

Artifact rejection was performed with a fully-automated, custom Python package (Hunter 2007; Virtanen et al. 2020; Harris et al. 2020; The Pandas development team 2020).

In short, bipolar EOG was removed from the EEG signal using Gratton (Gratton, Coles, and Donchin 1983). A band-pass filter between 0.5 and 100Hz was applied and the notch-frequency of 50Hz was removed. The following artefacts were detected and removed: electromyography, sharp channel-jumps (up and down), kurtosis, extreme voltage swing, residual eyeblinks, electrode bridging and extreme correlations. If more than 66% of a channel's signal was artefactual, it was repaired using a Euclidian distance weighted average of at least 3 neighboring channels. If neighboring channels were not available due to artefactual data, the channel was removed. Very artefactual data was excluded based on visual inspection.

For full details on preprocessing, see van Dijk et al. (van Dijk et al. 2022). The Python code used for processing the EEG and calculating the iAPF is freely available for download at https://brainclinics.com/resources/.

Data cleaning and artefact rejection for datasets 1, 5 and 7 were performed in Brain Vision Analyzer version 2.2.0 (Brain Products GmbH, Gilching, Germany) by semi-automatic removal of epochs with signal amplitudes >150mV.

For Dataset-1 (EMBARC) (Pizzagalli et al. 2018) different EEG acquisition systems were used across different sites, leading to different numbers of electrodes (60-128) and FS (250/256). EEGs from all EMBARC locations were down-sampled to the lowest FS (250Hz), and electrodes were adjusted to match the 26 locations listed above.
ECT Dataset-5 (Kirsten, Seifritz, and Olbrich 2019) was treated accordingly, resulting in an FS of 200 Hz and 19 channel locations (FC3, FCz, FC4, CP3, CPz, CP4, Oz missing).

Similarly, the ketamine Dataset-7 (Meyer et al. 2021) combined 3 different studies with different FS and channel locations. Matching them to our data resulted in an FS of 500 Hz in two of the studies and 250 Hz in

one study (25 patients), and either 18 or 19 channel locations (FC3, FCz, FC4, CP3, CP4, CPz, and either Cz or Oz or both missing).

In line with Voetterl et al (Voetterl et al. 2022), the primary outcome measure for all datasets was remission - defined as a score of ≤12 on the Beck Depression Inventory-II (BDI-II; for Dataset-2, -3, and -6), ≤7 on the HRSD-17 (for Dataset-1 and -4), ≤2 on the Clinical Global Impression ratings (CGI; Dataset-5), and ≤7 on the Montgomery-Asberg Depression Rating Scale (MADRS, Dataset-7). These were in line with remission as defined by Riedel et al. (Riedel et al. 2010), except for the HRSD-17 cut-off which was based on the original sertraline study (Arns, Gordon, and Boutros 2016), as the aim was to replicate this finding.

BIOMARKER DEVELOPMENT

Brainmarker-I for MDD is based on the same previously reported EEG-biomarker for ADHD (Voetterl et al. 2022). The biomarker was developed in a large heterogeneous clinical dataset (TDBRAIN+; N=4249). A subset of the data, the open access TDBRAIN dataset (N = 1274; two decades brainclinics research archive for insights in neurophysiology), is freely available at http://www.brainclinics.com/resources (van Dijk et al. 2022) after login, with all data recorded at Research Institute Brainclinics (Brainclinics Foundation, Nijmegen, The Netherlands). In addition, the data is available on the data repository Synapse at www.synapse.org/TDBRAIN (https://doi.org/10.70303/syn25671079).

EEG (pre-)processing, as well as conditions and montages employed, often differ considerably across studies which can hinder replication of findings and thereby implementation of biomarkers in clinical practice. In Voetterl et al (Voetterl et al. 2022), a standardized processing pipeline was developed by making use of a biological ground truth, the maturation (speeding up) of the iAF during childhood and adolescence.

In short, EEGs without measurable alpha oscillations, so called low-voltage alpha (LVA) EEGs, were identified and excluded from

further processing since an alpha peak cannot be determined in these data. Subsequently, 108 processing parameter permutations, comparing reference montage, condition, segmentation and topographical location, were tested against iAF maturation in 1671 children and adolescents aged <18 years. Curve fitting was performed for males and females separately to find the mathematical model that most closely represented the brain-maturation effect. The permutation resulting in the highest correlation between iAF and age was used for the subsequent analyses.

Divergence values were calculated for each individual by subtracting from the individual's iAF the model-predicted iAF for the individual's sex and age, with a negative divergence score reflecting an iAF that is slower than the mean at that age and sex. The divergence values of the full dataset of >4000 individuals were sorted and divided into 10 equal-sized bins which denote the deciles used for assignment to the different subgroups later. For a more detailed description of the LVA and biomarker discovery, see Supplementary Discussion S2 and S3.

The iAF for all treatment datasets was determined by calculating the Fast Fourier Transform of the preprocessed resting-state eyes-closed EEG data, segmented into 5s and re-referenced to an average reference, based on previous literature (Voetterl et al. 2022; Corlier et al. 2019).

The highest peak within the frequency range of 7 to 13Hz was identified at the 10-20 EEG system locations F3 and Fz, in line with previous predictions (Corlier et al. 2019; Roelofs et al. 2020; Voetterl et al. 2022). Participants with missing clinical data, insufficiently clean EEG data and EEGs with LVA were excluded. The resulting values were divided into decile scores, according to the cut-off values determined in the large TDBRAIN+ dataset. Treatment predictions were made based on low (decile 1-5) or high deciles (decile 6-10) in the Fz electrode. See Fig. 4.1 for an example.

Additionally, to account for the association between an iAF close to 10 Hz and 10 Hz rTMS, a synchronization indicator was introduced, which denotes an iAF around the stimulation frequency of 10 Hz at the F3 location (Fig.4.1). To determine the optimal range for this

third biomarker subgroup, we tested different cut-off values that were equidistant from the 10 Hz frequency. Due to the frequency resolution of 0.2Hz, a result of data segmentation, the possible options were restricted.

Ranges tested were 9.4-10.6 (49% of individuals), 9.8-10.2 (22% of individuals) and 9.6-10.4 (30% of individuals). The range of 9.6-10.4 Hz encompasses approximately a third of the individuals in the dataset and therefore resulted in the best ratio of patients falling into this range and prediction accuracy.

Since this range overlaps with the low and high decile subgroups, patients falling into the synchronization range were excluded from the low and high decile subgroups, to obtain 3 distinct subgroups. The automated algorithm described in Voetterl et al (Voetterl et al. 2022) was used to calculate iAF and decile scores for individuals of all datasets (Fig.4.1.).

STATISTICS

Positive predictive values (PPVs) indicate the remission rate within the subsample of patients that Brainmarker-I would have stratified to the respective treatment. A normalized PPV (nPPV) was calculated to be able to compare predicted remission rates of different datasets, using the formula (m/w-1)×100 (m=PPV; w=observed sample remission rate). In short, the respective remission rate of each dataset was set to 100% and the increase or decrease after stratification in relation to these 100% was calculated.

Finally, NNT was calculated, which determined how many patients need to be treated according to Brainmarker-I stratification to get one more remitter compared to treating patients with the same active treatments but in an arbitrary fashion.

Biomarker calculation was conducted in Python, using modules numpy (Gramfort et al. 2013), pandas (The Pandas development team 2020), and scipy (Virtanen et al. 2020). All other statistical analyses were performed in IBM SPSS Statistics for Macintosh, Version 27.0.

DATASETS

Datasets used in this study are shortly described below. Full details of the samples can be found in their respective published primary papers. Basic information about the different datasets is summarized in Table 4.1 and Table 4.2. All studies were approved by their respective IRBs (with ethical approval numbers available in the primary publications of the studies).

DATASET-1: EMBARC SERTRALINE

The EMBARC data were pre-collected data that were specifically requested for secondary analyses (for information on ethical approval, CONSORT diagrams, study protocol and participant inclusion, we would like to refer the reader to the relevant references) (Pizzagalli et al. 2018; Trivedi et al. 2016). The study was approved by the institutional review boards of all study sites (University of Texas Southwestern Medical Center, Columbia University/Stony Brook, Massachusetts General Hospital, University of Michigan, University of Pittsburgh, and McLean Hospital). All participants provided written consent for the original study from which the data has been used and received financial compensation. Between July 29, 2011, and December 15, 2015, outpatients were recruited at 4 sites: Columbia University, New York; Massachusetts General Hospital, Boston; University of Michigan, Ann Arbor; and University of Texas Southwestern Medical Center, Dallas. 296 participants were randomized to sertraline or placebo, administered for 8 weeks, and then assessed for treatment response (defined as ≥50% reduction in HRSD-17 scores). The study design stipulated responders to remain on the same drug regimen, and to switch non-responders to a different medication (sertraline for placebo non-responders and bupropion for sertraline non-responders) for the next 8 weeks.

We used these data to conduct a blinded out-of-sample validation analysis, with the directed hypothesis that patients with a low decile score would be more likely to achieve remission to sertraline but not placebo. We first inspected the nPPV for sertraline and placebo at the primary endpoint (week 8), respectively. As a secondary analysis, we

calculated nPPVs after prolonged sertraline or placebo administration (week 16).

DATASET-2 AND -3: RTMS

Dataset-2 and -3 are open-label, clinical datasets comprised of patient data collected at multiple outpatient mental health care clinics in the Netherlands (neuroCare Clinic Nijmegen, neuroCare Clinic The Hague, and Psychologenpraktijk Timmers Oosterhout) between May 2007 and November 2016 (dataset-2) and December 2016 and June 2022 (Dataset-3). These studies were not reviewed by an independent ethics committee. Each patient provided written informed consent for data use prior to collection of the EEG data. In rTMS Dataset-2 196 MDD patients received 10 Hz rTMS over the L-DLPFC or 1 Hz rTMS over the R-DLPFC (at 120% resting motor threshold, 1500 or 1200 pulses, respectively) concurrent with psychotherapy (Roelofs et al. 2020). In rTMS Dataset-3, 39 patients received only 1 Hz R-DLPFC stimulation and psychotherapy. All other parameters were the same as in rTMS Dataset-2.

DATASET-4 AND -5: ECT

ECT Dataset-4 comprises data from the Study on Neuroimaging predictors of Outcome in ECT Patients (SNOEP) which was approved by Rijnstate Hospital and the medisch-ethische toetsingscommissies (METC) Arnhem/Nijmegen. Patients who were referred for ECT treatment at Rijnstate hospital between August 2016 and June 2022 were included. All patients provided written consent for the original study from which the data has been used prior to study start. Since these data were collected as part of a clinical trajectory, participants did not receive financial compensation. Thirty-nine outpatients with MDD were treated with ECT, 19 of whom had complete EEG and outcome data. Fourteen received bifrontotemporal (BL) stimulation and 5 right unilateral (RUL; according to d'Elia (D'Elia 1974)) stimulation with stimulus dose relative to seizure threshold (SDRST; i.e., 6 times seizure threshold [ST] in RUL and 2.5 times ST in BL ECT) and using 0.5 ms pulse width (PW). Resting-state EEG data and HRSD-17-

score were collected prior to ECT and two weeks post-ECT-course.

ECT Dataset-5 comprised data of 60 patients who underwent ECT treatment at University Hospital Zurich between 2006 and 2015. This study was not reviewed by an independent ethics committee. All participants provided written consent for the original study from which the data has been used. Since these data were collected as part of clinical treatment, participants received no financial compensation. As part of clinical treatment, patients were treated with 6-12 sessions of bifrontal ECT (PW=0.5ms, SDRST=1.5 of ST) (Kirsten, Seifritz, and Olbrich 2019), with outcome analyzed by CGI.

STRATIFICATION BETWEEN BRAIN STIMULATION TECHNIQUES

Discovery analyses were conducted for brain stimulation techniques except for the 10 Hz rTMS prediction since the direction of effect was informed on previous findings. Since this finding has already been independently replicated (Roelofs et al. 2020), no blinded out-of-sample validation was conducted. Instead, in Dataset-2, remission was predicted in patients with an iAF in the synchronization range (iAF between 9.6-10.4) who had received 10 Hz rTMS.

For 1 Hz and ECT datasets, all possible directions of effect were tested, i.e., low decile score (1-5) excl. synchronization range, synchronization range, and high decile (6-10) excl. synchronization range. Potential findings were subsequently evaluated in blinded, out-of-sample validations in rTMS Dataset-3 and ECT Dataset-5.

Lastly, we conducted a simulation for stratification between brain stimulation interventions.

Since patients were not evenly distributed across the different subgroups (low decile, synchronization and high decile), we first determined the percentage of patients that can be expected to be stratified to each subgroup based on all our rTMS and ECT datasets. Subsequently, we used these percentages to calculate the weighted average of the PPVs that were previously determined for each treatment. The resulting PPV and nPPV were the expected remission rate and

normalized remission rate following stratification to rTMS and ECT with Brainmarker-I.

DATASET-6 AND -7: EXPLORATORY ANALYSES - PSYCHOTHERAPY, KETAMINE AND BUPROPION

Dataset-6 comprised patient data from three outpatient mental health care clinics (Synaeda Leeuwaarden Fonteinland, Synaeda Drachten, Synaeda Heerenveen), and was therefore not reviewed by an independent ethics committee. Each patient provided written informed consent for data use prior to EEG and treatment start. Since these data were collected as part of clinical treatment, participants received no financial compensation.

Approval for all three ketamine studies used in Dataset-7 was obtained from the Ethical committee of Prague Psychiatric Centre/National Institute of Mental Health, Czech Republic prior to patient enrollment. Outpatients were recruited for study participation at Prague Psychiatric Centre, Czech Republic between 2010 and 2022. All patients provided written informed consent. No financial compensation was offered (for more information on ethical approval, study protocol and participant inclusion, we refer the reader to the relevant trial registration and reference (Meyer et al. 2021)).

Exploratory analyses were performed in Dataset-6 for psychotherapy (Meijs et al. 2022), in Dataset-7 for ketamine treatment and in Dataset-1 for the subgroup of sertraline non-responders, switched to Bupropion (N= 54) in accordance with the previous analyses, however, without a guided hypothesis. Datasets are described in more detail in Supplementary Discussion S4.

Clinical trials registration: Establishing Moderators and Biosignatures of Antidepressant Response for Clinical Care for Depression (EMBARC). Identifier: NCT01407094. URL: http://clinicaltrials.gov/show/NCT01407094.

QEEG cordance and EEG connectivity changes after administration of subanesthetic ketamine doses in depressive disorder patients. URL: https://www.clinicaltrialsregister.eu/ctr-search/trial/2009-010625-39/CZ.

The Role of mTOR (Mammalian Target of Rapamycin) Signaling Pathway in the Antidepressive Effect of Ketamine in Patients with Depressive Disorder. URL: https://www.clinicaltrialsregister.eu/ctr-search/trial/2013-000952-17/CZ.

Clinical and neurobiological predictors of response to ketamine: towards personalized treatment of depression. URL: https://www.clinicaltrialsregister.eu/ctr-search/trial/2018-001539-39/CZ.

DATA AVAILABILITY

The TDBRAIN EEG data is freely available for download at https://brainclinics.com/resources/. The EMBARC dataset is available from the National Institute of Mental Health Data Archive (https://nda.nih.gov/edit_collection.html?id=2199). Other data are available from the corresponding author on reasonable request. Since these data were kindly shared with us by collaborators and due to the provided consent by participants of the respective studies, we are not at liberty to make these accessible in a repository.

CODE AVAILABILITY

The Python code used for processing the EEG and calculating the iAPF was custom-made for this study and is freely available for download at https://brainclinics.com/resources/.

ACKNOWLEDGEMENTS

We would like to thank Mark Koppenberg for the design of figures and tables and acknowledge the contributors to the TDBRAIN+ dataset: Kristi R. Griffiths, Roger DeBeus, Mayuresh S. Korgaonkar, Rien Breteler, Johan Fekkes, Ben Reitsma, Paul du Jour, Rosalinde van Ruth, Jeanine Jansen.
Some data presented in this work (Dataset-7) was supported by Czech Science Foundation (projects 21-32608S and 21-14727K).
We would like to thank the EMBARC National Coordinating Center at UT Southwestern Medical Center, Madhukar H. Trivedi, M.D., Coordinating PI, and the Data Center at Columbia University.
Data and/or research tools used in the preparation of this manuscript were obtained from the National Institute of Mental Health (NIMH) Data Archive (NDA). NDA is a collaborative informatics system created by the National Institutes of Health to provide a national resource to support and accelerate research in mental health. Dataset identifier(s): 10.15154/1528638. This manuscript reflects the views of the authors and may not reflect the opinions or views of the NIH or of the Submitters submitting original data to NDA.

DISCLOSURES

The authors declare the following competing interests:
ATS is Chief Scientific Advisor of PlatoScience and Alphasys, Founder of Neurowear Medical B.V., received equipment support from MagVenture, Deymed, and MagStim Company, and is Scientific Director of the International Clinical TMS Certification Course (www.tmscourse.eu). SO is cofounder of DeepPsy, a company for EEG-biomarker report systems.
Over the past 3 years, DAP has received consulting fees from Albright Stonebridge Group, Boehringer Ingelheim, Compass Pathways, Engrail Therapeutics, Neumora Therapeutics (formerly BlackThorn Therapeutics), Neurocrine Biosciences, Neuroscience Software, Otsuka, Sage Therapeutics, Sunovion, and Takeda; he has received honoraria from the Psychonomic Society and the American Psychological Association (for editorial work) and Alkermes; he has received research funding from the Bird Foundation, Brain and Behavior Research Foundation, the Dana Foundation, Millennium Pharmaceu-

ticals, National Institute of Mental Health (NIMH), and Wellcome Leap; he has received stock options from Compass Pathways, Engrail Therapeutics, Neumora Therapeutics, and Neuroscience Software. No funding from these entities was used to support the current work, and all views expressed are solely those of the authors.
MB declares to have shares in "Psyon s.r.o" and is involved in Compass Pathways trials with psilocybin and MAPS clinical trial with MDMA outside the submitted work. He has founded "PSYRES - Psychedelic Research Foundation" and has shares in "Společnost pro podporu neurovědního výzkumu s.r.o".
MA holds equity/stock in neurocare and Sama Therapeutics, serves as consultant to Synaeda and Sama Therapeutics, and is named inventor on neurocare owned patent, but receives no royalties; Research Institute Brainclinics received research and consultancy support from neurocare and equipment support from neuroconn and Deymed.
Brainmarker-I as described in this manuscript is offered by the non-profit mental health clinic Synaeda PMC (Leeuwarden, The Netherlands), where it is also used in practice.
Authors HTSV, SS, RR, AP, NV, JAW, MB, IO, BR, JF, HD report no conflicts of interest.

SUPPLEMENTARY MATERIAL

Week	8	8	16	16	16
Treatment	Placebo	SER	Placebo	SER	SER-BUP
Remission rate	31%	42%	84%	72%	35%

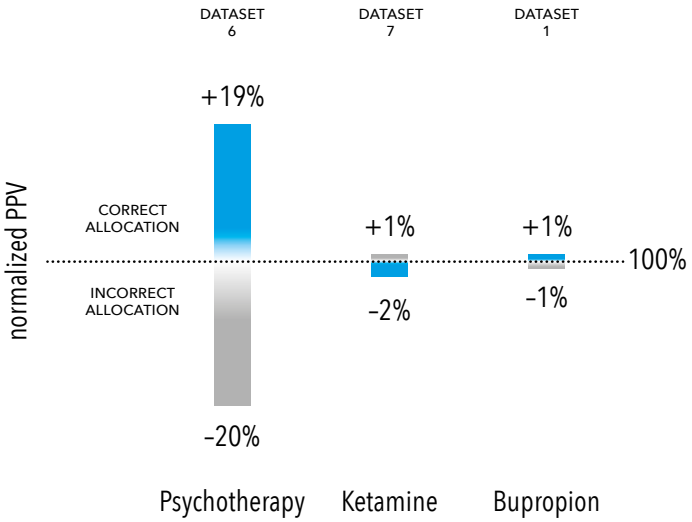
Supplementary Table S4.1.1. Full sample remission rate per dataset. *Remission rate Dataset-1 (EMBARC) for all treatment arms.*
SE*R= Sertraline,* BUP= *Bupropion*

Dataset	2	2	3	4	5	6	7
Sample size (N)	74	113	39	41	51	156	81
Treatment	10 Hz rTMS	1 Hz rTMS	1 Hz rTMS	ECT	ECT	Psychotherapy	Ketamine
Remission rate	60%	53%	43%	26%	61%	30%	26%

Supplementary Table S4.1.2. Remission rates remaining datasets.
rTMS = *repetitive transcranial magnetic stimulation,* **ECT** = *electroconvulsive therapy*

Dataset	Treatment	Low decile subgroup (PPV/nPPV)	SYNC subgroup (PPV/nPPV)	High decile subgroup (PPV/nPPV)
2	10 Hz rTMS	52 (-14%)	77 (+29%)	53 (-12%)
2	1 Hz rTMS	45 (-14%)	58 (+9%)	60 (+14%)
4	ECT	36 (+38%)	0 (-100%)	17 (-37%)

Supplementary Table S4.2 full results discovery analyses in Dataset-2 and -4.*Discovery analyses were conducted for rTMS and ECT, calculating remission rates and normalized remission rates (PPV/nPPV) in all different biomarker subgroups, i.e. the low decile (decile 1-5) subgroup, the SYNC subgroup (9.6-10.4 Hz), and the high decile (decile 6-10) subgroup.*
rTMS = *repetitive transcranial magnetic stimulation,* **ECT** = *electroconvulsive therapy,* **PPV** = *positive predictive value,* **nPPV** = *normalized positive predictive value*



Supplementary Figure S4.1. *Results of the Exploratory Analyses.*
Normalized PPV for low decile subgroup (blue) vs high decile subgroup (grey) for psychotherapy (any), ketamine and bupropion treatment.
Only part of Dataset-1 went into the analysis for bupropion, namely patients who did not respond to 8 weeks of sertraline administration (with response defined as ≥50% reduction in the 17-item Hamilton Rating Scale for Depression (HRSD-17) scores) and were subsequently switched to bupropion treatment. PPV= *positive predictive value*

SUPPLEMENTARY DISCUSSION S4.1. SIDE EFFECTS ECT

At the moment of the analysis, only side effect data from the ECT Dataset-5 were available.

The capacity of Brainmarker-I to predict which patient will have no side effects after ECT treatment was tested in the same way as treatment outcome was tested. Positive predictive values (PPVs) were calculated for patients falling into the low decile range compared to patients falling into the high decile range. The occurrence of side effects was high, with only 36% of patients experiencing no side effects. The PPV in the low decile range was 44% (nPPV= +24%) compared to a PPV of 25% (nPPV= -30%) in the high decile subgroup. This means that people who are more likely to respond to ECT treatment, ac-

cording to Brainmarker-I, are also more likely to experience no side effects. Although very promising and of high clinical relevance, this finding requires replication in an independent dataset.

Dataset	Treatment	Subgroup Low decile	Subgroup SYNC	Subgroup High decile
1	Sertraline (8 weeks)	41.6%	31.5%	27%
1	Placebo (8 weeks)	38.3%	40.2%	21.5%
1	Sertraline (16 weeks)	42.6%	36.2%	21.3%
1	Placebo (16 weeks)	37.5%	37.5%	25%
2	10 Hz rTMS	42.6%	32.4%	25%
2	1 Hz rTMS	47.3%	30.4%	22.3%
3	1 Hz rTMS	40.5%	37.8%	21.6%
4	ECT	57.9%	10.5%	31.6%
5	ECT	54.3%	28.3%	17.4%

Supplementary Table S4.3. *Distribution of biomarker subgroups across main datasets.*
rTMS = *repetitive transcranial magnetic stimulation,* ECT = *electroconvulsive therapy*

SUPPLEMENTARY DISCUSSION S4.2. ABSENCE OF ALPHA OSCILLATIONS

We previously developed a method of determining low-voltage alpha (LVA) in data based on the linked-mastoid (LM) reference montage (Voetterl et al. 2022). The individual alpha peak frequency (iAF) calculated with the LM montage led to a slightly higher correlation with age for children (6-18yrs.) as compared to iAF based on the average reference (AR) montage (ρ= 0.12 vs ρ=0.11). However, for adults the opposite was true (LE: ρ=0.05 vs AR: ρ= 0.06), i.e., the AR reflected most of the variance in iAF related to age, hence in the current study which focuses on an adult population, this reference montage was a priori selected. In addition, for the 10-Hz TMS SYNC findings, the necessity of using an AR has already been demonstrated before (Corlier et al. 2019).

We therefore recalculated the LVA threshold based on the AR, in the same way as the threshold was determined for an LM montage (Voetterl et al. 2022).

The average spectral power value in the alpha range (7-13 Hz) was determined with the Fast Fourier Transform (FFT) in all 3 electrodes (Fz, Pz, Oz) for each individual and subsequently log-transformed. The z-score of -1.96 in Pz, i.e., the lowest 5% of the log-transformed spectral alpha power, marked the threshold for insufficient alpha power. This value was validated by visually inspecting 1% of the recordings above and below this threshold for alpha oscillations. When the spectral power value in an electrode fell below this threshold the data of this electrode was discarded.

The resulting log amplitude threshold was 12.3. Applying this threshold led to 5% of data being discarded at 10-20 EEG system electrode Fz.

In light of the fact that different amplifiers lead to different voltage gains of the signal, the LVA threshold needed to be adapted according to the amplifier used. As several datasets (EMBARC Dataset-1, ECT Dataset-5 and ketamine Dataset-7) recorded data with a different amplifier from the one we used, we determined new LVA thresholds for these datasets in line with the previous calculation.

SUPPLEMENTARY DISCUSSION S4.3. BIOMARKER DISCOVERY PHASE

In Voetterl et al (Voetterl et al. 2022), a total of 108 algorithm permutations were tested with 1) condition (eyes closed (EC), eyes open (EO) or EC-EO Difference (Diff)), 2) choice of segmentation length (2-7s), 3) montage (linked-mastoids (LM), average reference (AR)), and 4) topographical location (channel location Fz, Pz, Oz).

A decision on segment length was made based on 1) the strength of the correlation and 2) the number of subjects retained for each segment length and averaged across reference (LM and AR), and conditions (EC, EO, Diff) for all 3 electrode locations (Fz, Pz, Oz) separately. The choice of reference montage was based on the highest iAF age correlation for the age range of interest, i.e., subjects above the age of 18.

In a post hoc analysis, comparing reference montages between children (age 6-18 yrs.) and adults, the iAF determined with a LM reference montage led to a slightly higher correlation with age for children (6-18yrs.) compared to the iAF determined with an AR montage (ρ= 0.12 vs ρ=0.11), while this was reversed for adults (ρ=0.05 vs ρ= 0.06), i.e., the AR reflected the adult data better.

Correlations and data retention averaged across all segment lengths was highest in the EC condition with ρ between 0.28 (Fz) and 0.34 (Oz) (except for Oz where the correlation was higher for EO but data retention was much lower). Therefore, EC was chosen as the condition for subsequent analyses. Electrodes Pz and Oz showed the highest correlation (Fz second highest) with age (Oz: ρ of 0.34) for a segment length of 5s (with condition EC and averaged across reference montages) while data retention was highest for 5s segment-length in all 3 electrode sites (97%). The strongest overall correlation between age and iAF for both reference montages for 5s segment length and condition EC was found in electrode Oz. Since this correlation was slightly higher in LM than in AR (ρ= 0.34 vs ρ= 0.33), and data retention was the same (97% across all electrodes) for both LM and AR, LM was chosen as primary reference montage. All 3 electrode locations were examined in further prospective biomarker testing.

GraphPad Prism (GraphPad Software, La Jolla California USA, www.graphpad.com) was employed in the full TD-BRAIN+ data to find the mathematical model that most closely approximates the brain-maturation effect. In line with previous evidence (van Dinteren et al. 2014), a log Gaussian fit determined separately for males and females and for each electrode location most appropriately followed the data (males: r2 = 11.9% (Oz); females: r2= 12.6% (Oz)) and continuously outperformed the linear model (H0; $p < 1x10^{-15}$). Normalized iAF values for each individual (divergence values), derived by subtraction of the model-predicted iAF from the real iAF, scattered around 0.

In order to validate the use of a clinical instead of a normative dataset, the full curve fitting procedure in GraphPad prism, specified above, was repeated in a normative dataset (Gerrits et al. 2019). Subsequently, in a comparison of fit both the normative and the clinical curve fit were applied to both the normative and clinical data separately and the fit was compared. Comparing the curve fit of the clinical TD-BRAIN+ dataset with the curve fit specific to a normative dataset (Gerrits et al. 2019) in GraphPad prism, indicated that the parameters of the clinical dataset generalized significantly better (p=0.03) to the normative data than the other way around (p=.21), suggesting that the clinical data is better capable of capturing and explaining variance.

STATISTICS BIOMARKER DISCOVERY

First, Spearman correlations between the various iAFs resulting from different EEG processing combinations and age in subjects below 18 years (N=1671) were calculated. To determine standardized iAF values independent of age, we derived non-linear regression models based on the full TD-BRAIN+ dataset that most closely fit the given data for each electrode (Fz, Pz, Oz). Different mathematical models following the developmental trajectory of the iAF (such as a Log gaussian model, in line with (van Dinteren et al. 2014)) were contrasted against a linear model (null hypothesis) and individually adjusted for females and males and for each site (channel). Divergence values represent where the individual's iAF lies compared to the mean iAF at that age and sex. These were calculated from the resulting models by sub-

tracting the model-derived average iAF for each subject's age from the person's actual iAF. Correlations between divergence values and age were conducted to confirm that the age effect had been eliminated from the data. The resulting divergence values were ranked from low to high and divided into 10 equal-sized bins (deciles) to improve interpretability by clinicians.

Curve fitting models were developed in GraphPad Prism version 8.4.0 for MacOS. Spearman correlations were conducted with Python modules scipy, and numpy.

All other statistical analyses were performed in IBM SPSS Statistics for Macintosh, Version 27.0.

SUPPLEMENTARY DISCUSSION S4.4. METHODS AND DISCUSSION OF EXPLORATORY ANALYSES

For Dataset-6, 175 patients were treated with psychotherapy (94 cognitive-behavioral therapy, 81 other therapy) in a clinical setting. Outcome was measured post-treatment with the BDI-II.

For Dataset-7, 81 patients were recruited for three different trials, a double-blind, randomized placebo-controlled trial, a single-blind, placebo-controlled trial, and an open-label trial. Ethical approval for all three studies was obtained from the Ethical committee of Prague Psychiatric Centre/National Institute of Mental Health, Czech Republic prior to patient enrollment (EudraCT Number: 2009-010625-39, 2013-000952-17 and 2018-001539-39).

All patients received ketamine via an intravenous catheter (0.54 mg/kg within 30 minutes) and outcome was tested with the MADRS 24hrs, 72hrs, and 7 days post-infusion.

In the exploratory analyses, only for psychotherapy an effect was found, with a higher remission rate in the low decile subgroup (nPPV = +11%) (see Supplementary Figure S4.1), while there were no effects for bupropion or ketamine treatment. This direction of effect for psychotherapy is in line with our previous finding of increased likelihood of remission to neurofeedback for boys with ADHD who fall

in the low decile range, since the therapy-like aspects of multimodal-neurofeedback such as clinician-attention are considered to play a major role in its mechanism of effect. However, since no replication data were available for the exploratory analyses, results need to be interpreted with caution and require further replication.

SUPPLEMENTARY REFERENCES

Voetterl H, van Wingen G, Michelini G, Griffiths KR, Gordon E, DeBeus R, et al. (2022): Brainmarker-I differentially predicts remission to various attention-deficit/hyperactivity disorder treatments: a blinded discovery, transfer and validation study. *Biological Psychiatry Cognitive Neurosci Neuroimaging*. https://doi.org/10.1016/j.bpsc.2022.02.007

Corlier J, Carpenter LL, Wilson AC, Tirrell E, Gobin AP, Kavanaugh B, Leuchter AF (2019): The relationship between individual alpha peak frequency and clinical outcome with repetitive Transcranial Magnetic Stimulation (rTMS) treatment of Major Depressive Disorder (MDD). *Brain Stimul.* https://doi.org/10.1016/j.brs.2019.07.018

van Dinteren R, Arns M, Jongsma MLA, Kessels RPC (2014): Combined frontal and parietal P300 amplitudes indicate compensated cognitive processing across the lifespan. *Front Aging Neurosci* 6: 294.

Gerrits B, Vollebregt MA, Olbrich S, van Dijk H, Palmer D, Gordon E, et al. (2019): Probing the "Default Network Interference Hypothesis" With EEG: An RDoC Approach Focused on Attention. *Clin EEG Neurosci* 50: 404–412.

5

DOES 18 HZ DEEP TMS BENEFIT A DIFFERENT SUBGROUP OF DEPRESSED PATIENTS RELATIVE TO 10 HZ RTMS?

THE ROLE OF THE INDIVIDUAL ALPHA FREQUENCY.

Based on Does 18 Hz deep TMS benefit a different subgroup of depressed patients relative to 10 Hz rTMS? The role of the individual alpha frequency.
Voetterl, H.T.S., Alyagon, U., Middleton. V.J., Downar, J., Zangen, A., Sack, A.T., van Dijk, H., Halloran, A., Donachie, N. & Arns, M. (under review at European Neuropsychopharmacology)

ABSTRACT

Both 10 Hz repetitive transcranial magnetic stimulation (rTMS) as well as 18 Hz deep TMS (dTMS) constitute effective, FDA-approved TMS treatment protocols for depression. However, not all patients experience sufficient symptom relief after either of these protocols. Biomarker-guided treatment stratification could aid in personalizing treatment and thereby enhancing improvement. An individual alpha frequency (iAF)-based EEG-biomarker, Brainmarker-I, can differentially stratify patients to depression treatments. For instance, an iAF close to 10 Hz was associated with better improvement to 10 Hz rTMS, possibly reflecting entrainment of endogenous oscillations to the stimulation frequency.

Accordingly, we examined whether 18 Hz dTMS would result in better improvement in individuals whose iAF lies around 9 Hz, a harmonic frequency of 18 Hz.

Curve fitting and regression analyses were conducted to assess the relation between iAF and improvement. For treatment stratification purposes, correlations with iAF-distance to 10 Hz compared 18 Hz dTMS (N=114) to 10 Hz rTMS (N=72).

We found a robust quadratic effect, indicating that patients with an iAF around 9 Hz exhibited least symptom improvement (r^2=.126, $p<.001$).

Improvement correlated positively with iAF-distance to 10 Hz (p=.003). A secondary analysis in 20 Hz figure-of-eight data confirmed this direction. A significant interaction of iAF-distance and stimulation frequency between 10 and 18 Hz datasets emerged (p=.026).
These results question entrainment of endogenous oscillations by their harmonic frequency for 18 Hz, and suggest that 10 Hz and 18 Hz TMS target different subgroups of depression patients. This study adds to iAF stratification, augmenting Brainmarker-I with alternative TMS protocols (18 Hz/20 Hz) for patients with a slower iAF, thereby broadening clinical applicability and relevance of the biomarker.

DOES 18 HZ DEEP TMS BENEFIT A DIFFERENT SUBGROUP OF DEPRESSED PATIENTS RELATIVE TO 10 HZ RTMS? THE ROLE OF THE INDIVIDUAL ALPHA FREQUENCY.

INTRODUCTION

Major depressive disorder (MDD) is one of the most prevalent disorders worldwide and carries a tremendous disease burden (Malhi & Mann, 2018; World Health Organization, 2008). Although many effective pharmacological treatments are available, many patients do not experience symptom relief despite multiple treatment attempts (Arns et al., 2022; Drysdale et al., 2017). Those patients are considered to suffer from a difficult-to-treat depression and often move on to other treatment options such as repetitive transcranial magnetic stimulation (rTMS). Treatment with rTMS is well-researched and shows good effectiveness at the group level, but a full treatment course is costly and time-consuming, and about 40% of patients still don't experience sufficient symptom relief post-treatment (Arns et al., 2022).
One way to speed up recovery and foster clinical improvement is predicting outcome by means of so-called biomarkers, for instance patterns measured with resting-state electroencephalography (EEG). Many biomarkers have been suggested for rTMS in depression (Drysdale et al., 2017; Elbau et al., 2023; Hunter et al., 2017; Wu et al., 2020), however to the best of our knowledge, none are widely used in clinical practice yet.

One approach to personalize treatment that might be particularly useful is treatment stratification, which describes the concept of identifying the best possible treatment for an individual out of multiple established treatment options using one or multiple biomarkers (Arns et al., 2022). Our group recently introduced a stratification biomarker (Brainmarker-I), that is able to differentially predict outcome to different treatments for attention-deficit hyperactivity disorder (ADHD) (Voetterl et al., 2022) and MDD (Voetterl, 2023). Brainmarker-I is based on the individual alpha frequency (iAF) in frontal brain areas, a pattern observed in the resting-state EEG that has been investigated extensively for treatment response (Arns, Drinkenburg, Fitzgerald, et al., 2012; Arns et al., 2017, 2018; Corlier et al., 2019; Krepel et al., 2020; Roelofs et al., 2020). The iAF is the modal frequency of an individual's oscillation in the alpha range (7-13 Hz) and is considered highly hereditary (Van Beijsterveldt & van Baal, 2002). Its relationship with symptom improvement was initially believed to be simple, with a slower iAF predicting reduced efficacy of depression treatments in general (Arns, Drinkenburg, Fitzgerald, et al., 2012; Ulrich et al., 1984). Recent evidence however, suggested a more multifaceted relation. Instead of a positive linear relationship, as suggested before, the distance of the iAF to 10 Hz was found to negatively correlate with symptom improvement after 10 Hz rTMS over the left dorsolateral prefrontal cortex (DLPFC) (Corlier et al., 2019). Corlier et al. hypothesized that this effect might reflect entrainment of intrinsic brain oscillations to the frequency of stimulation. This finding was independently replicated (Roelofs et al., 2020), while no effect of iAF on 1 Hz treatment outcome was found. The Brainmarker-I study, however, showed that patients with a higher iAF that was not close to 10 Hz, indicated by a high biomarker decile score, had a better chance to achieve remission after 1 Hz right-sided DLPFC stimulation (Voetterl, 2023). This was successfully replicated in an independent sample using a blinded out-of-sample validation. The subgroup of patients with a slow iAF (indicated by low decile scores) showed low remission rates for both 10 Hz and 1 Hz treatment but was found to respond best to ECT treatment which again was successfully replicated in a blinded out-of-sample analysis.

H-coils are a family of electromagnetic coils enabling deeper and broader stimulation volumes, relative to traditional figure-of-eight

rTMS coils (Zibman et al., 2021), and hence were termed Deep TMS coils (dTMS). Multiple H-coils gained FDA-approval for the treatment of different psychiatric disorders, including the H1-coil for the treatment of MDD using 18 Hz stimulation of bilateral prefrontal cortices with highest electric field induced over the left DLPFC (Tendler et al., 2016). More recently, 18 Hz rTMS using the H7 coil, which induces the strongest electrical field over the medial prefrontal cortex, has been shown to be equally effective for MDD treatment (Zangen et al., 2023). This was advocated by research implicating the dorsomedial prefrontal cortex as alternative appropriate target to treat depression (Downar et al., 2014; Downar & Daskalakis, 2013).

EEG-power based biomarkers for dTMS have been suggested recently (Alyagon et al., 2020; Zangen et al., 2023), however, the relationship between alpha peak frequency and symptom improvement after stimulation with 18 Hz dTMS has, to the best of our knowledge, never been tested before. Here, we examined whether iAF has predictive potential for 18 Hz dTMS delivered with either the H1 or the H7 coil. Although 18 Hz does not fall into the alpha frequency range, it represents the first harmonic of the 9 Hz frequency. Zmeykina et al. showed that phase entrainment not only happened at the individual alpha frequency at which TMS pulses were delivered, but also at the first harmonic frequency in the beta range (Zmeykina et al., 2020). This suggests that stimulating the endogenous frequency in the alpha range with TMS pulses at the harmonic frequency (here 18 Hz) could lead to a similar entrainment effect (Figure 5.1a). On the other hand, Zrenner et al. demonstrated that only the negative peak of the μ rhythm - the most prominent oscillation in the sensorimotor cortex (8-12Hz) - represents a high excitability state (Figure 5.1b) (Zrenner et al., 2018). Since the 18 Hz harmonic frequency would stimulate the same oscillatory cycle of the dominant alpha rhythm at opposite phases (see Figure 5.1b lower), and thereby affect at least one state of low excitability, it could push the oscillator out of balance.

In line with the first hypothesis that TMS might entrain the endogenous underlying frequency of the same rhythm (Corlier et al., 2019; Roelofs et al., 2020), we theorized that an iAF around 9 Hz would correlate with better response to 18 Hz TMS treatment. To examine whether any potential effects are more related to the specific design of the H-coils or to the higher stimulation frequency, the same direc-

tion of effect was tested in a secondary analysis in an independently acquired dataset that delivered higher frequency (20 Hz) stimulation with a traditional figure-of-eight coil.

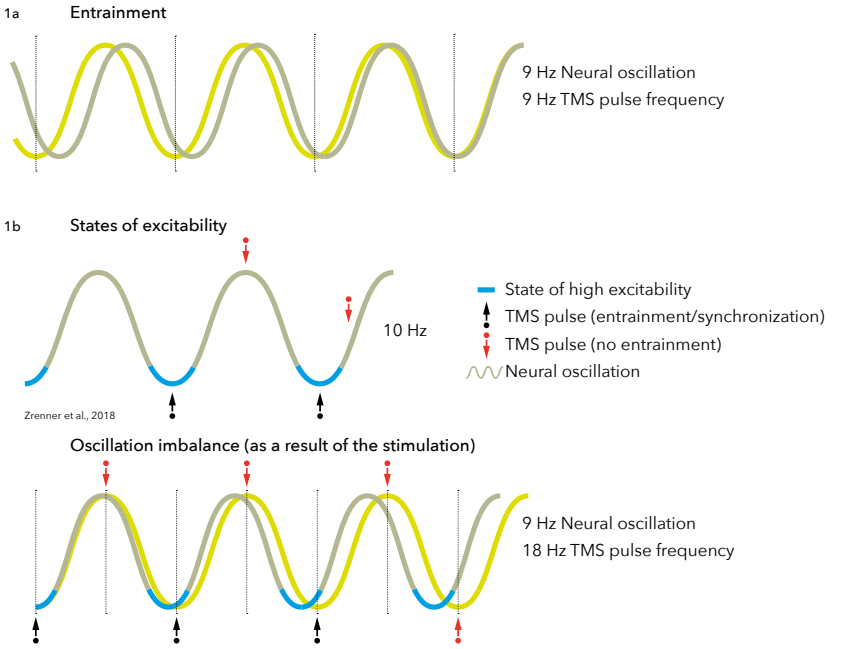


Figure 5.1. Entrainment hypotheses
1a. *The first hypothesis states that phase entrainment (grey line) might happen when an endogenous oscillation (lime green line) is stimulated by TMS pulses of the same frequency. It has been shown that this entrainment occurs not only at the endogenous frequency but also at the first harmonic frequency in the beta range (Zmeykina et al., 2020), suggesting that stimulation with a harmonic frequency (here:18 Hz) might also result in phase entrainment.*
1b (upper). *However, only the negative peak of the oscillation appears to represent a state of high excitability (Zrenner et al., 2018) while the positive peak and the random phase (oscillation slopes) are states of low excitability.*
1b (lower). *The second hypothesis therefore suggests that stimulating the endogenous oscillation (lime green line) with its harmonic frequency (e.g., 18 Hz) impacts the same oscillatory cycle twice at opposite phases (red and black arrows), thereby affecting at least one state of low excitability (red arrows). This might lead to an unbalancing of the oscillator (grey-blue line) instead of entrainment.*

MATERIALS AND METHODS

DATASET 1 – BRAINSWAY DTMS 18 HZ

In this double-blind, randomized, controlled trial, 114 patients with MDD received 24 sessions of 18 Hz rTMS with either the Brainsway H1 (mainly left prefrontal stimulation) or the H7 coil (mainly medial prefrontal stimulation), and had sufficient clean pre-treatment EEG data (Zangen et al., 2023). For the electric field maps of both coils, see Tendler et al. (Tendler et al., 2016).
Primary clinical outcome was the change in Hamilton Depression Rating Scale 21 score (HDRS-21) following 6 weeks of treatment, to quantify symptom improvement and treatment remission. In line with Voetterl et al. (Voetterl, 2023) primary outcome was remission, defined as <= 7, in accordance with Riedel et al. (Riedel et al., 2010).

DATASET 2 – TD-BRAIN 10 HZ RTMS

Data from Roelofs et al. (Roelofs et al., 2020) were used to contrast the 18 Hz dTMS results to the 10 Hz rTMS results. The exclusion criterion for insufficient alpha differed in this previous study, thus 73 MDD patients (instead of 59) were included, who received an average of 21 (at least 10) sessions of 10 Hz rTMS with a figure-of-eight coil. The primary outcome measure was remission, defined as <=12 on the Beck depression inventory II (BDI-II) (Riedel et al., 2010).

DATASET 3 – SALIENCE BILATERAL RTMS (20 HZ + 1 HZ)

96 patients (68 female, mean age: 39) with MDD had sufficiently clean pre-treatment EEG data and were treated with an average of 36 (at least 10) sessions of 1 Hz rTMS over right DLPFC (360 pulses, 120%MT) followed by 20 Hz rTMS over the left DLPFC (1200 pulses, 120%MT), using MagVenture equipment with a Cool B65 figure-of-eight coil.
Symptom improvement was measured with the self-rated 30-item Inventory of Depressive Symptomatology (IDS), with remission defined as a score <=14 (Rush et al., 2005).

DATA COLLECTION AND PREPROCESSING

For Dataset-1, 5 minutes of resting state EEG data (eyes closed) was collected using TMS-compatible 32 (8 sites) or 64-channel (one site) amplifiers (TMSi Ltd.), sampled at 2048 Hz (down-sampled offline to 500 Hz), with an online Cz reference and POz as ground. Impedance was kept <10 kΩ for all electrodes. Data was matched as closely as possible to the data used for previous Brainmarker-I predictions.
Matching channels to the previously used channel setup resulted in 23 channels for 111 participants (Fp1, Fp2, F7, F3, Fz, F4, F8, FC3, FC4, T7, C3, C4, T8, CP3, CP4, P7, P3, Pz, P4, P8, O1, Oz, O2) and 25 channels for 9 participants (Fp1, Fp2, F7, F3, Fz, F4, F8, FC3, FCz, FC4, T7, C3, C4, T8, CP3, CPz, CP4, P7, P3, Pz, P4, P8, O1, Oz, O2).
Due to low data quality in some recordings, a low-pass filter of 45Hz was applied to reduce excess artifacts from 50 and 60 Hz power line noise. Artifact rejection was performed with an automated, custom Python package (Harris et al., 2020; Hunter, 2007; The Pandas development team, 2020; Virtanen et al., 2020), in accordance with previous deartifacting procedures (Arns, Bruder, et al., 2016; van Dijk et al., 2022) (full code available under www.brainclinics.com/resources). Data was re-referenced offline to an average reference in line with previous Brainmarker-I predictions in adults (Voetterl, 2023).

Data of Dataset-2 was recorded in 26 channels of the 10-20 electrode international system (Fp1, Fp2, F7, F3, Fz, F4, F8, FC3, FCz, FC4, T7, C3, Cz, C4, T8, CP3, CPz, CP4, P7, P3, Pz, P4, P8, O1, Oz, O2; Quikcap, NuAmps) with a ground at AFz and an online linked-mastoid reference. Two minutes of eyes-open (EO) and two minutes of eyes-closed (EC) resting-state data were recorded at a sampling frequency of 500 Hz. Horizontal and vertical eye-movements were measured and skin resistance was kept <10 kΩ for all electrodes. The same artifact rejection procedure was followed as in Dataset-1 (see above), and data was re-referenced to an average reference.

Dataset-3 was recorded using 21 channels (Fp1, Fp2, F7, F3, Fz, F4, F8, T7, C3, Cz, C4, T8, P7, P3, Pz, P4, P8, O1, O2, M1, M2; Brainview EEG-21 Pro) with a ground at AFz and a CPz reference. Five minutes of EO and five minutes of EC resting-state data were recorded. Skin resistance was <10 kΩ for all electrodes. Sampling frequency was 500 Hz. A high

pass filter of 0.5 Hz was applied as well as a low pass filter of 60 Hz. Artifact rejection was conducted in Brain Vision Analyzer version 2.2.0 (Brain Products GmbH, Gilching, Germany) matching the preprocessing steps described above.

BIOMARKER DEVELOPMENT

Brainmarker-I was calculated in line with the previously reported EEG biomarker for ADHD and MDD (for a detailed description of the biomarker development see Voetterl (Voetterl et al., 2022)).
In short, the biomarker was developed in the large heterogeneous, clinical TDBRAIN+ dataset (N=4249) of which a part (n = 1274) is freely available for download at https://brainclinics.com/resources/ (van Dijk et al., 2022). The iAF was determined as the highest peak in the alpha range (7-13 Hz) determined by Fast Fourier Transform in the cleaned resting-state eyes-closed data at the 10-20 EEG system electrode sites F3 and Fz. Participants with insufficient (frontal) alpha oscillations were excluded (~5%) and iAFs were age- and sex-normalized. Resulting values were divided into decile scores (Fz) and a synchronization indicator (SYNC) was introduced, marking an iAF between 9.6-10.4 Hz at the F3 location for stratification to 10 Hz rTMS. The automated algorithm described in Voetterl et al. (Voetterl et al., 2022) was used to calculate iAF, synchronization and decile scores for individuals of all datasets.

STATISTICS

Curve fitting (including Loess curve fits) was used to examine the relationship between percent improvement on the HDRS-21 and iAF at electrode site Fz, in line with the previous Brainmarker-I analyses that were calculated based on this location and the broad electric fields of both coils, overlapping this area. An extra sum-of-squares F test compared one shared curve fit for the H1 and H7 datasets with the fit of individual curves for each dataset to test whether a potential effect significantly differed between the two coil datasets.
Accordingly, regression analyses were conducted for both coils together or separately.

In line with Roelofs et al. (Roelofs et al., 2020), who found a negative correlation between the distance of the endogenous iAF to 10 Hz and symptom improvement following 10 Hz rTMS treatment, Spearman correlation analyses were performed to examine whether there was an effect between symptom improvement and iAF distance to 10 Hz. Significance levels reported represent two-tailed significance testing for 18 Hz data and one-tailed testing for 10 Hz data, since this effect has been shown and replicated before.
To assess whether there was a difference in effect of distance to 10 Hz and symptom improvement between Dataset-1 and Dataset-2, a multiple regression analysis was conducted with the interaction term Dataset*10 Hz-distance added.
Positive predictive values (PPVs) that denoted the remission rate in each biomarker-classified subsample of patients, were calculated for synchronized (around 10 Hz) and non-synchronized groups at electrode location F3, since this is the location that was used for synchronized calculations before (Corlier et al., 2019; Roelofs et al., 2020; Voetterl, 2023). PPVs for Brainmarker-1 subgroups (i.e., synchronized, high decile and low decile subgroups) were based on Fz (in accordance with Voetterl et al. (Voetterl, 2023)). PPVs were normalized (nPPV) to make remission rates comparable across datasets in line with Voetterl et al. (Voetterl, 2023; Voetterl et al., 2022), using the formula (m/w-1)*100 (with m=PPV and w=observed sample remission rate).

To test whether a potential effect of 18 Hz treatment can be attributed to (higher) stimulation frequency, we repeated the same PPV calculations in a secondary analysis in Dataset-3 which delivered 20 Hz (+1 Hz) treatment with a traditional figure-of-eight coil.
To control for possible confounding factors, we conducted a Kruskal-Wallis test assessing whether there were differences between recording sites, with subsequent leave-one-out analyses for each research site, repeating the previous regression analyses 8 times (for 8 research sites). Possible differences in age between subgroups were tested with one-way ANOVAs for each dataset. To examine whether treatment resulted in a change in iAF, a repeated-measures ANOVA was performed to assess pre-/post-treatment iAF changes.
Biomarker calculation was conducted in Python, using modules numpy (Harris et al., 2020), pandas (The Pandas development team, 2020), and

scipy (Virtanen et al., 2020). GraphPad prism version 9.5.1 for MacOS (GraphPad Software, La Jolla California USA, www.graphpad.com) was used for curve fitting analyses. All other statistical analyses were performed in IBM SPSS Statistics for Macintosh, Version 27.0.

RESULTS

Since Dataset-1 is the main focus of the paper, all analyses were conducted on these data, unless indicated otherwise.
Details of all datasets are summarized in Table 5.1.

CURVE FITTING REVEALS QUADRATIC ASSOCIATION

Due to insufficient EEG data quality, data of 7 patients was discarded. This resulted in 113 resting-state EEG recordings being examined.
Results of the regression analysis are visualized in Figure 5.2.
Contrary to our hypothesis, Loess curve fits revealed a u-shaped effect, with a trough between 9 and 10 Hz, indicating least improvement for people with an iAF in this range. A quadratic fit most adequately followed the data, and this effect was present for both the H1 and the H7 coil, separately. An extra sum-of-squares F test showed that this effect did not significantly differ between the two coils and that one curve adequately fit both H1 and H7 datasets ($F(3,102)=0.1718$, $p=.9152$). Thus, all following analyses were conducted on the data from both coils combined (see Supplementary Figures S5.1. for quadratic effect in both coils, separately).

Table 5.1. Basic information for all datasets. *(next page)*
DBPC = *Double-blind placebo-controlled trial;* dTMS = *deep repetitive transcranial magnetic stimulation,* rTMS= *repetitive transcranial magnetic stimulation,* lDLPFC= *left dorsolateral prefrontal cortex,* rDLPFC= *right dorsolateral prefrontal cortex,* HDRS = *Hamilton Depression Rating Scale,* BDI= *Beck depression Inventory,* IDS = *Inventory of Depressive Symptomatology*

Full Datasets	Dataset-1	Dataset-2	Dataset-3
Study type	DBPC	Naturalistic	Naturalistic
Study site	Multicenter	Multicenter	Single site
Treatment protocol	18 Hz dTMS	10 Hz lDLPFC	Bilateral (1Hz rDLPFC + 20 Hz lDLPFC
TMS coil	H1/H7 dTMS coil	Figure-of-eight coil	Figure-of-eight coil
Sample size (N)	114	73	96
Females (%)	67 (59)	35 (48)	68 (71)
Mean Age (SD), years	45.8 (12.8)	41.1 (11.8)	39.1 (15.9)
Age range, years	22-69	19-65	14-79
Baseline severity	23.6 (HDRS)	30.4 (BDI)	37 (IDS)

QUADRATIC REGRESSION

A quadratic regression for both coils combined showed that there was a significant quadratic effect between iAF at Fz and percent improvement on the HDRS-21 ($F_{(2,105)}$=7.560, p<.001, r^2 = .126; Figure 5.2), with a trough around 9 Hz.

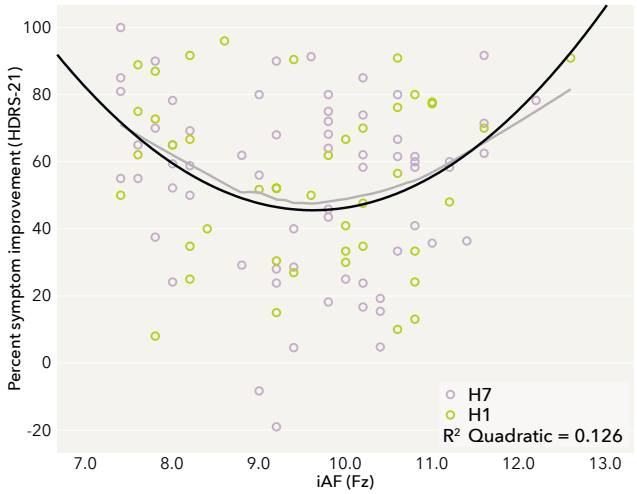


Figure 5.2. Quadratic effect between iAF and percent improvement
Quadratic curve fit (black) and Loess fit (dark grey) for iAF and percent symptom improvement for H1 and H7 coils combined.

RELATION BETWEEN SYMPTOM IMPROVEMENT AND DISTANCE TO 10 HZ

Due to the regression results revealing potential for differential stratification of patients to 18 Hz and 10 Hz stimulation, in line with Voetterl et al. (Voetterl, 2023), we focused the remaining analyses on the 10 Hz synchronization range (9.6-10.4 Hz) instead of 9 Hz.
In Dataset-1, we found a significant positive correlation between percent symptom improvement on the HDRS-21 and iAF distance to 10 Hz (rho=.269, p=.005). In line with previous findings, that correlation was negative for Dataset-2 (rho=-.208, p=.048).
To verify this stratification potential, we conducted a multiple regression analysis including an interaction of dataset (18 Hz Dataset-1 vs 10 Hz Dataset-2) on symptom improvement and iAF-distance to 10 Hz (Figure 5.3) which yielded a significant regression model (R^2=.053, $F_{(3,169)}$=3.168, p=.026). Larger distance to 10 Hz was related to less improvement to 10 Hz treatment but better improvement to 18 Hz treatment (Beta=-.246, p=.008). The model without the interaction term was not significant ($F_{(2,170)}$=1.143, p=.321).

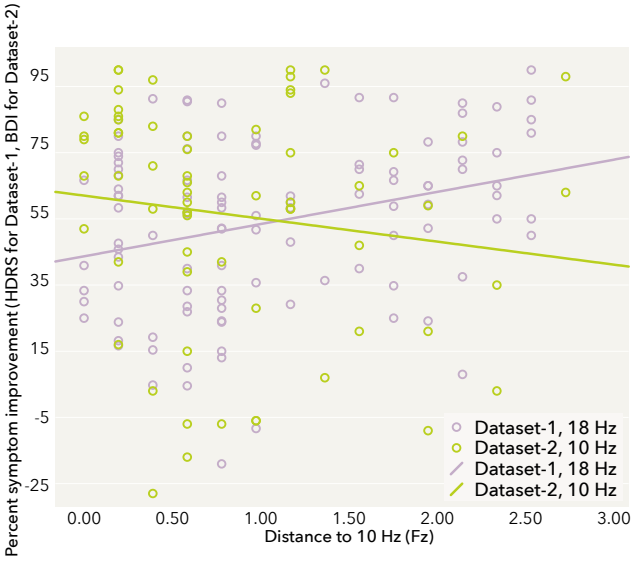


Figure 5.3. Visualization of interaction effect between 18 Hz and 10 Hz datasets
Interaction of Dataset (treatment protocol 18 Hz vs 10 Hz) on percent symptom improvement post-treatment and distance to 10 Hz. With increasing iAF-distance to 10 Hz, symptom improvement decreases for the 10 Hz protocol but increases for treatment with 18 Hz.
HDRS= *Hamilton Depression Rating Scale 21,* BDI= *Beck Depression Inventory*

POSITIVE PREDICTIVE VALUES

Full results can be found in Supplementary table and material S5.1; results for decile-based subgroups are visualized in Figure 5.4. For 18 Hz TMS, the strongest increment for decile-based subgroups was seen in the low decile subgroup (PPV=41%, nPPV=+23%).
To get an indication whether this quadratic effect for iAF and symptom improvement to 18 Hz dTMS treatment might be related to the high stimulation frequency of 18 Hz or whether it is a consequence of the dTMS coil design, we repeated the analysis in Dataset-3 (20 Hz rTMS with a figure-of-eight coil). Results were similar to Dataset-1, with the highest remission rate in the low decile subgroup (PPV=39%, nPPV=+23%).

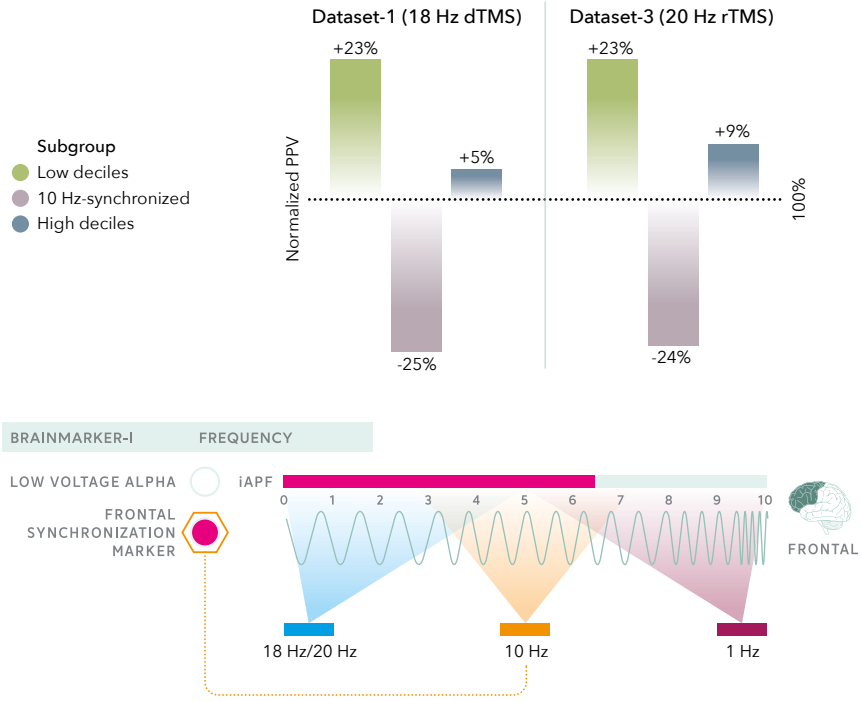


Figure 5.4. Subgroup analyses and Stratification recommendation
Results of the PPV/nPPV analyses of Dataset-1 and -3. The low decile subgroup presented with highest remission rates compared to the 10 Hz-synchronized and high decile subgroups. Taken together with the previous findings of Brainmarker-I, this suggests stratification of the low decile subgroup to 18 or 20 Hz TMS, the 10 Hz-synchronized subgroup to 10 Hz and the high decile subgroup to 1 Hz TMS.

CONFOUNDING FACTORS – DIFFERENCES BETWEEN RECORDING SITES

To rule out effects of recording site in Dataset-1, multiple analyses were conducted to examine whether the findings hold across recording sites. Since baseline HDRS-21 and percent change on the HDRS-21 were not normally distributed for multiple sites and there was heterogeneity of variance, a Kruskal-Wallis test was conducted to test whether baseline HDRS-21, percent change on the HDRS-21, iAF at Fz and age differed significantly between sites. Results showed that of these, only baseline HDRS-21 differed significantly between sites ($\chi^2(7)= 23.286$, $p=.002$).

We, therefore, conducted leave-one-out analyses for each site, repeating the previous quadratic regressions but excluding one site at a time. All 8 quadratic regressions remained highly significant ($p<.001$ to $p=.006$).

There were no age differences between biomarker subgroups of all datasets ($p>.216$).

Post-treatment EEG data was only available for Dataset-1. No significant differences were found between pre- and post-treatment iAF (see Supplementary Material S5.2).

DISCUSSION

In a previous study by our lab, we found and optimized an EEG biomarker based on the iAF, capable of transdiagnostically predicting remission to different ADHD and MDD treatments (Voetterl, 2023). For MDD, Brainmarker-1 predicted higher likelihood of remission to ECT for patients with a low decile score, i.e., a slow iAF, and to 1 Hz rTMS targeting the right DLPFC for patients with high decile scores. Moreover, 10 Hz rTMS targeting the left DLPFC worked best for patients with an iAF close to the stimulation frequency of 10 Hz, suggesting an underlying synchronization mechanism. In the present manuscript we further assessed this synchronization hypothesis in a dataset that administered 18 Hz stimulation, by examining the association between clinical improvement and an iAF of 9 Hz, a harmonic of the stimulation frequency of 18 Hz.

Intriguingly, we found an effect opposite to the entrainment hypothesis for the 18 Hz dTMS Dataset-1, namely a quadratic curve with least symptom improvement around 9 Hz iAF, which is opposite to the finding for 10 Hz rTMS. This effect was highly significant with moderate effect sizes, was replicated for both H1 and H7 coils, and was stable across different research sites.

These results complement and critically extend the Brainmarker-I findings (Voetterl, 2023), providing a less invasive alternative to the previously recommended ECT treatment for the low decile subgroup. Although the ECT finding is clinically meaningful, the prospect of more severe side effects often keeps patients from pursuing ECT treatment. Identifying rTMS protocols that work well for this patient group (i.e., supporting a three-way stratification) and lack these unwanted ECT side effects is, thus, very valuable.

Unlike the previous Brainmarker-I studies, the 18 Hz Dataset-1 used here administered treatment with H-coils which induce a larger electric field and stimulate larger cortical areas relative to standard figure-of-eight rTMS coils. This poses the question whether the effect found here is related to stimulation frequency or stimulation area. It is conceivable that this low decile subgroup requires a broader stimulation which is achievable by either ECT or deep TMS. However, the fact that the findings presented here hold for both H1 and H7 coil, and were validated in an independent 20 Hz dataset using a traditional figure-of-eight coil, suggests that coil type might be less relevant than frequency. For deep TMS, treatment with 10 Hz frequency has also been shown to be effective (Tendler et al., 2018). Despite the replicated quadratic effect of 10 Hz and iAF found with a figure-of-eight coil, it remains to be shown whether the same effect is present when delivering 10 Hz stimulation using an H-coil. The present results, thus, require replication of the 10 Hz finding in patients treated with an H-coil, as well as in patients who received higher frequency stimulation such as 18 or 20 Hz, delivered with a standard figure-of-eight coil. If successful, these findings would further advance and facilitate three-way treatment stratification, by stratifying patients with a 10 Hz-synchronized Brainmarker-I to 10 Hz rTMS, patients with higher iAF (high decile scores) to 1 Hz TMS (as shown previously(Voetterl, 2023)) and the low iAF subgroup to higher TMS frequencies (18 Hz or 20 Hz), independent of coil type (Figure 5.4).

Although the present study did not systematically investigate neural entrainment, the current findings might provide further evidence to support the existing literature. The hypothesis that rTMS can entrain oscillations, has been subject of extensive investigation, with good evidence for neural entrainment when stimulation is applied at the individual oscillatory rhythm (Leuchter et al., 2013; Thut, Schyns, et al., 2011; Thut, Veniero, et al., 2011). For instance, administering rhythmic (at the iAF) but not arhythmic or sham rTMS led to an immediate increase in phase alignment of the underlying oscillations with the TMS pulses (Zmeykina et al., 2020). That phase synchronization occurred both in posterior alpha rhythms and in their harmonic in the beta range, albeit to a lesser extent, seems to contradict our findings. This discrepancy could be explained by differing states of excitability (Zrenner et al., 2018). A slower stimulation frequency could entrain the endogenous harmonic frequency by stimulating at every second oscillatory cycle, however, a harmonic frequency of the dominant alpha rhythm (such as 18 Hz in the present study), would stimulate the same oscillatory cycle at opposite phases (see Figure 5.1b) potentially leading to an oscillator unbalance.
This is in line with findings that showed that the effectiveness of phase-locking (Thut, Veniero, et al., 2011; Zmeykina et al., 2020) and corticospinal excitability is phase-dependent (Schilberg et al., 2018, 2021; Zrenner et al., 2018) and that long-term-potentiation-like plasticity only occurred when the negative peak was triggered but not in the positive or random phase (slope of the oscillation) condition. This might explain a lack of effect when the oscillatory cycle is triggered at the wrong phase.

To date there are few studies that examined the relationship between synchronized TMS (sTMS) and clinical outcome. A large double-blind, randomized, sham-controlled trial and a randomized, sham-controlled pilot trial found significantly greater improvement for sTMS compared to sham (Jin & Phillips, 2014) and significantly less improvement in patients who received sTMS stimulation at the incorrect iAF (Leuchter et al., 2015).
These findings suggest that there is strong evidence for the entrainment hypothesis, and that entrainment depends on pre-stimulation alpha phase, which could explain our findings and lend support to

the hypothesis that stimulation with a faster harmonic frequency might not entrain but unbalance alpha oscillators. Since we did not test entrainment and phase synchronization systematically, these are merely hypotheses, and more research is necessary to corroborate this theory and to test whether there is an association with reduced clinical improvement. The stratification potential of high frequency stimulation as treatment for the low decile subgroup, on the other hand, represents a novel finding that has direct clinical implications.

LIMITATIONS

Data was recorded with different EEG systems and at different recording sites, introducing heterogeneity. Extensive analyses were conducted to assess whether the quadratic effect for 18 Hz rTMS on percent symptom improvement was affected by differing baseline severity at the different recording sites, and showed that the effect remained highly stable and significant (all $p<.006$) across multiple leave-one-out analyses. This suggests that no single treatment site individually contributed to the presented association.

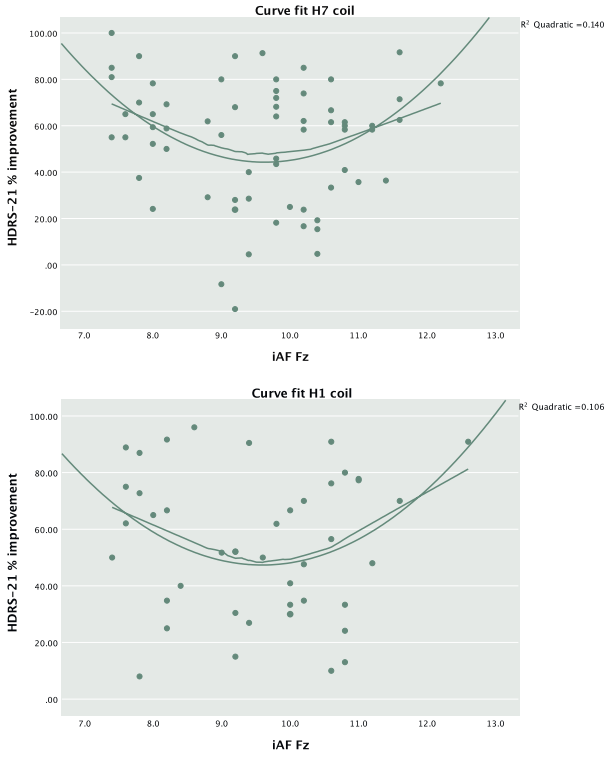
Since it is unknown whether the quadratic relationship between iAF and outcome to 18 Hz treatment is related to the treatment stimulation frequency or to the broader stimulation area of H-coils or to both these factors, it is perceivable that the here presented findings only hold for dTMS set-ups but not for 18 Hz stimulation administered with a standard figure-of-eight coil. Although the finding in the 20 Hz rTMS Dataset-3 seems to corroborate the 18 Hz finding, suggesting that frequency plays a bigger role than area, the fact that patients received bilateral treatment including 1 Hz stimulation weakens that argument.

The present study did not examine other potentially predictive factors, such as baseline severity since different scales were used to measure depression severity. The Brainmarker-I study (Voetterl, 2023), however, detected no differences in baseline severity between 5 datasets. Furthermore, Krepel et al. (Krepel et al., 2019) only found limited prediction potential of baseline variables for rTMS response. No age differences were found between subgroups.

CONCLUSION

Our findings enhance the clinical applicability and relevance of Brainmarker-I by suggesting an alternative treatment for ECT for the low decile subgroup.
Despite strong evidence for the entrainment hypothesis, at least stimulation at a faster harmonic frequency of the endogenous oscillations does not seem to entrain the oscillator but rather have a disruptive effect, thereby reducing symptom improvement. These results thus suggest that 10 Hz and 18 Hz TMS target different subgroups of MDD patients.

SUPPLEMENTARY MATERIAL



Supplementary Figures S5.1. *Quadratic effects for H1 and H7 coils separately.*

Dataset	Treatment	Low decile subgroup (PPV/nPPV)	SYNC subgroup (PPV/nPPV)	High decile subgroup (PPV/nPPV)
1	18 Hz dTMS	41 (+23%)	25 (-25%)	35 (+5%)
2	10 Hz rTMS	52 (-14%)	77 (+29%)	53 (-12%)
3	20 Hz rTMS	39 (+23%)	24 (-24%)	35 (+9%)

Supplementary table S5.1 Full subgroup analyses in all datasets. *Full subgroup analyses were conducted for all stimulation frequencies, calculating remission rates and normalized remission rates (PPV/nPPV) in all biomarker subgroups, i.e. low decile (decile 1-5), SYNC (9.6-10.4 Hz), and high decile (decile 6-10) subgroup.* dTMS= *deep transcranial magnetic stimulation,* rTMS = *repetitive transcranial magnetic stimulation,* PPV = *positive predictive value,* nPPV = *normalized positive predictive value*

SUPPLEMENTARY MATERIAL S5.1. POSITIVE PREDICTIVE VALUES

For 18 Hz TMS, an nPPV of -25% for the 10 Hz-synchronized subgroup (PPV=25%) and +14% for the non-synchronized subgroup (PPV=38%) was found. In Dataset-3 (20 Hz rTMS with a figure-of-eight coil), results resembled those of Dataset-1, with an nPPV of -24% (PPV=24%) in the 10 Hz-synchronized subgroup (i.e. the 20 Hz-harmonic), and an nPPV of +17% (PPV=38%) in the non-synchronized subgroup.

SUPPLEMENTARY MATERIAL S5.2. PRE- POST-TREATMENT CHANGES IN IAF

A repeated-measures ANOVA testing differences between iAF (at Fz) pre-treatment and post-treatment at group level was not significant ($F(1,98)=0.000$, $p=.984$). The mean of iAF at pre-treatment was 9.562 and 9.564 at post-treatment, remaining exceedingly stable. When repeating this analysis with remission as between-subject factor, there was no interaction between pre-/post-iAF and remission ($F_{(1,97)}=0.882$, $p=.350$). The mean of remitters was slightly lower at both pre- and post-treatment (9.321 and 9.455, respectively) compared to non-remitters (9.682 and 9.618, respectively). The same analysis repeated with response as between-subject factor again showed no significant interaction or main effects ($F_{(1,97)}=0.001$, $p=.979$). As expected, iAF remained highly stable from pre- to post-treatment, corroborating previous reports of high reproducibility of iAF and the idea of iAF representing a trait marker.

6

DISCUSSION

DISCUSSION

KEY FINDINGS

The book starts with a general introduction to depression, TMS and the need for robust treatment prediction biomarkers in Chapter 1.

Depression is a severe and debilitating disorder that is highly prevalent and, especially when unsuccessfully treated, can have detrimental effects on the lives of those affected, one of the most severe effects being a heightened suicide risk. Finding the best treatment for the individual early on in the course of the illness is therefore crucial. However, despite overall effectiveness of different approved depression treatments, a substantial number of patients experiences no or only little symptom improvement. With increasing failed treatment attempts, the likelihood of achieving remission diminishes, and after multiple futile treatment courses a patient is considered difficult-to-treat. These patients often move on to receive noninvasive brain stimulation, such as rTMS, which shows good efficacy at the group level, but still results in a considerable number of non-remitters. One strategy to enhance treatment

response is individualizing treatment to the patient, so called precision psychiatry. Within the context of precision psychiatry, stratified psychiatry can aid in identifying the best out of multiple approved treatments for a given disorder, for instance by means of imaging biomarkers. Throughout the years, a multitude of MRI and EEG markers has been examined, however valid criticism has been raised regarding the quality of EEG biomarker research, for instance a lack of out-of-sample validations, insufficient direct replication, differing processing methods, and general publication bias, concluding that EEG biomarkers are not yet ready for clinical implementation.

These issues are once more addressed in Chapter 2, which systematically reviews robustness of imaging biomarkers and compiles the state-of-the-art of robust imaging biomarkers for depression and TMS. Only MRI and EEG biomarker studies with sample sizes of ≥ 88 patients or studies that attempted external replication were included, ensuring robustness of any potential identified biomarker. In total 18 different studies were included per modality, however only 2 MRI and 2 EEG biomarkers were identified as robust: network mapping and DLPFC-sgACC functional connectivity for MRI, and frontal-midline rACC theta and iAF for EEG.

The subsequent 3 chapters contributed to the evidence for the robustness of one of these identified robust biomarkers, the iAF, which had already been indicated as predictor of treatment outcome for both ADHD and MDD although findings were initially inconsistent. One potential reason for the discrepant findings could be attributed to differences in methodology utilized across these studies, such as data cleaning, which had been one of the points of critique for prior biomarker studies.

Thus, in Chapter 3, EEG processing was first optimized in a large clinical dataset of over 4000 people for optimal detection of the alpha peak, by testing and validating 108 different processing permutations, relying on the well-known finding that iAF indexes brain maturation (N=4249). The resulting biomarker, Brainmarker-I, was then normalized for age and sex and tested for treatment stratification in ADHD. It was demonstrated that Brainmarker-I can successfully stratify boys with a rela-

tively faster iAF to stimulant medication and those with a slower iAF to multimodal neurofeedback treatment (MM-NFB). These findings were subsequently replicated in unseen independent datasets through blinded prediction of treatment outcome, thereby lending credibility to the presented findings and improving on criticism pertaining prior biomarker research, as well as demonstrating generalizability to different recording hardware and set-ups. Two other treatment options, guanfacine and atomoxetine, were evaluated, with guanfacine showing a similar prediction direction as MPH and atomoxetine as MM-NFB. For these two medications, no out-of-sample validations were conducted, therefore these findings still require replication.

In Chapter 4, the same iAF-based biomarker was shown to extend to depression treatments, first independently replicating the previous finding of better response to sertraline for the slow iAF subgroup in the EMBARC dataset in first-line depression treatment while proving a lack of predictive effect for placebo. For second-line MDD treatment, in addition to replicating the previous finding that patients with an iAF close to the stimulation frequency of 10 Hz are more likely to remit to 10 Hz rTMS, higher remission rates were found in a fast-iAF subgroup for 1 Hz rTMS and in a slow-iAF group for ECT. These new findings were once again replicated in independent, out-of-sample validations in a blinded fashion. Lastly, the study explored the biomarker's stratification potential for psychotherapy, bupropion and ketamine, with a trend of better remission to psychotherapy in the slow iAF subgroup but no stratification potential for bupropion and ketamine. The finding for psychotherapy still requires out-of-sample replication. This complements the fourth chapter, showing that Brainmarker-I represents a transdiagnostic treatment stratification biomarker for ADHD and MDD. This marker is both easy to collect and to interpret, and, thanks to the stratification approach, can be applied without risk, which aids in straightforward realization in clinical practice.

In Chapter 5, stratification potential of Brainmarker-I for depression was further investigated for deep-TMS administered at a higher stimulation frequency of 18 Hz. This TMS protocol was shown to be a valid alternative for the slow iAF subgroup that would have been advised ECT treatment before. A successful replication in 20 Hz TMS data that

administered stimulation with a figure-of-eight coil - instead of a deep TMS coil - suggested a higher relevance of stimulation frequency than coil shape, enabling Brainmarker-I stratification between TMS protocols only. Due to its benign side effect profile and outpatient setting, rTMS treatment can be considered a less intensive intervention than ECT when considering a stepped-care model. This is supported by patient preference for TMS before ECT treatment which makes full stratification to TMS protocols all the more valuable. Furthermore, it enables full implementation of Brainmarker-I at TMS clinics that have no means to provide ECT. The same study, furthermore, discusses the indication that stimulation with a frequency of 18 Hz might not lead to entrainment of endogenous oscillations at the harmonic frequency of 9 Hz, and suggests that it might even lead to an unbalancing of the oscillators.

The previous chapters introduce an iAF-based, age- and sex-normalized transdiagnostic stratification biomarker for ADHD and MDD, which is convenient to measure and calculate and has been directly replicated through blinded outcome prediction in independent, unseen datasets. The following chapter discusses the implications of such a biomarker.

THE STRATIFICATION APPROACH

One important aspect of stratification, that might be considered unconventional compared to standard scientific practice, is the more pragmatic approach that is adopted with regard to statistical methodology (Arns et al., 2023). For instance, the heterogeneity between and large volume of different datasets used in the previous chapters necessitated an adjustment of standard accuracy measures. The PPV (also called precision) which is part of the customary accuracy test battery was normalized for imbalances in base remitter rates between cohorts, resulting in a normalized version, the nPPV. Through the principle of stratification, described in the introduction, statistical significance is less relevant, since the biomarker never discourages a specific treatment for a patient without proposing a more promising alternative. Instead, the respective best option out of a number of effective, evidence-based treatments is identified for each patient meaning that a stratification biomarker will provide a treatment recommendation for

each individual. Hence the capacity to predict non-remission, as other accuracy measures such as the negative predictive value (NPV) and specificity do, does not add substantial value to a stratification marker, since a patient will always be assigned an effective treatment (Arns et al., 2022). Since significance testing is not based on the PPV or nPPV alone, but automatically takes into account other accuracy measures (NPV, sensitivity, specificity), the value of these other measures might be inflated while that of the measure of interest, the PPV, is discounted. This might mean that an increase in remission rate in a biomarker subgroup is valuable although it does not reach significance (when other accuracy measures are considered, too). Blinded out-of-sample validations are then crucial to ascertain that such non-significant results do not represent spurious findings.

As the aforementioned meta-analyses demonstrated, there are no differences between the standard stimulation protocols with respect to response and remission rates at the group level (Brunoni et al., 2016) (however, note that back when this meta-analysis was published newer approaches such as accelerated, deep and synchronized TMS were not found to be more effective than sham). For the most commonly used protocols – high frequency (HF) stimulation over the left DLPFC and low frequency (LF) stimulation over the right DLPFC- multiple meta-analyses demonstrated no differences in treatment efficacy (Brunoni et al., 2016; Cao et al., 2018; J. Chen et al., 2014; Fitzgerald et al., 2019). Similarly, bilateral stimulation, applied simultaneously or sequentially over left and right DLPFC, is equally but not more effective than unilateral HF left or LF right stimulation alone (Berlim et al., 2013; J. Chen et al., 2014). For iTBS, Blumberger et al. were the first to convincingly show that there was no difference in treatment outcome with HF left rTMS (Blumberger et al., 2018). Even response rates of less traditional higher frequency protocols such as 5, 18 and 20 Hz did not differ from the standard protocols (Filipčić et al., 2019; Leggett et al., 2015; T. Zhang et al., 2021). The stratification approach is, therefore, in line with the "primum non nocere" ("first, do no harm") principle, meaning that even a "wrong" treatment recommendation would not adversely affect a patient since they would still be prescribed an overall equally effective treatment. In fact, one might even go as far as to say that the worst that could happen when following the stratification approach is

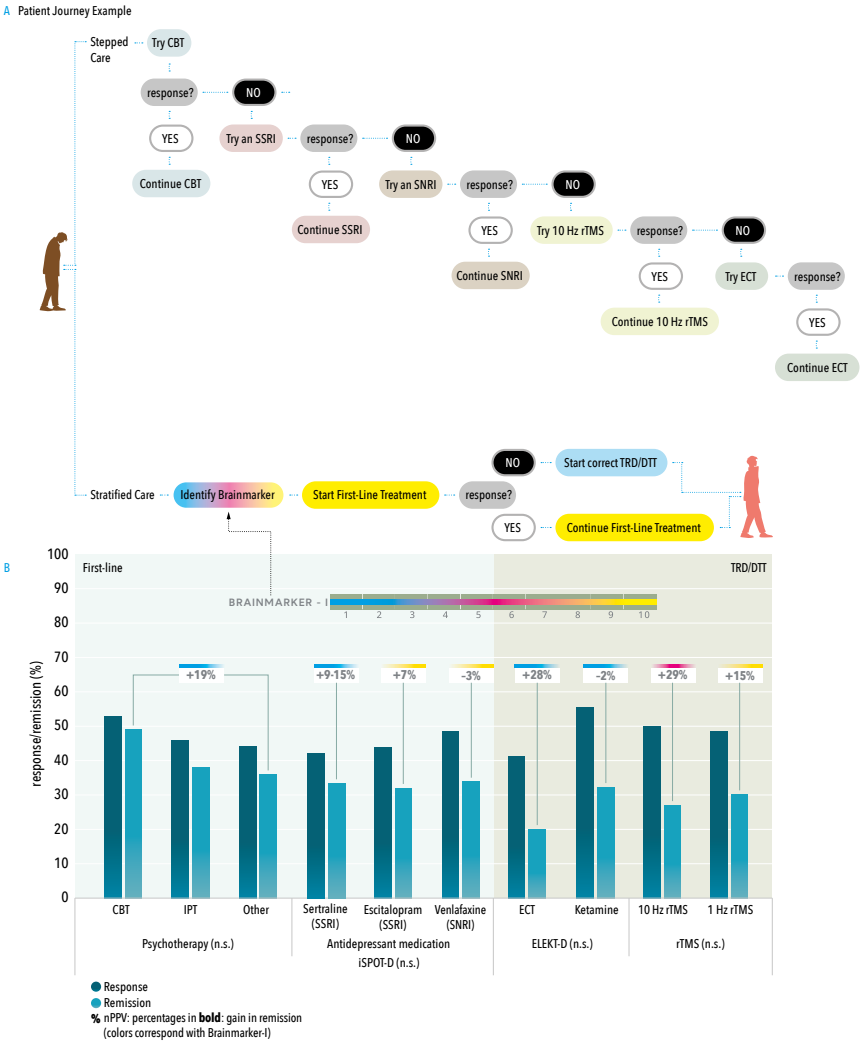
to reproduce the status quo of a one-size-fits-all prescription practice. Since Brainmarker-I can be determined based on a short resting-state EEG measurement and calculated in a straight-forward way, additional effort and cost for patient and clinic are limited compared to other biomarkers. Considering the blinded validations conducted for most of the tested treatments, it is very likely, however, that remission rates will in fact be enhanced after stratification.

In addition to the PPV/nPPV, the NNT provides further information on the size of the effect that we could expect when assigning patients of different biomarker subgroups to their biomarker-recommended treatment, compared to treatment-as-usual on a trial-and-error basis. The NNTs reported in the previous chapters can be considered a stratification simulation in pre-collected EEG datasets. They indicate in each biomarker subgroup how many patients need to be treated with the biomarker-recommended active treatment to get one more remitter than in the full, unstratified group that received active treatment and for which EEG data was available. The resulting NNTs for Brainmarker-I in depression demonstrate that the effect we found is of the same order of magnitude as the NNT reported for tricyclic and SSRI treatment compared to placebo (Arroll et al., 2009). Since we do not compare active vs inactive treatment, but stratified active vs non-stratified active treatment, this effect is clinically meaningful and arguably stronger than the antidepressant vs placebo effects.

As the amount of determined Brainmarker-I subgroups is rather small (3 groups total) and the investigated interventions numerous, some overlap of more first-line and more second-line treatment options was to be expected. For instance, Brainmarker-I predicted better clinical improvement in response to both sertraline, ECT and higher frequency TMS (18/20 Hz) treatment for the slow iAF subgroup. Naturally that doesn't mean that the biomarker advises an individual with a first-line depression to be prescribed ECT, although some might argue for administering rTMS earlier during the disease course. Instead, these findings suggest a stepped stratified care approach where the biomarker aids in informing a stepped-care decision, i.e., out of multiple evidence-based treatments that can be considered first-line or mild interventions, the patient is stratified to the option with the best likelihood

of improvement, while for more difficult-to-treat patients, a different range of interventions is considered for stratification (see Figure 6.1 reprinted from (Arns et al., 2023)).
Self-evidently, the stratification advice should be seen as a tool that advises the treating physician to choose the best treatment while at the same time still taking into account the patient's treatment history, sensitivities (e.g., medication), failed treatment attempts and more.

Figure 6.1. Stepped care versus stratified psychiatry *(following page) (reprinted from Arns et al., 2023 with permission)*
A. Example of the traditional stepped care approach (top), highlighting the large number of treatment attempts that many patients have to go through. This is contrasted with stratified care (bottom), informing stepped care by means of biomarkers such as Brainmarker-I, thereby reducing the number of treatment steps. B. Response and remission rates of large clinical trials or meta-analyses for different first-line depression treatments (light blue area on the left) and interventions typically described for difficult-to-treat depression (light-grey area on the right), indicating no significant difference within the modalities. Percentages above bar graphs depict expected change in remission rate with stratification as advised by Brainmarker-I. These percentages reflect normalized positive predictive values (nPPV), demonstrating the expected change in remission rate after stratification relative to group remission rate. Brainmarker-I advice is summarized by the color scale, e.g. nPPV percentages marked in blue indicate better chances of remission for the low decile subgroup, pink for the synchronized subgroup and yellow for the higher decile subgroup.
CBT: cognitive behavior therapy; IPT: interpersonal therapy; NS: not significant, SSRI: selective serotonin reuptake inhibitor; SNRI: serotonin norepinephrine reuptake inhibitor; ECT: electroconvulsive therapy; rTMS: repetitive transcranial magnetic stimulation; TRD/DTT: treatment-resistant depression/difficult-to-treat depression.


A Patient Journey Example
Stepped Care
Try CBT
response?
NO
YES
Continue CBT
Try an SSRI
response?
NO
YES
Continue SSRI
Try an SNRI
response?
NO
YES
Continue SNRI
Try 10 Hz rTMS
response?
NO
YES
Continue 10 Hz rTMS
Try ECT
response?
YES
Continue ECT
Stratified Care
Identify Brainmarker
Start First-Line Treatment
response?
NO
Start correct TRD/DTT
YES
Continue First-Line Treatment
B
First-line
TRD/DTT
BRAINMARKER - I
1 2 3 4 5 6 7 8 9 10
response/remission (%)
100 90 80 70 60 50 40 30 20 10 0
+19%
+9-15%
+7%
-3%
+28%
-2%
+29%
+15%
CBT
IPT
Other
Sertraline (SSRI)
Escitalopram (SSRI)
Venlafaxine (SNRI)
ECT
Ketamine
10 Hz rTMS
1 Hz rTMS
Psychotherapy (n.s.)
Antidepressant medication
iSPOT-D (n.s.)
ELEKT-D (n.s.)
rTMS (n.s.)
Response
Remission
% nPPV: percentages in bold: gain in remission
(colors correspond with Brainmarker-I)

As mentioned above, the nature of the stratification approach can aid in bridging the translational gap, bringing biomarkers into the clinic. The here adopted approach to identify biomarkers, examines effectiveness, as assessed by clinical trials, as opposed to efficacy, tested in RCTs. Although RCTs are usually regarded as the gold standard of research trials, there are some limitations to their real-world value. One issue with RCTs is that they often deploy strict exclusion criteria, meaning that only a limited sample of the actual patient population is effectively included in these studies (see (Taipale et al., 2022) for a review on stringent exclusion criteria in psychosis research). This limits clinical applicability since most patient populations, especially more difficult-to-treat patients usually have a host of comorbidities, complementary medication and a complicated treatment history, all of which are often exclusion criteria (Koslow et al., 2010). The approach that this project has followed may be less controlled (due to heterogeneous clinical datasets including patients with medication and comorbidities) but allows for broader implementation which is crucial if the biomarker is supposed to be of any real-life value.
In addition to the lack of heterogeneity in most RCTs, it has furthermore been shown that treatment response to the active intervention in RCTs decreases with increasing likelihood of receiving placebo treatment. This effect, coined "lessebo" effect by Sinyor et al. (Sinyor et al., 2010), is hypothesized to be related to the heightened expectation of a negative event occurring (here: receiving placebo instead of active treatment) – a thought pattern pervasive in depression patients. This phenomenon could potentially be relevant in clinical practice, too, however, the impact of it has been shown to decrease with a lower chance of placebo administration. One study where such effects can be observed is the EMBARC study which showed no significant difference in HDRS outcomes between placebo and active compound sertraline at week 8 due to a high placebo remission rate of 33% and simultaneous comparatively low improvement following sertraline (39% remission compared to 47% in iSPOT-D which tested only active treatments) (Arns, Bruder, et al., 2016; Nguyen et al., 2022; Webb et al., 2019). RCTs might thus underestimate treatment effects achieved in clinical practice, if the likelihood to receive the placebo is

high. On the other hand, RCTs can also exaggerate treatment effect compared to real-life findings due to more intensive clinician attention in the RCT than is feasible in standard practice and due to higher levels of comorbidities in the effectiveness trial that are generally harder to treat (van der Lem et al., 2012; Singal et al., 2014).
Moreover, setting up and conducting a well-run RCT requires resources and time; often many years pass from study design to eventual publication (Ross et al., 2012). During this time, patients would still be treated according to a trial-and-error approach. The question arises whether it can be ethically justified to deny patients a tool that can potentially lead to faster and stronger clinical improvement and can do no harm. Although it will be valuable to confirm the presented results on the iAF biomarker in a more controlled, systematic way, making RCTs a prerequisite for clinical implementation of biomarkers could be considered unethical.

A blinded out-of-sample replication approach, as utilized in the previous chapters could function as an alternative to double-blind placebo-controlled RCTs to guarantee robustness of potential findings and clinical value of the identified biomarker.

ADHD POTENTIAL WORKING MECHANISMS & BIOMARKERS

Although the main focus of this thesis lies on biomarkers for depression treatment, Brainmarker-I has first shown merit for ADHD treatment stratification. The study summarized in Chapter 3 was informed on earlier findings, for instance in iSPOT-A where a slow iAF indicated non-response to stimulant medication in adolescent boys (Arns et al., 2008, 2018), and by Krepel et al., suggesting better remission rates for neurofeedback in boys with a slow iAF (Krepel et al., 2020).

Although a lengthy discussion is outside the scope of this work, a note on NFB seems warranted due to persistent criticism and inconclusive meta-analyses that question its value (Cortese et al., 2016). The reputation of NFB for ADHD has long suffered from a lack of restrictions and consistency, for instance in the protocols, EEG set-ups, reward thresholding and reward signals used (Arns, Heinrich, et al., 2014). When NFB protocols are chosen that have been tested, are backed

by research, and preferentially informed by a baseline EEG, response and remission rates are suggested to be high (Arns, Drinkenburg, & Kenemans, 2012; Krepel et al., 2020; Monastra et al., 2002). However, despite being able to show significant symptom improvement after NFB, a recent double-blind RCT could not demonstrate a significant difference in treatment effect between NFB and control group (The Neurofeedback Collaborative Group et al., 2021), albeit treatment effects at 2-year follow-up were similar to treatment with methylphenidate or multicomponent behavior therapy (The Neurofeedback Collaborative Group et al., 2022). One reason for the lack of difference in this RCT and other studies might be the multimodal nature of many neurofeedback interventions, that often provide lifestyle and sleep advice, in addition to 30-40 intensive training sessions where patients may enjoy undivided attention by the NFB clinician. These interventions are usually also applied in the sham group and might contribute as much to treatment success as the NFB training itself. Nonetheless, a recent review of RCTs and open label trials, applying standard protocols, found medium to large effect sizes, confirming that NFB can be considered a valuable alternative to pharmacological treatment in ADHD (Arns et al., 2020) with sustained benefit beyond 2 years.

In the iSPOT-A study, that served as basis for Chapter 3, a lack of the typical maturational iAF trajectory was found in non-responders, arguing for a stunted development in these patients, and suggesting a different underlying etiology in this slow iAF subgroup. Although the iAF is traditionally considered a trait marker with a solid hereditary component (Van Beijsterveldt & van Baal, 2002; Posthuma et al., 2001) which remains fairly stable across time (Kondacs & Szabó, 1999), some studies have suggested that iAF can normalize with successful treatment. For instance, slowed iAF in chronic pain patients - potentially slowed down as a result of the chronicity - gradually increased after successful pain relief (Sarnthein et al., 2006). Ulrich et al. already found early on that responders to antidepressant medication showed a slight increase in iAF of 0.5 Hz while this was not the case in non-responders who were characterized by a slow iAF (Ulrich et al., 1984). Since there is no iAF post-treatment data available for the slow iAF subgroup that showed a higher likelihood for remission in Krepel et al. (Krepel et al., 2020), it is not possible to draw any conclu-

sions on potential changes in iAF based on the study described in this thesis. However, it is conceivable that a normalization of iAF occurs with symptom relief, through the alleviation of sleeping problems. ADHD is commonly characterized by diminished sleep quality, leading to increased daytime sleepiness and thus to the typically experienced symptoms of decreased attention and heightened impulsivity – assumed to be a mechanism counteracting said drowsiness (Ryan et al., 2019). Sleep deprivation has been shown to decrease iAF, which correlated with an increase in sleepiness (Quiquempoix et al., 2023). Klimesch proposed that this was driven by an increase in power in the slow alpha band and a decrease in the fast alpha band, potentially reflecting a means to counteract drowsiness in those individuals (Klimesch, 1999). This could explain the iAF slowing in this subgroup which might be ameliorated by an improvement in sleep quality. As mentioned before, NFB is often augmented by sleep hygiene protocols, and has been shown to mitigate sleeping problems which correlated with treatment effect (Ryan et al., 2019). It might therefore treat a distinct ADHD phenotype, characterized by sleeping problems and iAF slowing.
Stimulant medication, on the other hand, might be more adequate for an ADHD subgroup presenting with no sleep problems, since stimulants can aggravate already dysfunctional sleep patterns and provide better symptom relief in patients who have longer sleep duration (Ryan et al., 2019).

Biomarkers for ADHD have mostly focused on the diagnostic aspect. A recent review summarized the literature on predictive imaging biomarkers for various treatments of ADHD including the two previously mentioned studies by Arns and Krepel, the here presented Brainmarker-I findings and one more study, that found no difference in iAF between responders and non-responders to MPH or dexamphetamine (Michelini et al., 2022). The review concludes that prospective testing of biomarkers is required before actual clinical implementation. Although this would indeed be very valuable, we propose that Brainmarker-I can already be implemented, since the stratification approach ensures stratification to an evidence-based intervention without denying a specific treatment, as argued above.

For NFB, even fewer predictive biomarkers have been suggested and sufficiently tested (Buitelaar et al., 2022). Identifying valuable biomarkers is thus even more relevant. Biomarkers measured with EEG might be particularly valuable since the EEG could simultaneously be utilized to assist NFB protocol choice.

HYPOTHESES OF THE UNDERLYING NEUROBIOLOGY IN DEPRESSION

Importantly, Brainmarker-I is not suggested or intended to explain the working mechanism of any tested treatment. Although this work ventures to propose potential hypotheses (e.g. see entrainment and phase dependence hypotheses in Chapter 5), the studies presented here have not tested potential neurobiological bases systematically and the following paragraphs are therefore mere conjecture.

One idea presented in this work is that of TMS pulses entraining endogenous oscillations. There is broad evidence supporting this hypothesis (see Chapter 5). However, the question remains whether this entrainment also plays a role in enhancing treatment response. In Chapter 4 and 5, and in several prior studies (Corlier et al., 2019; Roelofs et al., 2020), we saw that administering 10 Hz rTMS to patients whose iAF was already close to the stimulation frequency of 10 Hz results in higher remission rates, suggesting that entrainment of the endogenous oscillations indeed fosters symptom improvement. This moreover suggests that stimulating each patient at their individual best frequency might enhance clinical improvement. A very recent study tested this by stimulating each patient at their respective iAF (George et al., 2023). In addition, one patient group received TMS stimulation that was synchronized to the patient's preferred oscillatory phase, i.e. the phase where the individual's prefrontal alpha oscillation led to strongest BOLD response in the dorsal ACC. One conclusion from this small study was that entrainment is only achieved in the subgroup that received synchronized stimulation at their individual best frequency, not in the unsynchronized subgroup. The level of entrainment across sessions in this group was correlated with personal clinical improvement. However, entrainment did not lead to better overall improvement in the synchronized group compared to the unsynchronized group. Moreover, the overall treat-

ment effect for both groups did not seem to differ – or might even be smaller - compared to what would be expected for standard 10 Hz treatment, although this is difficult to conclude since a 10 Hz control group was lacking. One issue could be that only the first pulse was synchronized to the preferred phase and not each pulse, as is common in a closed-loop system. Applying a pulse at that preferred phase would shift the oscillation. This means that the following pulses would not be applied at that same preferred point of the oscillation anymore, which might potentially still result in entrainment but not in stimulation at the preferred phase. In summary, this study indicates that stimulation individualized to the iAF might not result in improved treatment outcome which is supported by another small pilot study that found best treatment outcome to 10 Hz stimulation in patients with a 9 Hz iAF (Arns et al., 2010).

One possible explanation why not all entrainment to the iAF might lead to improved treatment response could be that, although differences in the iAF are also common in the general population, the average iAF of a middle-aged healthy adult lies around 9.8 Hz, while slowing of iAF is often related to mental health problems, organic causes, white matter damage - for instance in Alzheimer's disease - and other disorders (Puttaert et al., 2021; Rodriguez et al., 1999), as summarized in Chapter 3 and 4. Entraining such an unfavorable rhythm might thus not be beneficial since it is not a preferred rhythm of the brain to begin with. A faster (than usual) iAF is generally considered to be more beneficial, likely reflecting faster information transfer in a thalamocortical loop; it is, moreover, related to improved cognitive performance, as discussed before, and linked to acute (e.g. task-related) vigilance and arousal (Ramsay et al., 2021). However, a general high state of arousal might be reflective of a different pathological subtype, characterized by hyper-arousal, which is possibly related to trauma (Wahbeh & Oken, 2013). Entraining this faster, hypervigilant rhythm, again, might not be therapeutic. Such a hypervigilant biotype, preliminarily proposed by Williams, is characterized by hyperconnectivity in the frontoparietal attention network which shows hypoconnectivity with the positive affect circuit and has been observed in social anxiety (Williams, 2016). Why these or similar subtypes might respond better to very high-frequency 18 Hz stimulation or ECT (for slow iAF) or to low-frequency 1 Hz stimulation (for the faster, hypervigilant

biotype) is unknown. One explanation might be that TMS can elicit different synaptic plasticity mechanism, such as inhibitory long-term depression (LTD) and facilitatory long-term potentiation (LTP) that are responsible for the longer-lasting effects of magnetic stimulation. Although low-frequency TMS is commonly assumed to be inhibitory and high frequency stimulation to be excitatory, this relationship has proven more complicated. At the single cell level, single-pulse TMS can have both excitatory and inhibitory effects (Romero et al., 2019), and both LTD-like and LTP-like plasticity have been observed after low-frequency TMS, depending on the stimulated phase (Baur et al., 2020). However, despite a trend towards LTP-like plasticity when the trough of the oscillatory phase, i.e., the high-excitability state, was stimulated with LF stimulation, LTD-like plasticity was generally more readily induced than LTP-like plasticity. Inhibiting this aforementioned hyperconnected frontoparietal attention network in the hypervigilant biotype, that Williams et al. proposed, through 1 Hz stimulation might, thus, induce LTD-like plasticity that could normalize network connectivity. High frequency stimulation, including 18 or 20 Hz protocols, on the other hand, usually have excitatory effects (Huerta & Volpe, 2009). However, it has been suggested that whether TMS leads to excitatory and inhibitory effect depends on multiple diverse internal and external factors (Hartwigsen & Silvanto, 2023). For instance, the brain's excitability state fluctuates depending on task-demand, resting condition or other internal factors such as fatigue, but is also influenced by external aspects, such as stimulation intensity, frequency and duration.

The potential mechanism underlying the findings of lower remission rates in patients with a 9 Hz iAF receiving higher frequency stimulation has been discussed in Chapter 5. Taking into account the results by Baur et al., suggesting that plasticity is linked to excitability states, it is indeed conceivable that 18 Hz simultaneously inhibits and excites neurons within one cycle (i.e., through stimulating both the peak and the trough), cancelling out these effects. In a slower alpha oscillation, the TMS pulse would stimulate a different oscillatory phase at each cycle, thereby being more likely to trigger more of the high-excitability states.

We are only beginning to understand how rTMS affects different neuronal and non-neuronal types of plasticity and exerts its effect in mental disorders. Testing the above hypotheses systematically should be a focus of future research.

IMPLICATIONS OF THIS WORK & FUTURE DIRECTIONS

As shortly touched upon in the introduction, depression is a life-threatening disorder.
Although one might think that prescribing a non-effective treatment (for the individual) first and only finding an effective alternative at a later disease stage only postpones symptom relief, research has shown that that is not the case. As mentioned before, an untreated (or unsuccessfully treated) depression often becomes more severe (Fils et al., 2010; Fogel et al., 2006; Ghio et al., 2014), and longer disease course and higher symptom severity make it harder to treat a depression effectively and elevate the risk for relapse (Buckman et al., 2018; O'Leary et al., 2000). Even more importantly, severe depression and time to full remission are strong predictors of suicide attempt, and higher depression severity increases the risk of death by suicide (Riera-Serra et al., 2023). Thus, identifying the best option out of the available first-line, second-line or late-line interventions can potentially be life-saving.

An evident research question arising from the here presented suggestions on entrainment with 10 Hz stimulation has been discussed above. Studies investigating whether stimulation at the iAF would lead to even better symptom improvement are at best inconsistent. A more promising approach might take into account the individual brain state during or prior to stimulation, as suggested by Sack et al. (Sack et al., 2023). For instance, applying the TMS pulse at the preferred phase might have clinical merit, if phase-locking is done correctly (Schilberg et al., 2021). However, a personalized closed-loop EEG-TMS approach is not viable for clinical implementation yet and more research is needed to bring this set-up to the clinic (Zrenner & Ziemann, 2023).

In the meantime, complementing Brainmarker-I by other, potentially more nuanced predictors to refine treatment advice would likely result in even higher likelihood to achieve remission. As summarized in

Chapter 2, there are multiple candidates to augment Brainmarker-I. In addition, besides EEG and fMRI markers, cardiac, genetic, neuroendocrinal and behavioral predictors could likely add relevant information not provided by neurophysiology.
For ADHD, Brainmarker-I was shown to be able to stratify boys to different treatments, however the available data in girls was limited such that no prediction could be made. Since ADHD is more commonly diagnosed in boys, a male-female gap exists in ADHD research, making findings for girls even more relevant (Faraone et al., 2000). Similarly, since ADHD often persists in adulthood, studies examining stratification biomarkers for adults with ADHD would be valuable.

Naturally, identifying biomarkers for treatment outcome in disorders other than depression and ADHD, is highly relevant. Since iAF shows transdiagnostic value, it would be interesting to investigate whether it has any predictive potential in other disorders.
To be able to follow a rigorous testing and validation approach in heterogeneous clinical datasets in order to identify the best treatment for each subgroup from a broad range of treatment options, data sharing is of utmost importance. The research presented in this thesis would not have been possible without collaborators sharing their clinical and neurophysiological data. While data sharing initiatives are underway (e.g. TDBRAIN) (Arns et al., 2023; van Dijk et al., 2022; Koslow et al., 2010), more work is needed to facilitate direct replication attempts and pave the way for more robust research in precision psychiatry.

THE NEW STATE-OF-THE-ART OF EEG BIOMARKERS IN DEPRESSION

This thesis starts with an overview of the current issues that biomarker research is facing and summarizes the state-of-the-art of imaging biomarkers for TMS treatment in depression.

The different chapters contributed to and improved on this research by presenting a robust transdiagnostic treatment stratification biomarker based on the optimized iAF. What makes this biomarker particularly valuable is the replication effort in independent data that all findings had to undergo, strengthening its robustness. As Widge and colleagues explain, there are multiple methods that can be de-

ployed in biomarker research "that increase the reliability of conclusions. Chief among these is independent sample verification or cross-validation—reporting the algorithm's predictive performance on a sample of patients separate from those originally used to develop it. Cross-validation has repeatedly been highlighted as essential in the development of a valid biomarker" (Widge et al., 2019). Validating new findings in unseen, independent samples (preferably from other labs and locations) in blinded predictions, furthermore ensures that the identified marker can adequately deal with the high heterogeneity of many mental disorders, allowing for broader applicability of the marker to diverse patient populations. This is, moreover, fostered both by the stratification approach and the ubiquitous nature of the alpha rhythm.

Other points of critique regarding EEG biomarkers were underreporting of negative findings, small sample size, and differences in methodology applied in studies of the same predictor, resulting in a lack of direct replication (Widge et al., 2019). In Chapter 1 of this thesis, imaging biomarker studies, amongst them iAF studies, were specifically chosen based on robustness including a minimum sample size of 88 and/or independent out-of-sample replications. Available null findings and unsuccessful replications were taken into account in evaluating a biomarker's robustness. Lastly, differences in methodology in prior studies were acknowledged and EEG processing was optimized and unified in the development of Brainmarker-I to make findings comparable and to ensure direct replication.

Five years after this meta-analysis by Widge et al. was published, the same author specifically argued for the value of the stratification principle in biomarker research. Once more the relevance of external validation is stressed - here mainly focused on machine learning predictors, where external validation is even more crucial (see discussion in Chapter 1) (Grzenda & Widge, 2023) - and the authors caution that the translation to clinical practice is often forfeited due to lack of such external validation and methodological flaws. The present work took many of the mentioned criticisms to heart and adopted important measures to address this translational gap. Thanks to rigorous testing and thorough validation, this thesis can present a robust EEG

biomarker, that can inform treatment for ADHD and depression - at different levels of depression severity - and that is currently being evaluated prospectively in clinical practice. If this trial run in a clinical setting indeed proves enhanced treatment outcome following Brainmarker-I advice, this would mark the beginning of broad biomarker implementation in clinical practice.

REFERENCES

Allsopp, K., Read, J., Corcoran, R. & Kinderman, P. (2019). Heterogeneity in psychiatric diagnostic classification. *Psychiatry Research*, 279, 15–22. https://doi.org/10.1016/j.psychres.2019.07.005

Alyagon, U., Shahar, H., Hadar, A., Barnea-Ygael, N., Lazarovits, A., Shalev, H. & Zangen, A. (2020). Alleviation of ADHD symptoms by non-invasive right prefrontal stimulation is correlated with EEG activity. *NeuroImage: Clinical, 26*, 102206. https://doi.org/10.1016/j.nicl.2020.102206

Arns, M. (2012). EEG-Based Personalized Medicine in ADHD: Individual Alpha Peak Frequency as an Endophenotype Associated with Nonresponse. *Journal of Neurotherapy, 16*(2), 123–141. https://doi.org/10.1080/10874208.2012.677664

Arns, M., Bruder, G., Hegerl, U., Spooner, C., Palmer, D. M., Etkin, A., Fallahpour, K., Gatt, J. M., Hirshberg, L. & 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*(1), 509–519. https://doi.org/10.1016/j.clinph.2015.05.032

Arns, M., Cerquera, A., Gutiérrez, R. M., Hasselman, F. & Freund, J. A. (2014). Non-linear EEG analyses predict non-response to rTMS treatment in major depressive disorder. *Clinical Neurophysiology, 125*(7), 1392–1399. https://doi.org/10.1016/j.clinph.2013.11.022

Arns, M., Clark, C. R., Trullinger, M., deBeus, R., Mack, M. & Aniftos, M. (2020). Neurofeedback and Attention-Deficit/Hyperactivity-Disorder (ADHD) in Children: Rating the Evidence and Proposed Guidelines. *Applied Psychophysiology and Biofeedback, 45*(2), 39–48. https://doi.org/10.1007/s10484-020-09455-2

Arns, M., van Dijk, H., Luykx, J. J., van Wingen, G. & Olbrich, S. (2022). Stratified psychiatry: Tomorrow's precision psychiatry? *European Neuropsychopharmacology, 55*, 14–19. https://doi.org/10.1016/j.euroneuro.2021.10.863

Arns, M., Drinkenburg, W. H., Fitzgerald, P. B. & Kenemans, J. L. (2012). Neurophysiological predictors of non-response to rTMS in depression. B*rain Stimulation, 5*(4), 569–576. https://doi.org/10.1016/j.brs.2011.12.003

Arns, M., Drinkenburg, W. & Kenemans, J. L. (2012). The Effects of QEEG-Informed Neurofeedback in ADHD: An Open-Label Pilot Study. *Applied Psychophysiology and Biofeedback, 37*(3), 171–180. https://doi.org/10.1007/s10484-012-9191-4

Arns, M., Feddema, I. & Kenemans, J. L. (2014). Differential effects of theta/beta and SMR neurofeedback in ADHD on sleep onset latency. *Frontiers in Human Neuro*

science, 8, 1019. https://doi.org/10.3389/fnhum.2014.01019
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. https://doi.org/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, 07*(03), 421–438. https://doi.org/10.1142/s0219635208001897
Arns, M., Heinrich, H. & Strehl, U. (2014). Evaluation of neurofeedback in ADHD: The long and winding road. *Biological Psychology, 95*, 108–115. https://doi.org/10.1016/j.biopsycho.2013.11.013
Arns, M. & Kenemans, J. L. (2014). Neurofeedback in ADHD and insomnia: Vigilance stabilization through sleep spindles and circadian networks. *Neuroscience & Biobehavioral Reviews, 44*, 183–194. https://doi.org/10.1016/j.neubiorev.2012.10.006
Arns, M., Olbrich, S. & Sack, A. T. (2023). Biomarker-driven stratified psychiatry: from stepped-care to matched-care in mental health. *Nature Mental Health, 1–3*. https://doi.org/10.1038/s44220-023-00156-3
Arns, M., Spronk, D. & Fitzgerald, P. B. (2010). Potential differential effects of 9 Hz rTMS and 10 Hz rTMS in the treatment of depression. *Brain Stimulation: Basic, Translational, and Clinical Research in Neuromodulation, 3*(2), 124–126. https://doi.org/10.1016/j.brs.2009.07.005
Arns, M., Vollebregt, M. A., Palmer, D., Spooner, C., Gordon, E., Kohn, M., Clarke, S., Elliott, G. R. & Buitelaar, J. K. (2018). Electroencephalographic biomarkers as predictors of methylphenidate response in attention-deficit/hyperactivity disorder. *European Neuropsychopharmacology, 28*(8), 881–891. https://doi.org/10.1016/j.euroneuro.2018.06.002
Arroll, B., Chin, W., Martis, W., Goodyear-Smith, F., Mount, V., Kingsford, D., Humm, S., Blashki, G. & MacGillivray, S. (2016). Antidepressants for treatment of depression in primary care: a systematic review and meta-analysis. *Journal of Primary Health Care, 8*(4), 325–334. https://doi.org/10.1071/hc16008
Arroll, B., Elley, C. R., Fishman, T., Goodyear-Smith, F. A., Kenealy, T., Blashki, G., Kerse, N. & MacGillivray, S. (2009). Antidepressants versus placebo for depression in primary care. *Cochrane Database of Systematic Reviews, 3,* CD007954. https://doi.org/10.1002/14651858.cd007954
Atkinson, A., Colburn, W., Degruttola, V., Demets, D., Downing, G. J., Hoth, D., Oates, J. A., peck, C. C., Schooley, R., Spilker, B., Woodcock, J. & Zeger, S. (2001). Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clinical Pharmacology & Therapeutics, 69*(3), 89–95. https://doi.org/10.1067/mcp.2001.113989
Baeken, C., van Beek, V., Vanderhasselt, M., Duprat, R. & Klooster, D. (2021). Cortical Thickness in the Right Anterior Cingulate Cortex Relates to Clinical Response to Left Prefrontal Accelerated Intermittent Theta Burst Stimulation: An Exploratory Study. *Neuromodulation: Technology at the Neural Interface.* https://doi.org/10.1111/ner.13380
Baeken, C., Brem, A.-K., Arns, M., Brunoni, A. R., Filipčić, I., Ganho-Ávila, A., Langguth, B., Padberg, F., Poulet, E., Rachid, F., Sack, A. T., Vanderhasselt, M.-A. & Bennabi, D. (2019). Repetitive transcranial magnetic stimulation treatment for depressive disorders. *Current Opinion in Psychiatry, 32*(5), 409–415. https://doi.

org/10.1097/yco.0000000000000533
Baeken, C., Duprat, R., Wu, G.-R., Raedt, R. D. & van Heeringen, K. (2017). Subgenual Anterior Cingulate–Medial Orbitofrontal Functional Connectivity in Medication-Resistant Major Depression: A Neurobiological Marker for Accelerated Intermittent Theta Burst Stimulation Treatment? *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 2*(7), 556–565. https://doi.org/10.1016/j.bpsc.2017.01.001
Baeken, C., Marinazzo, D., Wu, G.-R., Schuerbeek, P. V., Mey, J. D., Marchetti, I., Vanderhasselt, M.-A., Remue, J., Luypaert, R. & Raedt, R. D. (2014). Accelerated HF-rTMS in treatment-resistant unipolar depression: Insights from subgenual anterior cingulate functional connectivity. *The World Journal of Biological Psychiatry, 15*(4), 286–297. https://doi.org/10.3109/15622975.2013.872295
Bailey, N., Hoy, K., Rogasch, N., Thomson, R., McQueen, S., Elliot, D., Sullivan, C., Fulcher, B., Daskalakis, Z. & Fitzgerald, P. (2019). Differentiating responders and non-responders to rTMS treatment for depression after one week using resting EEG connectivity measures. *Journal of Affective Disorders, 242*, 68–79. https://doi.org/10.1016/j.jad.2018.08.058
Bailey, N. W., Krepel, N., van Dijk, H., Leuchter, A. F., Vila-Rodriguez, F., Blumberger, D. M., Downar, J., Wilson, A., Daskalakis, Z. J., Carpenter, L. L., Corlier, J., Arns, M. & Fitzgerald, P. B. (2021). Resting EEG theta connectivity and alpha power to predict repetitive transcranial magnetic stimulation response in depression: A non-replication from the ICON-DB consortium. *Clinical Neurophysiology, 132*(2), 650–659. https://doi.org/10.1016/j.clinph.2020.10.018
Banaschewski, T., Becker, K., Döpfner, M., Holtmann, M., Rösler, M. & Romanos, M. (2017). Attention-Deficit/Hyperactivity Disorder: A Current Overview. *Deutsches Aerzteblatt Online, 114*(9), 149–159. https://doi.org/10.3238/arztebl.2017.0149
Bares, M., Brunovsky, M., Novak, T., Kopecek, M., Stopkova, P., Sos, P. & Höschl, C. (2015). QEEG Theta Cordance in the Prediction of Treatment Outcome to Prefrontal Repetitive Transcranial Magnetic Stimulation or Venlafaxine ER in Patients With Major Depressive Disorder. *Clinical EEG and Neuroscience, 46*(2), 73–80. https://doi.org/10.1177/1550059413520442
Barker, A. T., Jalinous, R. & Freeston, I. L. (1985). Non-invasive magnetic stimulation of human motor cortex. *The Lancet, 325*(8437), 1106–1107. https://doi.org/10.1016/s0140-6736(85)92413-4
Baur, D., Galevska, D., Hussain, S., Cohen, L. G., Ziemann, U. & Zrenner, C. (2020). Induction of LTD-like corticospinal plasticity by low-frequency rTMS depends on pre-stimulus phase of sensorimotor μ-rhythm. *Brain Stimulation, 13*(6), 1580–1587. https://doi.org/10.1016/j.brs.2020.09.005
Bazanova, O. M., Auer, T. & Sapina, E. A. (2018). On the Efficiency of Individualized Theta/Beta Ratio Neurofeedback Combined with Forehead EMG Training in ADHD Children. *Frontiers in Human Neuroscience, 12*, 3. https://doi.org/10.3389/fnhum.2018.00003
Bedard, K. & Witman, A. (2020). Family structure and the gender gap in ADHD. *Review of Economics of the Household, 18*(4), 1101–1129. https://doi.org/10.1007/s11150-019-09476-9
van Beijsterveldt, C. E. M. & van Baal, G. C. M. (2002). Twin and family studies of the human electroencephalogram: a review and a meta-analysis. *Biological Psychology, 61*(1–2), 111–138. https://doi.org/10.1016/s0301-0511(02)00055-8
Bellato, A., Arora, I., Hollis, C. & Groom, M. J. (2020). Is autonomic nervous system

function atypical in attention deficit hyperactivity disorder (ADHD)? A systematic review of the evidence. *Neuroscience & Biobehavioral Reviews, 108*, 182–206. https://doi.org/10.1016/j.neubiorev.2019.11.001

Beniczky, S. & Schomer, D. L. (2020). Electroencephalography: basic biophysical and technological aspects important for clinical applications. *Epileptic Disorders, 22*(6), 697–715. https://doi.org/10.1684/epd.2020.1217

Berlim, M. T., Eynde, F. V. den & Daskalakis, Z. J. (2013). A systematic review and meta-analysis on the efficacy and acceptability of bilateral repetitive transcranial magnetic stimulation (rTMS) for treating major depression. *Psychological Medicine, 43*(11), 2245–2254. https://doi.org/10.1017/s0033291712002802

Berlim, M. T., van den Eynde, F., Tovar-Perdomo, S. & Daskalakis, Z. J. (2014). Response, remission and drop-out rates following high-frequency repetitive transcranial magnetic stimulation (rTMS) for treating major depression: a systematic review and meta-analysis of randomized, double-blind and sham-controlled trials. *Psychological Medicine, 44*(2), 225–239. https://doi.org/10.1017/s0033291713000512

Biomarkers Definitions Working Group(2001). Biomarkers and surrogate endpoints: Preferred definitions and conceptual framework. *Clinical Pharmacology & Therapeutics, 69*(3), 89–95. https://doi.org/10.1067/mcp.2001.113989

Björkholm, C., & Monteggia, L. M. (2016). BDNF - a key transducer of antidepressant effects. *Neuropharmacology, 102*, 72–79. https://doi.org/10.1016/j.neuropharm.2015.10.034

Blumberger, D. M., Vila-Rodriguez, F., Thorpe, K. E., Feffer, K., Noda, Y., Giacobbe, P., Knyahnytska, Y., Kennedy, S. H., Lam, R. W., Daskalakis, Z. J. & Downar, J. (2018). Effectiveness of theta burst versus high-frequency repetitive transcranial magnetic stimulation in patients with depression (THREE-D): a randomised non-inferiority trial. *The Lancet, 391*(10131), 1683–1692. https://doi.org/10.1016/s0140-6736(18)30295-2

Boes, A. D., Prasad, S., Liu, H., Liu, Q., Pascual-Leone, A., Caviness, V. S. & Fox, M. D. (2015). Network localization of neurological symptoms from focal brain lesions. *Brain, 138*(10), 3061–3075. https://doi.org/10.1093/brain/awv228

Boes, A. D., Uitermarkt, B. D., Albazron, F. M., Lan, M. J., Liston, C., Pascual-Leone, A., Dubin, M. J. & Fox, M. D. (2018). Rostral anterior cingulate cortex is a structural correlate of repetitive TMS treatment response in depression. *Brain Stimulation, 11*(3), 575–581. https://doi.org/10.1016/j.brs.2018.01.029

Boord, P., Siddall, P. J., Tran, Y., Herbert, D., Middleton, J. & Craig, A. (2008). Electroencephalographic slowing and reduced reactivity in neuropathic pain following spinal cord injury. *Spinal Cord, 46*(2), 118–123. https://doi.org/10.1038/sj.sc.3102077

Boutros, N. N. (1996). Diffuse electroencephalogram slowing in psychiatric patients: a preliminary report. *Journal of Psychiatry & Neuroscience : JPN, 21*(4), 259–263.

Brunoni, A. R., Chaimani, A., Moffa, A. H., Razza, L. B., Gattaz, W. F., Daskalakis, Z. J. & Carvalho, A. F. (2016). Repetitive Transcranial Magnetic Stimulation for the Acute Treatment of Major Depressive Episodes: A Systematic Review With Network Meta-analysis. *JAMA Psychiatry, 74*(2), 143. https://doi.org/10.1001/jamapsychiatry.2016.3644

Brunoni, A. R., Lopes, M., Kaptchuk, T. J. & Fregni, F. (2009). Placebo Response of Non-Pharmacological and Pharmacological Trials in Major Depression: A Systematic Review and Meta-Analysis. *PLoS ONE, 4*(3), e4824. https://doi.org/10.1371/journal.pone.0004824

Buckman, J. E. J., Underwood, A., Clarke, K., Saunders, R., Hollon, S. D., Fearon, P. &

Pilling, S. (2018). Risk factors for relapse and recurrence of depression in adults and how they operate: A four-phase systematic review and meta-synthesis. *Clinical Psychology Review, 64*, 13–38. https://doi.org/10.1016/j.cpr.2018.07.005

Buitelaar, J., Bölte, S., Brandeis, D., Caye, A., Christmann, N., Cortese, S., Coghill, D., Faraone, S. V., Franke, B., Gleitz, M., Greven, C. U., Kooij, S., Leffa, D. T., Rommelse, N., Newcorn, J. H., Polanczyk, G. V., Rohde, L. A., Simonoff, E., Stein, M., Vitiello, B., Yazgan, Y., Roesler, M., Doepfner, M. & Banaschewski, T. (2022). Toward Precision Medicine in ADHD. *Frontiers in Behavioral Neuroscience, 16*, 900981. https://doi.org/10.3389/fnbeh.2022.900981

Bymaster, F. P., Katner, J. S., Nelson, D. L., Hemrick-Luecke, S. K., Threlkeld, P. G., Heiligenstein, J. H., Morin, S. M., Gehlert, D. R. & Perry, K. W. (2002). Atomoxetine Increases Extracellular Levels of Norepinephrine and Dopamine in Prefrontal Cortex of Rat: A Potential Mechanism for Efficacy in Attention Deficit/Hyperactivity Disorder. *Neuropsychopharmacology, 27*(5), 699–711. https://doi.org/10.1016/s0893-133x(02)00346-9

Cai, N., Choi, K. W. & Fried, E. I. (2020). Reviewing the genetics of heterogeneity in depression: Operationalizations, manifestations, and etiologies. *Human Molecular Genetics, 29*(R1), ddaa115-. https://doi.org/10.1093/hmg/ddaa115

Cao, X., Deng, C., Su, X. & Guo, Y. (2018). Response and Remission Rates Following High-Frequency vs. Low-Frequency Repetitive Transcranial Magnetic Stimulation (rTMS) Over Right DLPFC for Treating Major Depressive Disorder (MDD): A Meta-Analysis of Randomized, Double-Blind Trials. *Frontiers in Psychiatry, 9*, 413. https://doi.org/10.3389/fpsyt.2018.00413

Carpenter, L. L., Janicak, P. G., Aaronson, S. T., Boyadjis, T., Brock, D. G., Cook, I. A., Dunner, D. L., Lanocha, K., Solvason, H. B. & Demitrack, M. A. (2012).Transcranial magnetic stimulation (TMS) for major depression: a multisite, naturalistic, observational study of acute treatment outcomes in clinical practice.*Depression and Anxiety, 29*(7), 587–596. https://doi.org/10.1002/da.21969

Carvalho, A. F., Solmi, M., Sanches, M., Machado, M. O., Stubbs, B., Ajnakina, O., Sherman, C., Sun, Y. R., Liu, C. S., Brunoni, A. R., Pigato, G., Fernandes, B. S., Bortolato, B., Husain, M. I., Dragioti, E., Firth, J., Cosco, T. D., Maes, M., Berk, M., Lanctot, K.L., Vieta, E., Pizzagalli, D.A., Smith, L., Fusar-Poli, P., Kurdyak, P.A., Fornaro, M., Rehm, J. & Herrmann, N. (2020). Evidence-based umbrella review of 162 peripheral biomarkers for major mental disorders. *Translational Psychiatry, 10*(1), 152. https://doi.org/10.1038/s41398-020-0835-5

Cash, R. F. H., Cocchi, L., Lv, J., Fitzgerald, P. B. & Zalesky, A. (2020). Functional Magnetic Resonance Imaging–Guided Personalization of Transcranial Magnetic Stimulation Treatment for Depression. *JAMA Psychiatry*. https://doi.org/10.1001/jamapsychiatry.2020.3794

Cash, R. F. H., Cocchi, L., Lv, J., Wu, Y., Fitzgerald, P. B. & Zalesky, A. (2021). Personalized connectivity-guided DLPFC-TMS for depression: Advancing computational feasibility, precision and reproducibility. *Human Brain Mapping, 42*(13), 4155–4172. https://doi.org/10.1002/hbm.25330

Cash, R. F. H., Müller, V. I., Fitzgerald, P. B., Eickhoff, S. B. & Zalesky, A. (2023). Altered brain activity in unipolar depression unveiled using connectomics. *Nature Mental Health*, 174–185. https://doi.org/10.1038/s44220-023-00038-8

Cash, R. F. H., Weigand, A., Zalesky, A., Siddiqi, S. H., Downar, J., Fitzgerald, P. B. & Fox, M. D. (2021). Using Brain Imaging to Improve Spatial Targeting of Transcranial

Magnetic Stimulation for Depression. *Biological Psychiatry, 90*(10), 689–700. https://doi.org/10.1016/j.biopsych.2020.05.033

Cash, R. F. H., Zalesky, A., Thomson, R. H., Tian, Y., Cocchi, L. & Fitzgerald, P. B. (2019). Subgenual Functional Connectivity Predicts Antidepressant Treatment Response to Transcranial Magnetic Stimulation: Independent Validation and Evaluation of Personalization. *Biological Psychiatry, 86*(2), e5–e7. https://doi.org/10.1016/j.biopsych.2018.12.002

Castrén, E., & Monteggia, L. M. (2021). Brain-Derived Neurotrophic Factor Signaling in Depression and Antidepressant Action. *Biological psychiatry, 90*(2), 128–136. https://doi.org/10.1016/j.biopsych.2021.05.008

Challman, T. D. & Lipsky, J. J. (2000). Methylphenidate: Its Pharmacology and Uses. *Mayo Clinic Proceedings, 75*(7), 711–721. https://doi.org/10.4065/75.7.711

Chen, J., Liu, Z., Zhu, D., Li, Q., Zhang, H., Huang, H., Wei, Y., Mu, J., Yang, D. & Xie, P. (2014). Bilateral vs. unilateral repetitive transcranial magnetic stimulation in treating major depression: A meta-analysis of randomized controlled trials. *Psychiatry Research, 219*(1), 51–57. https://doi.org/10.1016/j.psychres.2014.05.010

Chen, L., Klooster, D. C. W., Tik, M., Thomas, E. H. X., Downar, J., Fitzgerald, P. B., Williams, N. R. & Baeken, C. (2023). Accelerated Repetitive Transcranial Magnetic Stimulation to Treat Major Depression: The Past, Present, and Future. *Harvard Review of Psychiatry, 31*(3), 142–161. https://doi.org/10.1097/hrp.0000000000000364

Clark, C. R., Veltmeyer, M. D., Hamilton, R. J., Simms, E., Paul, R., Hermens, D. & Gordon, E. (2004). Spontaneous alpha peak frequency predicts working memory performance across the age span. *International Journal of Psychophysiology, 53*(1), 1–9. https://doi.org/10.1016/j.ijpsycho.2003.12.011

Cohen, S. L., Bikson, M., Badran, B. W. & George, M. S. (2022). A visual and narrative timeline of US FDA milestones for Transcranial Magnetic Stimulation (TMS) devices. *Brain Stimulation, 15*(1), 73–75. https://doi.org/10.1016/j.brs.2021.11.010

Cole, E. J., Phillips, A. L., Bentzley, B. S., Stimpson, K. H., Nejad, R., Barmak, F., Veerapal, C., Khan, N., Cherian, K., Felber, E., Brown, R., Choi, E., King, S., Pankow, H., Bishop, J. H., Azeez, A., Coetzee, J., Rapier, R., Odenwald, N., Carreon, D., Hawkins, J., Chang, M., Jennifer Keller, J., Raj, K., DeBattista, C., Jo, B., Espil, F.M., Schatzberg, A.F., Sudheimer, K.D. & Williams, N. R. (2021). Stanford Neuromodulation Therapy (SNT): A Double-Blind Randomized Controlled Trial. *American Journal of Psychiatry*, appiajp202120101429. https://doi.org/10.1176/appi.ajp.2021.20101429

Corlier, J., Carpenter, L. L., Wilson, A. C., Tirrell, E., Gobin, A. P., Kavanaugh, B. & Leuchter, A. F. (2019). The relationship between individual alpha peak frequency and clinical outcome with repetitive Transcranial Magnetic Stimulation (rTMS) treatment of Major Depressive Disorder (MDD). *Brain Stimulation*. https://doi.org/10.1016/j.brs.2019.07.018

Cortese, S., Adamo, N., Giovane, C. D., Mohr-Jensen, C., Hayes, A. J., Carucci, S., Atkinson, L. Z., Tessari, L., Banaschewski, T., Coghill, D., Hollis, C., Simonoff, E., Zuddas, A., Barbui, C., Purgato, M., Steinhausen, H.-C., Shokraneh, F., Xia, J. & Cipriani, A. (2018). Comparative efficacy and tolerability of medications for attention-deficit hyperactivity disorder in children, adolescents, and adults: a systematic review and network meta-analysis. *The Lancet Psychiatry, 5*(9), 727–738. https://doi.org/10.1016/s2215-0366(18)30269-4

Cortese, S., Ferrin, M., Brandeis, D., Holtmann, M., Aggensteiner, P., Daley, D., Santosh, P., Simonoff, E., Stevenson, J., Stringaris, A., Sonuga-Barke, E. J. S., European ADHD

Guidelines Group (EAGG). (2016). Neurofeedback for Attention-Deficit/Hyperactivity Disorder: Meta-Analysis of Clinical and Neuropsychological Outcomes From Randomized Controlled Trials. *Journal of the American Academy of Child & Adolescent Psychiatry, 55*(6), 444–455. https://doi.org/10.1016/j.jaac.2016.03.007

Cotovio, G., Ventura, F., Silva, D. R. da, Pereira, P. & Oliveira-Maia, A. J. (2023). Regulatory Clearance and Approval of Therapeutic Protocols of Transcranial Magnetic Stimulation for Psychiatric Disorders. *Brain Sciences, 13*(7), 1029. https://doi.org/10.3390/brainsci13071029

D'Elia, G. (1974). Unilateral electroconvulsive therapy. In M. Fink (Ed.), *Psychobiology of Convulsive Therapy*. VH Winston.

Deng, K., Li, H. & Guan, Y. (2020). Treatment Stratification of Patients with Metastatic Castration-Resistant Prostate Cancer by Machine Learning. *Science, 23*(2), 100804. https://doi.org/10.1016/j.isci.2019.100804

DeRubeis, R. J. & Strunk, D. R. (Eds.). (2017). *The Oxford Handbook of Mood Disorders.* https://doi.org/10.1093/oxfordhb/9780199973965.001.0001

Dickinson, A., DiStefano, C., Senturk, D. & Jeste, S. S. (2018). Peak alpha frequency is a neural marker of cognitive function across the autism spectrum. *European Journal of Neuroscience, 47*(6), 643–651. https://doi.org/10.1111/ejn.13645

van Dijk, H., van Wingen, G., Denys, D., Olbrich, S., van Ruth, R. & Arns, M. (2022). The two decades brainclinics research archive for insights in neurophysiology (TD BRAIN) database. *Scientific Data, 9*(1), 333. https://doi.org/10.1038/s41597-022-01409-z

Dinga, R., Schmaal, L. & Marquand, A. F. (2020). A Closer Look at Depression Biotypes: Correspondence Relating to Grosenick et al. (2019). *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 5*(5), 554–555. https://doi.org/10.1016/j.bpsc.2019.09.011

Dinga, R., Schmaal, L., Penninx, B., van Tol, M. J., Veltman, D. J., van Velzen, L., Mennes, M., van der Wee, N. & Marquand, A. (2019). Evaluating the evidence for biotypes of depression: Methodological replication and extension of Drysdale et al. (2017). *NeuroImage: Clinical, 22*(J. Consult. Clin. Psychol. 56 1988), 101796. https://doi.org/10.1016/j.nicl.2019.101796

van Dinteren, R., Arns, M., Jongsma, M. L. A. & Kessels, R. P. C. (2014). Combined frontal and parietal P300 amplitudes indicate compensated cognitive processing across the lifespan. *Frontiers in Aging Neuroscience, 6*, 294. https://doi.org/10.3389/fnagi.2014.00294

van Dinteren, R., Arns, M., Kenemans, L., Jongsma, M. L. A., Kessels, R. P. C., Fitzgerald, P., Fallahpour, K., Debattista, C., Gordon, E. & Williams, L. M. (2015). Utility of event-related potentials in predicting antidepressant treatment response: an iSPOT-D report. *European Neuropsychopharmacology, 25*(11), 1981–1990. https://doi.org/10.1016/j.euroneuro.2015.07.022

Donse, L., Padberg, F., Sack, A. T., Rush, A. J., & Arns, M. (2018). Simultaneous rTMS and psychotherapy in major depressive disorder: Clinical outcomes and predictors from a large naturalistic study. *Brain stimulation, 11*(2), 337–345. https://doi.org/10.1016/j.brs.2017.11.004

Doren, J. V., Arns, M., Heinrich, H., Vollebregt, M. A., Strehl, U. & Loo, S. K. (2018). Sustained effects of neurofeedback in ADHD: a systematic review and meta-analysis. *European Child & Adolescent Psychiatry, 28*(3), 293–305. https://doi.org/10.1007/s00787-018-1121-4

Downar, J. & Daskalakis, Z. J. (2013). New Targets for rTMS in Depression: A Review of Convergent Evidence. *Brain Stimulation, 6*(3), 231–240. https://doi.org/10.1016/j.brs.2012.08.006

Downar, J., Geraci, J., Salomons, T. V., Dunlop, K., Wheeler, S., McAndrews, M. P., Bakker, N., Blumberger, D. M., Daskalakis, Z. J., Kennedy, S. H., Flint, A. J. & Giacobbe, P. (2014). Anhedonia and Reward-Circuit Connectivity Distinguish Nonresponders from Responders to Dorsomedial Prefrontal Repetitive Transcranial Magnetic Stimulation in Major Depression. *Biological Psychiatry, 76*(3), 176–185. https://doi.org/10.1016/j.biopsych.2013.10.026

Drysdale, A. T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., Fetcho, R. N., Zebley, B., Oathes, D. J., Etkin, A., Schatzberg, A. F., Sudheimer, K., Keller, J., Mayberg, H. S., Gunning, F. M., Alexopoulos, G. S., Fox, M. D., Pascual-Leone, A., Voss, H. U., Casey,B.J., Dubon, M.J. & Liston, C. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nature Medicine, 23*(1), 28–38. https://doi.org/10.1038/nm.4246

Elbau, I. G., Lynch, C. J., Downar, J., Vila-Rodriguez, F., Power, J. D., Solomonov, N., Daskalakis, Z. J., Blumberger, D. M. & Liston, C. (2023a). Functional Connectivity Mapping for rTMS Target Selection in Depression. *American Journal of Psychiatry, 180*(3), 230–240. https://doi.org/10.1176/appi.ajp.20220306

Erguzel, T. T., Ozekes, S., Gultekin, S., Tarhan, N., Sayar, G. H. & Bayram, A. (2014). Neural Network Based Response Prediction of rTMS in Major Depressive Disorder Using QEEG Cordance. *Psychiatry Investigation, 12*(1), 61–65. https://doi.org/10.4306/pi.2015.12.1.61

Erguzel, T. T. & Tarhan, N. (2016). Machine Learning Approaches to Predict Repetitive Transcranial Magnetic Stimulation Treatment Response in Major Depressive Dis order. *SAI Intelligent Systems Conference*, 391–401. https://doi.org/10.1007/978-3-319-56991-8_29

Faraone, S. V., Biederman, J., Mick, E., Williamson, S., Wilens, T., Spencer, T., Weber, W., Jetton, J., Kraus, I., Pert, J. & Zallen, B. (2000). Family Study of Girls With Attention Deficit Hyperactivity Disorder. *American Journal of Psychiatry, 157*(7), 1077–1083. https://doi.org/10.1176/appi.ajp.157.7.1077

Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods, 39*(2), 175–191. https://doi.org/10.3758/bf03193146

Fernandes, B. S., Williams, L. M., Steiner, J., Leboyer, M., Carvalho, A. F. & Berk, M. (2017). The new field of 'precision psychiatry.' *BMC Medicine, 15*(1), 80. https://doi.org/10.1186/s12916-017-0849-x

Filipčić, I., Filipčić, I. Š., Milovac, Ž., Sučić, S., Gajšak, T., Ivezić, E., Bašić, S., Bajić, Ž. & Heilig, M. (2019). Efficacy of repetitive transcranial magnetic stimulation using a figure-8-coil or an H1-Coil in treatment of major depressive disorder; A randomized clinical trial. *Journal of Psychiatric Research, 114*, 113–119. https://doi.org/10.1016/j.jpsychires.2019.04.020

Fils, J. M., Penick, E. C., Nickel, E. J., Othmer, E., DeSouza, C., Gabrielli, W. F. & Hunter, E. E. (2010). Minor Versus Major Depression: A Comparative Clinical Study. *The Primary Care Companion to The Journal of Clinical Psychiatry, 12*(1), PCC.08m00752. https://doi.org/10.4088/pcc.08m00752blu

Fitzgerald, P. B., Fountain, S. & Daskalakis, Z. J. (2006). A comprehensive review of the effects of rTMS on motor cortical excitability and inhibition. *Clinical Neurophysi-*

ology, 117(12), 2584–2596. https://doi.org/10.1016/j.clinph.2006.06.712
Fitzgerald, P. B., George, M. S. & Pridmore, S. (2022). The evidence is in: Repetitive transcranial magnetic stimulation is an effective, safe and well-tolerated treatment for patients with major depressive disorder. *Australian & New Zealand Journal of Psychiatry, 56*(7), 745–751. https://doi.org/10.1177/00048674211043047
Fitzgerald, P. B., Hoy, K. E., Reynolds, J., Singh, A., Gunewardene, R., Slack, C., Ibrahim, S. & Daskalakis, Z. J. (2019). A pragmatic randomized controlled trial exploring the relationship between pulse number and response to repetitive transcranial magnetic stimulation treatment in depression. *Brain Stimulation.* https://doi.org/10.1016/j.brs.2019.09.001
Fogel, J., Eaton, W. W. & Ford, D. E. (2006). Minor depression as a predictor of the first onset of major depressive disorder over a 15-year follow-up. *Acta Psychiatrica Scandinavica, 113*(1), 36–43. https://doi.org/10.1111/j.1600-0447.2005.00654.x
Fox, H. & Sinha, R. (2014). Chapter Six The Role of Guanfacine as a Therapeutic Agent to Address Stress-Related Pathophysiology in Cocaine-Dependent Individuals. *Advances in Pharmacology, 69*, 217–265. https://doi.org/10.1016/b978-0-12-420118-7.00006-8
Fox, M. D., Buckner, R. L., White, M. P., Greicius, M. D. & Pascual-Leone, A. (2012). Efficacy of transcranial magnetic stimulation targets for depression is related to intrinsic functional connectivity with the subgenual cingulate. *Biological Psychiatry, 72*(7), 595–603. https://doi.org/10.1016/j.biopsych.2012.04.028
Garcés, P., Vicente, R., Wibral, M., Pineda-Pardo, J. Á., López, M. E., Aurtenetxe, S., Marcos, A., Andrés, M. E. de, Yus, M., Sancho, M., Maestú, F. & Fernández, A. (2013). Brain-wide slowing of spontaneous alpha rhythms in mild cognitive impairment. *Frontiers in Aging Neuroscience, 5*, 100–100. https://doi.org/10.3389/fnagi.2013.00100
Gaynes, B. N., Lux, L., Gartlehner, G., Asher, G., Forman-Hoffman, V., Green, J., Boland, E., Weber, R. P., Randolph, C., Bann, C., Coker-Schwimmer, E., Viswanathan, M. & Lohr, K. N. (2020). Defining treatment-resistant depression. *Depression and Anxiety, 37*(2), 134–145. https://doi.org/10.1002/da.22968
Ge, R., Downar, J., Blumberger, D. M., Daskalakis, Z. J. & Vila-Rodriguez, F. (2020). Functional connectivity of the anterior cingulate cortex predicts treatment outcome for rTMS in treatment-resistant depression at 3-month follow-up. *Brain Stimulation, 13*(1). https://doi.org/10.1016/j.brs.2019.10.012
George, M. S., Huffman, S., Doose, J., Sun, X., Dancy, M., Faller, J., Li, X., Yuan, H., Goldman, R. I., Sajda, P. & Brown, T. R. (2023). EEG synchronized left prefrontal transcranial magnetic stimulation (TMS) for treatment resistant depression is feasible and produces an entrainment dependent clinical response: A randomized controlled double blind clinical trial. *Brain Stimulation, 16*(6), 1753–1763. https://doi.org/10.1016/j.brs.2023.11.010
George, M. S., Wassermann, E. M., Williams, W. A., Callahan, A., Ketter, T. A., Basser, P., Hallett, M. & Post, R. M. (1995). Daily repetitive transcranial magnetic stimulation (rTMS) improves mood in depression. *NeuroReport, 6*(14), 1853–1856. https://doi.org/10.1097/00001756-199510020-00008
Ghio, L., Gotelli, S., Marcenaro, M., Amore, M. & Natta, W. (2014). Duration of untreated illness and outcomes in unipolar depression: A systematic review and meta-analysis. J*ournal of Affective Disorders, 152*, 45–51. https://doi.org/10.1016/j.jad.2013.10.002

Glover, G. H. (2011). Overview of Functional Magnetic Resonance Imaging. *Neurosurgery Clinics of North America, 22*(2), 133–139. https://doi.org/10.1016/j.nec.2010.11.001

Golan, N., Shahar, E., Ravid, S. & Pillar, G. (2004). Sleep Disorders and Daytime Sleepiness in Children with Attention-Deficit/ Hyperactive Disorder. *Sleep, 27*(2), 261–266. https://doi.org/10.1093/sleep/27.2.261

Goldberg, D. (2011). The heterogeneity of "major depression." *World Psychiatry, 10*(3), 226–228. https://doi.org/10.1002/j.2051-5545.2011.tb00061.x

Goldstein-Piekarski, A. N., Holt-Gosselin, B., O'Hora, K. & Williams, L. M. (2020). Integrating sleep, neuroimaging, and computational approaches for precision psychiatry. *Neuropsychopharmacology, 45*(1), 192–204. https://doi.org/10.1038/s41386-019-0483-8

Gramfort, A., Luessi, M., Larson, E., Engemann, D. A., Strohmeier, D., Brodbeck, C., Goj, R., Jas, M., Brooks, T., Parkkonen, L. & Hämäläinen, M. (2013). MEG and EEG data analysis with MNE-Python. *Frontiers in Neuroscience, 7*, 267. https://doi.org/10.3389/fnins.2013.00267

Grandy, T. H., Werkle-Bergner, M., Chicherio, C., Lövdén, M., Schmiedek, F. & Lindenberger, U. (2013). Individual alpha peak frequency is related to latent factors of general cognitive abilities. *NeuroImage, 79*, 10–18. https://doi.org/10.1016/j.neuroimage.2013.04.059

Gratton, G., Cole, M. G. H., & Donchin, E. (1983). A New Method for Off-Line Removal of Ocular Artifact. *Electroencephalography and Clinical Neurophysiology 55* (4): 468–84. https://doi.org/10.1016/0013-4694(83)90135-9.

Griffiths, K. R., Leikauf, J. E., Tsang, T. W., Clarke, S., Hermens, D. F., Efron, D., Williams, L. M. & Kohn, M. R. (2018). Response inhibition and emotional cognition improved by atomoxetine in children and adolescents with ADHD: The ACTION randomized controlled trial. *Journal of Psychiatric Research, 102*(Depress. Anxiety 26 3 2009), 57–64. https://doi.org/10.1016/j.jpsychires.2018.03.009

Grosenick, L. & Liston, C. (2020). Reply to: A Closer Look at Depression Biotypes: Correspondence Relating to Grosenick et al. (2019). *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging, 5*(5), 556. https://doi.org/10.1016/j.bpsc.2019.11.002

Grover, V. P. B., Tognarelli, J. M., Crossey, M. M. E., Cox, I. J., Taylor-Robinson, S. D. & McPhail, M. J. W. (2015). Magnetic Resonance Imaging: Principles and Techniques: Lessons for Clinicians. *Journal of Clinical and Experimental Hepatology, 5*(3), 246–255. https://doi.org/10.1016/j.jceh.2015.08.001

Grzenda, A. & Widge, A. S. (2023). Electronic health records and stratified psychiatry: bridge to precision treatment? *Neuropsychopharmacology*, 1–6. https://doi.org/10.1038/s41386-023-01724-y

Harris, C. R., Millman, K. J., van der Walt, S. J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M.H., Brett, M., Haldane, A., Fernández del Río, J., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C. & Oliphant, T. E.(2020). Array programming with NumPy. *Nature, 585*, 357–362.

Hartwigsen, G. & Silvanto, J. (2023). Noninvasive Brain Stimulation: Mul_ tiple Effects on Cognition. *The Neuroscientist, 29*(5), 639–653. https://doi.org/10.1177/10738584221113806

Hasanzadeh, F., Mohebbi, M. & Rostami, R. (2019). Prediction of rTMS treatment response in major depressive disorder using machine learning techniques and non-

linear features of EEG signal. *Journal of Affective Disorders, 256*, 132–142. https://doi.org/10.1016/j.jad.2019.05.070

Hawton, K., Comabella, C. C. i, Haw, C. & Saunders, K. (2013). Risk factors for suicide in individuals with depression: A systematic review. *Journal of Affective Disorders, 147*(1–3), 17–28. https://doi.org/10.1016/j.jad.2013.01.004

Henriques, J. B. & Davidson, R. J. (1991). Left frontal hypoactivation in depression. *Journal of Abnormal Psychology, 100*(4), 535–545. https://doi.org/10.1037/0021-843x.100.4.535

Hermens, D. F., Kohn, M. R., Clarke, S. D., Gordon, E. & Williams, L. M. (2005). Sex differences in adolescent ADHD: findings from concurrent EEG and EDA. *The Meridian Clinic*. https://doi.org/10.1016/j.clinph.2005.02.012

Herrman, H., Patel, V., Kieling, C., Berk, M., Buchweitz, C., Cuijpers, P., Furukawa, T. A., Kessler, R. C., Kohrt, B. A., Maj, M., McGorry, P., Reynolds, C. F., Weissman, M. M., Chibanda, D., Dowrick, C., Howard, L. M., Hoven, C. W., Knapp, M., Mayberg, H. S., Penninx, B.W.J.H., Xiao, S., Trivedi, M., Uher, R., Vijayakumar, L. & Wolpert, M. (2022). Time for united action on depression: a Lancet–World Psychiatric Association Commission. *The Lancet, 399*(J Gerontol B Psychol Sci Soc Sci 70 2015), 957–1022. https://doi.org/10.1016/s0140-6736(21)02141-3

Ho, S. Y., Phua, K., Wong, L. & Goh, W. W. B. (2020). Extensions of the External Validation for Checking Learned Model Interpretability and Generalizability. *Patterns, 1*(8), 100129. https://doi.org/10.1016/j.patter.2020.100129

Höflich, G., Kasper, S., Hufnagel, A., Ruhrmann, S. & Möller, H. J. (1993). Application of transcranial magnetic stimulation in treatment of drug-resistant major depres sion - a report of two cases. *Human Psychopharmacology: Clinical and Experimental, 8*(5), 361–365. https://doi.org/10.1002/hup.470080510

Hopman, H. J., Chan, S. M. S., Chu, W. C. W., Lu, H., Tse, C.-Y., Chau, S. W. H., Lam, L. C. W., Mak, A. D. P. & Neggers, S. F. W. (2021). Personalized prediction of transcranial magnetic stimulation clinical response in patients with treatment-refractory depression using neuroimaging biomarkers and machine learning. *Journal of Affective Disorders, 290*, 261–271. https://doi.org/10.1016/j.jad.2021.04.081

Huang, W. A., Stitt, I. M., Negahbani, E., Passey, D. J., Ahn, S., Davey, M., Dannhauer, M., Doan, T. T., Hoover, A. C., Peterchev, A. V., Radtke-Schuller, S., & Fröhlich, F. (2021). Transcranial Alternating Current Stimulation Entrains Alpha Oscillations by Preferential Phase Synchronization of Fast-Spiking Cortical Neurons to Stimulation Waveform. *Nature Communications 12* (1): 3151. https://doi.org/10.1038/s41467-021-23021-2.

Huerta, P. T. & Volpe, B. T. (2009). Transcranial magnetic stimulation, synaptic plasticity and network oscillations. *Journal of NeuroEngineering and Rehabilitation, 6*(1), 7. https://doi.org/10.1186/1743-0003-6-7

Hunter, A. M., Nghiem, T. X., Cook, I. A., Krantz, D. E., Minzenberg, M. J. & Leuchter, A. F. (2017). Change in Quantitative EEG Theta Cordance as a Potential Predictor of Repetitive Transcranial Magnetic Stimulation Clinical Outcome in Major Depressive Disorder. *Clinical EEG and Neuroscience, 49*(5), 306–315. https://doi.org/10.1177/1550059417746212

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science & Engineering, 9*(3), 90--95. https://doi.org/10.1109/mcse.2007.55

Iseger, T. A., Korgaonkar, M. S., Kenemans, J. L., Grieve, S. M., Baeken, C., Fitzgerald, P. - B. & Arns, M. (2017). EEG connectivity between the subgenual anterior cingulate

and prefrontal cortices in response to antidepressant medication. *European Neuropsychopharmacology, 27*(4), 301–312. https://doi.org/10.1016/j.euronero.2017.02.002
Jahanshahi, M. & Rothwell, J. (2000). Transcranial magnetic stimulation studies of - cognition: an emerging field. *Experimental Brain Research, 131*(1), 1–9. https://doi.org/10.1007/s002219900224
Jannati, A., Oberman, L. M., Rotenberg, A. & Pascual-Leone, A. (2023). Assessing the mechanisms of brain plasticity by transcranial magnetic stimulation. *Neuropsychopharmacology, 48*(1), 191–208. https://doi.org/10.1038/s41386-022-01453-8
Jin, Y. & Phillips, B. (2014). A pilot study of the use of EEG-based synchronized Transcranial Magnetic Stimulation (sTMS) for treatment of Major Depression. *BMC Psychiatry, 14*(1), 13. https://doi.org/10.1186/1471-244x-14-13
Kaster, T. S., Downar, J., Vila-Rodriguez, F., Thorpe, K. E., Feffer, K., Noda, Y., Giacobbe, P., Knyahnytska, Y., Kennedy, S. H., Lam, R. W., Daskalakis, Z. J. & Blumberger, D. M. (2019). Trajectories of Response to Dorsolateral Prefrontal rTMS in Major Depression: A THREE-D Study. *American Journal of Psychiatry, 176*(5), 367–375. https://doi.org/10.1176/appi.ajp.2018.18091096
Kato, M. & Manabe, A. (2018). Treatment and biology of pediatric acute lymphoblastic l eukemia. *Pediatrics International, 60*(1), 4–12. https://doi.org/10.1111/ped.13457
Kedzior, K. K., Azorina, V. & Reitz, S. K. (2014). More female patients and fewer stimuli per session are associated with the short-term antidepressant properties of repetitive transcranial magnetic stimulation (rTMS): a meta-analysis of 54 sham-controlled studies published between 1997–2013. *Neuropsychiatric Disease and Treatment, 10*, 727–756. https://doi.org/10.2147/ndt.s58405
Keizer, A. W. (2019). Standardization and Personalized Medicine Using Quantitative EEG in Clinical Settings. *Clinical EEG and Neuroscience*, 1550059419874945. https://doi.org/10.1177/1550059419874945
Kirsten, A., Seifritz, E. & Olbrich, S. (2019). Electroencephalogram Source Connectivity in the Prediction of Electroconvulsive Therapy Outcome in Major Depressive Disorder. *Clinical EEG and Neuroscience, 51*(1), 10–18. https://doi.org/10.1177/1550059419888338
Kishi, T., Sakuma, K., Matsuda, Y., Kito, S. & Iwata, N. (2023). Intermittent theta burst stimulation vs. high-frequency repetitive transcranial magnetic stimulation for major depressive disorder: A systematic review and meta-analysis. *Psychiatry Research, 328*, 115452. https://doi.org/10.1016/j.psychres.2023.115452
Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Research Reviews, 29*(2–3), 169–195. https://doi.org/10.1016/s0165-0173(98)00056-3
Kondacs, A. & Szabó, M. (1999). Long-term intra-individual variability of the background EEG in normals. *Clinical Neurophysiology, 110*(10), 1708–1716. https://doi.org/10.1016/s1388-2457(99)00122-4
Koslow, S. H., Williams, L. M. & Gordon, E. (2010). Personalized medicine for the brain: a call for action. *Molecular Psychiatry, 15*(3), 229–230. https://doi.org/10.1038/mp.2009.147
Krepel, N., Egtberts, T., Sack, A. T., Heinrich, H., Ryan, M. & Arns, M. (2020). A multicenter effectiveness trial of QEEG-informed neurofeedback in ADHD: Replication and treatment prediction. *NeuroImage: Clinical,* 102399. https://doi.org/10.1016/j.nicl.2020.102399
Krepel, N., Rush, A. J., Iseger, T. A., Sack, A. T. & Arns, M. (2019). Can psychological

features predict antidepressant response to rTMS? A Discovery–Replication approach. P*sychological Medicine, 50*(2), 264–272. https://doi.org/10.1017/s0033291718004191

Krepel, N., Sack, A. T., Kenemans, J. L., Fitzgerald, P. B., Drinkenburg, W. H. & Arns, M. (2018). Non-replication of neurophysiological predictors of non-response to rTMS in depression and neurophysiological data-sharing proposal. *Brain Stimulation, 11*(3), 639–641. https://doi.org/10.1016/j.brs.2018.01.032

Kupfer, D. J., Frank, E. & Phillips, M. L. (2012). Major depressive disorder: new clinical, neurobiological, and treatment perspectives. *The Lancet, 379*(9820), 1045–1055. https://doi.org/10.1016/s0140-6736(11)60602-8

Lam, R. W., Milev, R., Rotzinger, S., Andreazza, A. C., Blier, P., Brenner, C., Daskalakis, Z. J., Dharsee, M., Downar, J., Evans, K. R., Farzan, F., Foster, J. A., Frey, B. N., Geraci, J., Giacobbe, P., Feilotter, H. E., Hall, G. B., Harkness, K. L., Hassel, S.,CAN-BIND Investigator Team. (2016). Discovering biomarkers for antidepressant response: protocol from the Canadian biomarker integration network in depression (CAN-BIND) and clinical characteristics of the first patient cohort. *BMC Psychiatry, 16*(1), 105. https://doi.org/10.1186/s12888-016-0785-x

Lefaucheur, J.-P., Aleman, A., Baeken, C., Benninger, D. H., Brunelin, J., Lazzaro, V. D., Filipović, S. R., Grefkes, C., Hasan, A., Hummel, F. C., Jääskeläinen, S. K., Langguth, B., Leocani, L., Londero, A., Nardone, R., Nguyen, J.-P., Nyffeler, T., Oliveira-Maia, A. J., Oliviero, A., Padberg, F., Palm, U., Paulus, W., Poulet, E., Quartarone, A., Rachid, F., Rektorová, I., Rossi, S., Sahlsten, H., Schecklmann, M., David Szekely, D. & Ziemann, U. (2020). Evidence-based guidelines on the therapeutic use of repetitive transcranial magnetic stimulation (rTMS): An update (2014–2018). *Clinical Neurophysiology*. https://doi.org/10.1016/j.clinph.2019.11.002

Leggett, L. E., Soril, L. J. J., Coward, S., Lorenzetti, D. L., MacKean, G. & Clement, F. M. (2015). Repetitive Transcranial Magnetic Stimulation for Treatment-Resistant Depression in Adult and Youth Populations. *The Primary Care Companion For CNS Disorders.* https://doi.org/10.4088/pcc.15r01807

Leikauf, J. E., Griffiths, K. R., Saggar, M., Hong, D. S., Clarke, S., Efron, D., Tsang, T. W., Hermens, D. F., Kohn, M. R. & Williams, L. M. (2017). Identification of biotypes in Attention-Deficit/Hyperactivity Disorder, a report from a randomized, controlled trial. *Personalized Medicine in Psychiatry, 3*, 8–17. https://doi.org/10.1016/j.pmip.2017.02.001

van der Lem, R., van der Wee, N. J. A., van Veen, T. & Zitman, F. G. (2012). Efficacy versus Effectiveness: A Direct Comparison of the Outcome of Treatment for Mild to Moderate Depression in Randomized Controlled Trials and Daily Practice. *Psychotherapy and Psychosomatics, 81*(4), 226–234. https://doi.org/10.1159/000330890

Lépine, J.-P. & Briley, M. (2011). The increasing burden of depression. *Neuropsychiatric Disease and Treatment, 7*(Suppl 1), 3–7. https://doi.org/10.2147/ndt.s19617

Leuchter, A. F., Cook, I. A., Feifel, D., Goethe, J. W., Husain, M., Carpenter, L. L., Thase, M. E., Krystal, A. D., Philip, N. S., Bhati, M. T., Burke, W. J., Howland, R. H., Sheline, Y. I., Aaronson, S. T., Iosifescu, D. V., O'Reardon, J. P., Gilmer, W. S., Jain, R., Burgoyne, K. S., Phillips, B., Manberg, P.J., Massaro, J., Hunter, A.M., Lisanby, S.H. & George, M. S. (2015). Efficacy and Safety of Low-field Synchronized Transcranial Magnetic Stimulation (sTMS) for Treatment of Major Depression. *Brain Stimulation, 8*(4), 787–794. https://doi.org/10.1016/j.brs.2015.05.005

Leuchter, A. F., Cook, I. A., Jin, Y. & Phillips, B. (2013). The relationship between brain

oscillatory activity and therapeutic effectiveness of transcranial magnetic stimulation in the treatment of major depressive disorder. *Frontiers in Human Neuroscience, 7*, 37. https://doi.org/10.3389/fnhum.2013.00037

Leuchter, A. F., Cook, I. A., Lufkin, R. B., Dunkin, J., Newton, T. F., Cummings, J. L., Mackey, J. K. & Walter, D. O. (1994). Cordance: A New Method for Assessment of Cerebral Perfusion and Metabolism Using Quantitative Electroencephalography. *NeuroImage, 1*(3), 208–219. https://doi.org/10.1006/nimg.1994.1006

Levkovitz, Y., Isserles, M., Padberg, F., Lisanby, S. H., Bystritsky, A., Xia, G., Tendler, A., Daskalakis, Z. J., Winston, J. L., Dannon, P., Hafez, H. M., Reti, I. M., Morales, O. G., Schlaepfer, T. E., Hollander, E., Berman, J. A., Husain, M. M., Sofer, U., Stein, A., Adler, S., Deutsch, L., Deutsch, F., Roth, Y., George, M.S. & Zangen, A. (2015). Efficacy and safety of deep transcranial magnetic stimulation for major depression: a prospective multicenter randomized controlled trial. *World Psychiatry, 14*(1), 64–73. https://doi.org/10.1002/wps.20199

Li, C.-T., Cheng, C.-M., Juan, C.-H., Tsai, Y.-C., Chen, M.-H., Bai, Y.-M., Tsai, S.-J. & Su, T.-P. (2021). Task-Modulated Brain Activity Predicts Antidepressant Responses of Prefrontal Repetitive Transcranial Magnetic Stimulation: A Randomized Sham-Control Study. *Chronic Stress, 5*, 24705470211006856. https://doi.org/10.1177/24705470211006855

Li, C.-T., Hsieh, J.-C., Huang, H.-H., Chen, M.-H., Juan, C.-H., Tu, P.-C., Lee, Y.-C., Wang, S.-J., Cheng, C.-M. & Su, T.-P. (2016). Cognition-Modulated Frontal Activity in Prediction and Augmentation of Antidepressant Efficacy: A Randomized Controlled Pilot Study. *Cerebral Cortex, 26*(1), 202–210. https://doi.org/10.1093/cercor/bhu191

Lindsley, D. B. (1936). Brain potentials in children and adults. *Science, 84*(2181), 354–354. https://doi.org/10.1126/science.84.2181.354

Liston, C., Chen, A. C., Zebley, B. D., Drysdale, A. T., Gordon, R., Leuchter, B., Voss, H. U., Casey, B. J., Etkin, A. & Dubin, M. J. (2014). Default mode network mechanisms of transcranial magnetic stimulation in depression. *Biological Psychiatry, 76*(7), 517–526. https://doi.org/10.1016/j.biopsych.2014.01.023

Loo, S. K., Bilder, R. M., Cho, A. L., Sturm, A., Cowen, J., Walshaw, P., Levitt, J., Del'Homme, M., Piacentini, J., McGough, J. J. & McCracken, J. T. (2016). Effects of d-Methylphenidate, Guanfacine, and Their Combination on Electroencephalogram Resting State Spectral Power in Attention-Deficit/Hyperactivity Disorder. *Journal of the American Academy of Child & Adolescent Psychiatry, 55*(8), 674-682.e1. https://doi.org/10.1016/j.jaac.2016.04.020

van Luijtelaar, G., Verbraak, M., van den Bunt, M., Keijsers, G. & Arns, M. (2010). EEG Findings in Burnout Patients. *The Journal of Neuropsychiatry and Clinical Neurosciences, 22*(2), 208–217. https://doi.org/10.1176/jnp.2010.22.2.208

Luo, Y., Weibman, D., Halperin, J. M. & Li, X. (2019). A Review of Heterogeneity in Attention Deficit/Hyperactivity Disorder (ADHD). *Frontiers in Human Neuroscience, 13*, 42. https://doi.org/10.3389/fnhum.2019.00042

Lynch, C. J., Power, J. D., Scult, M. A., Dubin, M., Gunning, F. M. & Liston, C. (2020). Rapid Precision Functional Mapping of Individuals Using Multi-Echo fMRI. *Cell Reports, 33*(12), 108540. https://doi.org/10.1016/j.celrep.2020.108540

Malaspina, D., Devanand, D. P., Krueger, R. B., Prudic, J. & Sackeim, H. A. (1994). The significance of clinical EEG abnormalities in depressed patients treated with ECT. *Convulsive Therapy, 10*(4), 259–266.

Malhi, G. S. & Mann, J. J. (2018). Depression. *The Lancet, 392*(10161), 2299–2312. https://doi.org/10.1016/s0140-6736(18)31948-2

Marwaha, S., Palmer, E., Suppes, T., Cons, E., Young, A. H. & Upthegrove, R. (2023). Novel and emerging treatments for major depression. *The Lancet, 401*(10371), 141–153. https://doi.org/10.1016/s0140-6736(22)02080-3

Matsuda, Y., Yamazaki, R., Shigeta, M. & Kito, S. (2021). Transcranial magnetic stimulation modalities for psychiatric disorders: Publication trends from 1985 to 2019. *Neuropsychopharmacology Reports, 41*(4), 538–543. https://doi.org/10.1002/npr2.12212

McCracken, J. T., McGough, J. J., Loo, S. K., Levitt, J., Del'Homme, M., Cowen, J., Sturm, A., Whelan, F., Hellemann, G., Sugar, C. & Bilder, R. M. (2016). Combined Stimulant and Guanfacine Administration in Attention-Deficit/Hyperactivity Disorder: A Controlled, Comparative Study. *Journal of the American Academy of Child & Adolescent Psychiatry, 55*(8), 657-666.e1. https://doi.org/10.1016/j.jaac.2016.05.015

McDonald, W. M., Durkalski, V., Ball, E. R., Holtzheimer, P. E., Pavlicova, M., Lisanby, S. H., Avery, D., Anderson, B. S., Nahas, Z., Zarkowski, P., Sackeim, H. A. & George, M. S. (2011). Improving the antidepressant efficacy of transcranial magnetic stimulation: maximizing the number of stimulations and treatment location in treatment-resistant depression. *Depression and Anxiety, 28*(11), 973–980. https://doi.org/10.1002/da.20885

McLaughlin, K. A. (2011). The Public Health Impact of Major Depression: A Call for Interdisciplinary Prevention Efforts. *Prevention Science, 12*(4), 361–371. https://doi.org/10.1007/s11121-011-0231-8

McLean, A. L. (2019). Publication trends in transcranial magnetic stimulation: a 30-year panorama. *Brain Stimulation, 12*(3), 619–627. https://doi.org/10.1016/j.brs.2019.01.002

Meeuwissen, J. A. C., Feenstra, T. L., Smit, F., Blankers, M., Spijker, J., Bockting, C. L. H., van Balkom, A. J. L. M. & Buskens, E. (2019). The cost-utility of stepped-care algorithms according to depression guideline recommendations – Results of a state-transition model analysis. *Journal of Affective Disorders, 242*, 244–254. https://doi.org/10.1016/j.jad.2018.08.024

Meijs, H., Prentice, A., Lin, B. D., Wilde, B. D., Hecke, J. V., Niemegeers, P., van Eijk, K., Luykx, J. J. & Arns, M. (2022). A polygenic-informed approach to a predictive EEG signature empowers antidepressant treatment prediction: A proof-of-concept study. *European Neuropsychopharmacology, 62*, 49–60. https://doi.org/10.1016/j.euroneuro.2022.07.006

Meyer, T., Brunovsky, M., Horacek, J., Novak, T., Andrashko, V., Seifritz, E. & Olbrich, S. (2021). Predictive value of heart rate in treatment of major depression with ketamine in two controlled trials. *Clinical Neurophysiology, 132*(6), 1339–1346. https://doi.org/10.1016/j.clinph.2021.01.030

Michelini, G., Norman, L. J., Shaw, P. & Loo, S. K. (2022). Treatment biomarkers for ADHD: Taking stock and moving forward. *Translational Psychiatry, 12*(1), 444. https://doi.org/10.1038/s41398-022-02207-2

Milev, R. V., Giacobbe, P., Kennedy, S. H., Blumberger, D. M., Daskalakis, Z. J., Downar, J., Modirrousta, M., Patry, S., Vila-Rodriguez, F., Lam, R. W., MacQueen, G. M., Parikh, S. V., Ravindran, A. V. & the CANMAT Depression Work Group. (2016). Canadian Network for Mood and Anxiety Treatments (CANMAT) 2016 Clinical Guidelines for the Management of Adults with Major Depressive Disorder. *The Canadian Journal of Psychiatry, 61*(9), 561–575. https://doi.org/10.1177/0706743716660033

Moffitt, T. E., Caspi, A., Taylor, A., Kokaua, J., Milne, B. J., Polanczyk, G. & Poulton, R. (2010). How common are common mental disorders? Evidence that lifetime prevalence rates are doubled by prospective versus retrospective ascertainment. *Psychological Medicine, 40*(6), 899–909. https://doi.org/10.1017/s0033291709991036

Molina, B. S. G., Hinshaw, S. P., Swanson, J. M., Arnold, L. E., Vitiello, B., Jensen, P. S., Epstein, J. N., Hoza, B., Hechtman, L., Abikoff, H. B., Elliott, G. R., Greenhill, L. L., Newcorn, J. H., Wells, K. C., Wigal, T., Gibbons, R. D., Hur, K., Houck, P. R. & The MTA Cooperative Group (2009). The MTA at 8 Years. *Journal of the American Academy of Child & Adolescent Psychiatry, 48*(5), 484–500. https://doi.org/10.1097/chi.0b013e31819c23d0

Monastra, V. J., Monastra, D. M. & George, S. (2002). The Effects of Stimulant Therapy, EEG Biofeedback, and Parenting Style on the Primary Symptoms of Attention-Deficit/Hyperactivity Disorder. *Applied Psychophysiology and Biofeedback, 27*(4), 231–249. https://doi.org/10.1023/a:1021018700609

Moncrieff, J., Cooper, R. E., Stockmann, T., Amendola, S., Hengartner, M. P. & Horowitz, M. A. (2023). The serotonin theory of depression: a systematic umbrella review of the evidence. *Molecular Psychiatry, 28*(8), 3243–3256. https://doi.org/10.1038/s41380-022-01661-0

Morris, S. S. J., Musser, E. D., Tenenbaum, R. B., Ward, A. R., Martinez, J., Raiker, J. S., Coles, E. K. & Riopelle, C. (2020). Emotion Regulation via the Autonomic Nervous System in Children with Attention-Deficit/Hyperactivity Disorder (ADHD): Replication and Extension. *Journal of Abnormal Child Psychology, 48*(3), 361–373. https://doi.org/10.1007/s10802-019-00593-8

Mumtaz, W., Rasheed, S. & Irfan, A. (2021). Review of challenges associated with the EEG artifact removal methods. *Biomedical Signal Processing and Control, 68*, 102741. https://doi.org/10.1016/j.bspc.2021.102741

Murphy, M. & Öngür, D. (2019). Decreased peak alpha frequency and impaired visual evoked potentials in first episode psychosis. *NeuroImage: Clinical, 22*(Schizophr. Bull. 40 2014), 101693. https://doi.org/10.1016/j.nicl.2019.101693

Musser, E. D., Backs, R. W., Schmitt, C. F., Ablow, J. C., Measelle, J. R. & Nigg, J. T. (2011). Emotion Regulation via the Autonomic Nervous System in Children with Attention-Deficit/Hyperactivity Disorder (ADHD). *Journal of Abnormal Child Psychology, 39*(6), 841–852. https://doi.org/10.1007/s10802-011-9499-1

Newson, J. J. & Thiagarajan, T. C. (2019). EEG Frequency Bands in Psychiatric Disorders: A Review of Resting State Studies. Frontiers in Human Neuroscience, 12, 521. https://doi.org/10.3389/fnhum.2018.00521Nilsonne, G. & Harrell, F. E. (2020). EEG-based model and antidepressant response. *Nature Biotechnology,* 1–1. https://doi.org/10.1038/s41587-020-00768-5

Nir, R.-R., Sinai, A., Raz, E., Sprecher, E. & Yarnitsky, D. (2010). Pain assessment by continuous EEG: Association between subjective perception of tonic pain and peak frequency of alpha oscillations during stimulation and at rest. *Brain Research, 1344*, 77–86. https://doi.org/10.1016/j.brainres.2010.05.004

Olbrich, S. & Arns, M. (2013). EEG biomarkers in major depressive disorder: Discriminative power and prediction of treatment response. *International Review of Psychiatry, 25*(5), 604–618. https://doi.org/10.3109/09540261.2013.816269

Olbrich, S., van Dinteren, R. & Arns, M. (2015). Personalized Medicine: Review and Perspectives of Promising Baseline EEG Biomarkers in Major Depressive Disorder and Attention Deficit Hyperactivity Disorder. *Neuropsychobiology, 72*(3–4), 229–240.

https://doi.org/10.1159/000437435
O'Leary, D., Costello, F., Gormley, N. & Webb, M. (2000). Remission onset and relapse in depression An 18-month prospective study of course for 100 first admission patients. *Journal of Affective Disorders, 57*(1–3), 159–171. https://doi.org/10.1016/s0165-0327(99)00086-5
O'Reardon, J. P., Solvason, H. B., Janicak, P. G., Sampson, S., Isenberg, K. E., Nahas, Z., McDonald, W. M., Avery, D., Fitzgerald, P. B., Loo, C., Demitrack, M. A., George, M. S. & Sackeim, H. A. (2007). Efficacy and Safety of Transcranial Magnetic Stimulation in the Acute Treatment of Major Depression: A Multisite Randomized Controlled Trial. *Biological Psychiatry, 62*(11), 1208–1216. https://doi.org/10.1016/j.biopsych.2007.01.018
Orr, K. E. & McHugh, K. (2019). The new international neuroblastoma response criteria. *Pediatric Radiology, 49*(11), 1433–1440. https://doi.org/10.1007/s00247-019-04397-2
Ösby, U., Brandt, L., Correia, N., Ekbom, A. & Sparén, P. (2001). Excess Mortality in Bipolar and Unipolar Disorder in Sweden. *Archives of General Psychiatry, 58*(9), 844–850. https://doi.org/10.1001/archpsyc.58.9.844
Pahor, A. & Jaušovec, N. (2016). Making Brains run Faster: are they Becoming Smarter? *The Spanish Journal of Psychology, 19*, E88. https://doi.org/10.1017/sjp.2016.83
Pascual-Leone, A, Gomez-Tortosa, E., Grafman, J., Alway, D., Nichelli, P. & Hallett, M. (1994). Induction of visual extinction by rapid-rate transcranial magnetic stimulation of parietal lobe. *Neurology, 44*(3 Part 1), 494–494. https://doi.org/10.1212/wnl.44.3_part_1.494
Pascual-Leone, A. & Hallett, M. (1994). Induction of errors in a delayed response task by repetitive transcranial magnetic stimulation of the dorsolateral prefrontal cortex. *NeuroReport, 5*(18), 2517–2520. https://doi.org/10.1097/00001756-199412000-00028
Pascual-Leone, A., Rubio, B., Pallardó, F. & Catalá, M. D. (1996). Rapid-rate transcranial magnetic stimulation of left dorsolateral prefrontal cortex in drug-resistant depression. *The Lancet, 348*(9022), 233–237. https://doi.org/10.1016/s0140-6736(96)01219-6
Paykel, E. S. (2008). Partial remission, residual symptoms, and relapse in depression. *Dialogues in Clinical Neuroscience, 10*(4), 431–437. https://doi.org/10.31887/dcns.2008.10.4/espaykel
Persson, J., Struckmann, W., Gingnell, M., Fällmar, D. & Bodén, R. (2020). Intermittent theta burst stimulation over the dorsomedial prefrontal cortex modulates resting-state connectivity in depressive patients: A sham-controlled study. *Behavioural Brain Research, 394*, 112834. https://doi.org/10.1016/j.bbr.2020.112834
Pigott, H. E. (2014). The STAR*D Trial: It is Time to Reexamine the Clinical Beliefs That Guide the Treatment of Major Depression. *The Canadian Journal of Psychiatry, 60*(1), 9–13. https://doi.org/10.1177/070674371506000104
Pigott, H. E., Kim, T., Xu, C., Kirsch, I. & Amsterdam, J. (2023). What are the treatment remission, response and extent of improvement rates after up to four trials of antidepressant therapies in real-world depressed patients? A reanalysis of the STAR*D study's patient-level data with fidelity to the original research protocol. *BMJ Open, 13*(7), e063095. https://doi.org/10.1136/bmjopen-2022-063095
Pimenta, M. G., Brown, T., Arns, M. & Enriquez-Geppert, S. (2021). Treatment Efficacy and Clinical Effectiveness of EEG Neurofeedback as a Personalized and Multimodal Treatment in ADHD: A Critical Review. *Neuropsychiatric Disease and Treatment, Volume 17*, 637–648. https://doi.org/10.2147/ndt.s251547

Pizzagalli, D. A. (2011). Frontocingulate Dysfunction in Depression: Toward Biomarkers of Treatment Response. *Neuropsychopharmacology, 36*(1), 183–206. https://doi.org/10.1038/npp.2010.166

Pizzagalli, D. A., Webb, C. A., Dillon, D. G., Tenke, C. E., Kayser, J., Goer, F., Fava, M., McGrath, P., Weissman, M., Parsey, R., Adams, P., Trombello, J., Cooper, C., Deldin, P., Oquendo, M. A., McInnis, M. G., Carmody, T., Bruder, G. & Trivedi, M. H. (2018). Pretreatment Rostral Anterior Cingulate Cortex Theta Activity in Relation to Symptom Improvement in Depression: A Randomized Clinical Trial. *JAMA Psychiatry, 75*(6), 547. https://doi.org/10.1001/jamapsychiatry.2018.0252

Posthuma, D., Neale, M. C., Boomsma, D. I. & Geus, E. J. C. de. (2001). Are Smarter Brains Running Faster? Heritability of Alpha Peak Frequency, IQ, and Their Inter relation. *Behavior Genetics, 31*(6), 567–579. https://doi.org/10.1023/a:1013345411774

Pressman, M. R. & Fry, J. M. (1989). Relationship of Autonomic Nervous System Activity to Daytime Sleepiness and Prior Sleep. *Sleep, 12*(3), 239–245. https://doi.org/10.1093/sleep/12.3.239

Puttaert, D., Wens, V., Fery, P., Rovai, A., Trotta, N., Coquelet, N., Breucker, S. D., Sadeghi, N., Coolen, T., Goldman, S., Peigneux, P., Bier, J.-C. & Tiège, X. D. (2021). Decreased Alpha Peak Frequency Is Linked to Episodic Memory Impairment in Pathological Aging. *Frontiers in Aging Neuroscience, 13*, 711375. https://doi.org/10.3389/fnagi.2021.711375

Quiquempoix, M., Drogou, C., Erblang, M., Beers, P. V., Guillard, M., Tardo-Dino, P.-E., Rabat, A., Léger, D., Chennaoui, M., Gomez-Merino, D., Sauvet, F. & the Percaf Investigator Group. (2023). Relationship between Habitual Caffeine Consumption, Attentional Performance, and Individual Alpha Frequency during Total Sleep Deprivation. *International Journal of Environmental Research and Public Health, 20*(6), 4971. https://doi.org/10.3390/ijerph20064971

Ramsay, I. S., Lynn, P. A., Schermitzler, B. & Sponheim, S. R. (2021). Individual alpha peak frequency is slower in schizophrenia and related to deficits in visual perception and cognition. *Scientific Reports, 11*(1), 17852. https://doi.org/10.1038/s41598-021-97303-6

Razza, L. B., Santos, L. A. dos, Borrione, L., Bellini, H., Branco, L. C., Cretaz, E., Duarte, D., Ferrão, Y., Galhardoni, R., Quevedo, J., Simis, M., Fregni, F., Correll, C. U., Padberg, F., Trevizol, A., Daskalakis, Z. J., Carvalho, A. F., Solmi, M. & Brunoni, A. R. (2020). Appraising the effectiveness of electrical and magnetic brain stimulation techniques in acute major depressive episodes: an umbrella review of meta-analyses of randomized controlled trials. *Brazilian Journal of Psychiatry, 43*(5), 514–524. https://doi.org/10.1590/1516-4446-2020-1169

Reynolds, E. H. & Wilson, J. V. K. (2013). Depression and anxiety in Babylon. *Journal of the Royal Society of Medicine, 106*(12), 478–481. https://doi.org/10.1177/0141076813486262

Riedel, M., Möller, H.-J., Obermeier, M., Schennach-Wolff, R., Bauer, M., Adli, M., Kronmüller, K., Nickel, T., Brieger, P., Laux, G., Bender, W., Heuser, I., Zeiler, J., Gaebel, W. & Seemüller, F. (2010). Response and remission criteria in major depression – A validation of current practice. *Journal of Psychiatric Research, 44*(15), 1063–1068. https://doi.org/10.1016/j.jpsychires.2010.03.006

Riera-Serra, P., Navarra-Ventura, G., Castro, A., Gili, M., Salazar-Cedillo, A., Ricci-Cabello, I., Roldán-Espínola, L., Coronado-Simsic, V., García-Toro, M., Gómez-Juanes, R. & Roca, M. (2023). Clinical predictors of suicidal ideation, suicide attempts and

suicide death in depressive disorder: a systematic review and meta-analysis. *European Archives of Psychiatry and Clinical Neuroscience*, 1–21. https://doi.org/10.1007/s00406-023-01716-5

Rivero-Santana, A., Perestelo-Perez, L., Alvarez-Perez, Y., Ramos-Garcia, V., Duarte-Diaz, A., Linertova, R., Garcia-Perez, L., Serrano-Aguilar, P. & PECUNIA Group (2021). Stepped care for the treatment of depression: a systematic review and meta-analysis. *Journal of Affective Disorders, 294*, 391–409. https://doi.org/10.1016/j.jad.2021.07.008

Rodriguez, G., Copello, F., Vitali, P., Perego, G. & Nobili, F. (1999). EEG spectral profile to stage Alzheimer's disease. *Clinical Neurophysiology, 110*(10), 1831–1837. https://doi.org/10.1016/s1388-2457(99)00123-6

Roelofs, C., Krepel, N., Corlier, J., Carpenter, L. L., Fitzgerald, P. B., Daskalakis, Z. J., Tendolkar, I., Wilson, A., Downar, J., Bailey, N. W., Blumberger, D. M., Vila-Rodriguez, F., Leuchter, A. F. & Arns, M. (2020). Individual alpha frequency proximity associated with repetitive transcranial magnetic stimulation outcome: An independent replication study from the ICON-DB consortium. *Clinical Neurophysiology, S1388-2457*(20), 30532–0. https://doi.org/10.1016/j.clinph.2020.10.017

Romero, M. C., Davare, M., Armendariz, M. & Janssen, P. (2019). Neural effects of transcranial magnetic stimulation at the single-cell level. *Nature Communications, 10*(1), 2642. https://doi.org/10.1038/s41467-019-10638-7

Ross, J. S., Tse, T., Zarin, D. A., Xu, H., Zhou, L. & Krumholz, H. M. (2012). Publication of NIH funded trials registered in ClinicalTrials.gov: cross sectional analysis. *BMJ, 344*(jan03 1), d7292. https://doi.org/10.1136/bmj.d7292

Rossi, S., Antal, A., Bestmann, S., Bikson, M., Brewer, C., Brockmöller, J., Carpenter, L. L., Cincotta, M., Chen, R., Daskalakis, J. D., Lazzaro, V. D., Fox, M. D., George, M. S., Gilbert, D., Kimiskidis, V. K., Koch, G., Ilmoniemi, R. J., Lefaucheur, J. P., Leocani, L., Lisanby, S.H., Miniussi, C., Padberg, F., Pascual-Leone, A., Paulus, W., Peterchev, A.V., Quartarone, A., Rotenberg, A., Rothwell, J., Rossini, P.M., Santarnecchi, E., Shafi, M.M., Siebner, H.R., Ugawa, Y., Wassermann, E.M., Zangen, A., Ziemann, U. & Hallett, M. 2020, T. basis of this article began with a C. S. from the I. W. on "Present, Future of TMS: Safety, Ethical Guidelines", Siena, October 17-20, 2018, updating through April. (2021). Safety and recommendations for TMS use in healthy subjects and patient populations, with updates on training, ethical and regulatory issues: Expert Guidelines. *Clinical Neurophysiology, 132*(1), 269–306. https://doi.org/10.1016/j.clinph.2020.10.003

Rossi, S., Hallett, M., Rossini, P. M., Pascual-Leone, A. & The Safety of TMS Consensus Group (2009). Safety, ethical considerations, and application guidelines for the use of transcranial magnetic stimulation in clinical practice and research. *Clinical Neurophysiology, 120*(12), 2008–2039. https://doi.org/10.1016/j.clinph.2009.08.016

Roth, B. J., Saypol, J. M., Hallett, M. & Cohen, L. G. (1991). A theoretical calculation of the electric field induced in the cortex during magnetic stimulation. *Electroencephalography and Clinical Neurophysiology/Evoked Potentials Section, 81*(1), 47–56. https://doi.org/10.1016/0168-5597(91)90103-5

Rush, A. J., Trivedi, M. H., Carmody, T. J., Ibrahim, H. M., Markowitz, J. C., Keitner, G. I., Kornstein, S. G., Arnow, B., Klein, D. N., Manber, R., Dunner, D. L., Gelenberg, A. J., Kocsis, J. H., Nemeroff, C. B., Fawcett, J., Thase, M. E., Russell, J. M., Jody, D. N., Borian, F. E. & Keller, M. B. (2005). Self-Reported Depressive Symptom Measures: Sensitivity to Detecting Change in a Randomized, Controlled Trial of Chronically Depressed, Nonpsychotic Outpatients. *Neuropsychopharmacology, 30*(2), 405–

416. https://doi.org/10.1038/sj.npp.1300614
Rush, A. J., Trivedi, M. H., Wisniewski, S. R., Nierenberg, A. A., Stewart, J. W., Warden, D., Niederehe, G., Thase, M. E., Lavori, P. W., Lebowitz, B. D., McGrath, P. J., Rosenbaum, J. F., Sackeim, H. A., Kupfer, D. J., Luther, J. & Fava, M. (2006). Acute and Longer-Term Outcomes in Depressed Outpatients Requiring One or Several Treatment Steps: A STAR*D Report. *American Journal of Psychiatry, 163*(11), 1905–1917. https://doi.org/10.1176/ajp.2006.163.11.1905
Ryan, M., Brown, T., Vollebregt, M. & Arns, M. (2019). What's sleep got to do with it? Circadian rhythm, sleep disorder, ADHD and neurofeedback. *Asia Pacific Journal of Neurotherapy (APJNT), 1*(1), 85–98.
Sack, A. T., Paneva, J., Küthe, T., Dijkstra, E., Zwienenberg, L., Arns, M. & Schuhmann, T. (2023). Target engagement and brain state dependence of transcranial magnetic stimulation: implications for clinical practice. *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2023.09.011
Salomons, T. V., Dunlop, K., Kennedy, S. H., Flint, A., Geraci, J., Giacobbe, P. & Downar, J. (2013). Resting-state cortico-thalamic-striatal connectivity predicts response to dorsomedial prefrontal rTMS in major depressive disorder. *Neuropsychopharmacology : Official Publication of the American College of Neuropsychopharmacology, 39*(2), 488–498. https://doi.org/10.1038/npp.2013.222
Sangkuhl, K., Klein, T. E. & Altman, R. B. (2009). Selective serotonin reuptake inhibitors pathway. *Pharmacogenetics and Genomics, 19*(11), 907–909. https://doi.org/10.1097/fpc.0b013e32833132cb
Sarnthein, J., Stern, J., Aufenberg, C., Rousson, V. & Jeanmonod, D. (2006). Increased EEG power and slowed dominant frequency in patients with neurogenic pain. *Brain, 129*(1), 55–64. https://doi.org/10.1093/brain/awh631
Sauseng, P., Klimesch, W., Doppelmayr, M., Pecherstorfer, T., Freunberger, R. & Hanslmayr, S. (2005). EEG alpha synchronization and functional coupling during top-down processing in a working memory task. *Human Brain Mapping, 26*(2), 148–155. https://doi.org/10.1002/hbm.20150
Scangos, K. W., State, M. W., Miller, A. H., Baker, J. T. & Williams, L. M. (2023). New and emerging approaches to treat psychiatric disorders. *Nature Medicine, 29*(2), 317–333. https://doi.org/10.1038/s41591-022-02197-0
Schilberg, L., Engelen, T., Oever, S. ten, Schuhmann, T., Gelder, B. de, Graaf, T. A. de & Sack, A. T. (2018). Phase of beta-frequency tACS over primary motor cortex modulates corticospinal excitability. *Cortex, 103*, 142–152. https://doi.org/10.1016/j.cortex.2018.03.001
Schilberg, L., Oever, S. T., Schuhmann, T. & Sack, A. T. (2021). Phase and power modulations on the amplitude of TMS-induced motor evoked potentials. *PLoS ONE, 16*(9), e0255815. https://doi.org/10.1371/journal.pone.0255815
Schomer, D. L. & Silva, F. H. L. da. (2011). *Niedermeyer's Electroencephalography: Basic Principles, Clinical Applications, and Related Fields* (6th ed.). Lippincott Williams & Wilkins.
Siddiqi, S. H. & Philip, N. S. (2023). Hitting the Target of Image-Guided Psychiatry? *American Journal of Psychiatry, 180*(3), 185–187. https://doi.org/10.1176/appi.ajp.20230015
Siddiqi, S. H., Schaper, F. L. W. V. J., Horn, A., Hsu, J., Padmanabhan, J. L., Brodtmann, A., Cash, R. F. H., Corbetta, M., Choi, K. S., Dougherty, D. D., Egorova, N., Fitzgerald, P. B., George, M. S., Gozzi, S. A., Irmen, F., Kuhn, A. A., Johnson, K. A., Naidech,

A. M., Pascual-Leone, A., Phan, T.G., Rouhl, R.P.W., Taylor, S.F., Voss, J.L., Zalesky, A., Grafman, J.H., Mayberg, H.S. & Fox, M. D. (2021). Brain stimulation and brain lesions converge on common causal circuits in neuropsychiatric disease. *Nature Human Behaviour*, 1–10. https://doi.org/10.1038/s41562-021-01161-1

Siddiqi, S. H., Weigand, A., Pascual-Leone, A. & Fox, M. D. (2021). Identification of Personalized Transcranial Magnetic Stimulation Targets Based on Subgenual Cingulate Connectivity: An Independent Replication. *Biological Psychiatry, 90*(10), e55–e56. https://doi.org/10.1016/j.biopsych.2021.02.015

Siebner, H. R., Funke, K., Aberra, A. S., Antal, A., Bestmann, S., Chen, R., Classen, J., Davare, M., Lazzaro, V. D., Fox, P. T., Hallett, M., Karabanov, A. N., Kesselheim, J., Beck, M. M., Koch, G., Liebetanz, D., Meunier, S., Miniussi, C., Paulus, W., Peterchev, A.V., Popa, T., Ridding, M.C., Thielscher, A., Ulf Ziemann, U., Rothwell, J.C. & Ugawa, Y. (2022). Transcranial magnetic stimulation of the brain: What is stimulated? – A consensus and critical position paper. *Clinical Neurophysiology, 140*, 59–97. https://doi.org/10.1016/j.clinph.2022.04.022

Singal, A. G., Higgins, P. D. R. & Waljee, A. K. (2014). A Primer on Effectiveness and Efficacy Trials. *Clinical and Translational Gastroenterology, 5*(1), e45. https://doi.org/10.1038/ctg.2013.13

Sinyor, M., Levitt, A. J., Cheung, A. H., Schaffer, A., Kiss, A., Dowlati, Y. & Lanctôt, K. L. (2010). Does Inclusion of a Placebo Arm Influence Response to Active Antidepressant Treatment in Randomized Controlled Trials?: Results From Pooled and Meta-Analyses. *The Journal of Clinical Psychiatry, 71*(03), 270–279. https://doi.org/10.4088/jcp.08r04516blu

Smit, D. J. A., Posthuma, D., Boomsma, D. I. & Geus, E. J. C. de. (2005). Heritability of background EEG across the power spectrum. *Psychophysiology, 42*(6), 691–697. https://doi.org/10.1111/j.1469-8986.2005.00352.x

Smith, J. R. (1938). The Electroencephalogram During Normal Infancy and Childhood: II. The Nature of the Growth of the Alpha Waves. *The Pedagogical Seminary and Journal of Genetic Psychology, 53*(2), 455–469. https://doi.org/10.1080/08856559.1938.10533821

Soares, J. C. & Mann, J. J. (1997). The anatomy of mood disorders—review of structural neuroimaging studies. *Biological Psychiatry, 41*(1), 86–106. https://doi.org/10.1016/s0006-3223(96)00006-6

Spellman, T. & Liston, C. (2020). Toward Circuit Mechanisms of Pathophysiology in Depression. *American Journal of Psychiatry, 177*(5), 381–390. https://doi.org/10.1176/appi.ajp.2020.20030280

Steele, M., Jensen, P. S. & Quinn, D. M. P. (2006). Remission versus response as the goal of therapy in ADHD: A new standard for the field? *Clinical Therapeutics, 28*(11), 1892–1908. https://doi.org/10.1016/j.clinthera.2006.11.006

Steinert, T. & Fröscher, W. (2017). Epileptic Seizures Under Antidepressive Drug Treatment: Systematic Review. *Pharmacopsychiatry, 51*(04), 121–135. https://doi.org/10.1055/s-0043-117962

Sterman, M. B., Howe, R. C. & Macdonald, L. R. (1970). Facilitation of Spindle-Burst Sleep by Conditioning of Electroencephalographic Activity While Awake. *Science, 167*(3921), 1146–1148. https://doi.org/10.1126/science.167.3921.1146

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), 171–175. https://doi.org/10.1177/155005940503600307

Swanson, J.M., Kraemer, H.C., Hinshaw, S.P., Arnold, L.E., Conners, C.K., Abikoff, H.B., Clevenger, W., Davies, M., Elliott, G.R., Greenhill, L.L., Hechtman, L., Hoza, B., Jensen, P.S., March, J.S., Newcorn, J.H., Owens, E.B., Pelham, W.E., Schiller, E. Simpson, J.B., Severe, S., Vitiello, B., Wells, K., Wigal, T. & Wu, M. (2001). Clinical Relevance of the Primary Findings of the MTA: Success Rates Based on Severity of ADHD and ODD Symptoms at the End of Treatment. *Journal of the American Academy of Child & Adolescent Psychiatry, 40*(2), 168–179. https://doi.org/10.1097/00004583-200102000-00011

Taipale, H., Schneider-Thoma, J., Pinzón-Espinosa, J., Radua, J., Efthimiou, O., Vinkers, C. H., Mittendorfer-Rutz, E., Cardoner, N., Pintor, L., Tanskanen, A., Tomlinson, A., Fusar-Poli, P., Cipriani, A., Vieta, E., Leucht, S., Tiihonen, J. & Luykx, J. J. (2022). Representation and Outcomes of Individuals With Schizophrenia Seen in Everyday Practice Who Are Ineligible for Randomized Clinical Trials. *JAMA Psychiatry, 79*(3). https://doi.org/10.1001/jamapsychiatry.2021.3990

Tendler, A., Gersner, R., Roth, Y. & Zangen, A. (2018). Alternate day dTMS combined with SSRIs for chronic treatment resistant depression: A prospective multicenter study. *Journal of Affective Disorders, 240*, 130–136. https://doi.org/10.1016/j.jad.2018.07.058

Tendler, A., Ygael, N. B., Roth, Y. & Zangen, A. (2016). Deep transcranial magnetic stimulation (dTMS) – beyond depression. *Expert Review of Medical Devices, 13*(10), 987–1000. https://doi.org/10.1080/17434440.2016.1233812

The Neurofeedback Collaborative Group, Arnold, L. E., Arns, M., Barterian, J. A., Connor, S., deBeus, R. J., Hollway, J. A., Kerson, C., Lightstone, H., Lubar, J. F., Mc Burnett, K., Monastra, V. J., Mulligan, A., Buchan-Page, K., Pan, X. (Jeff), Rice, R., Roley-Roberts, M. E., Schrader, C. A., Tan, Y. & Williams, C. (2022). Neurofeedback for Attention-Deficit/Hyperactivity Disorder: 25-Month Follow-up of Double-Blind Randomized Controlled Trial. *Journal of the American Academy of Child & Adolescent Psychiatry*. https://doi.org/10.1016/j.jaac.2022.07.862

The Neurofeedback Collaborative Group. (2021). Double-Blind Placebo-Controlled Randomized Clinical Trial of Neurofeedback for Attention-Deficit/Hyperactivity Disorder With 13-Month Follow-up. *Journal of the American Academy of Child & Adolescent Psychiatry, 60*(7), 841–855. https://doi.org/10.1016/j.jaac.2020.07.906

The Pandas development team. (2020). *pandas-dev/pandas* (V1.0.3) [Python]. Zenodo. https://doi.org/10.5281/zenodo.3715232

Thomson, A. C., Kenis, G., Tielens, S., Graaf, T. A. de, Schuhmann, T., Rutten, B. P. F. & Sack, A. T. (2020). Transcranial Magnetic Stimulation-Induced Plasticity Mechanisms: TMS-Related Gene Expression and Morphology Changes in a Human Neuron-Like Cell Model. *Frontiers in Molecular Neuroscience, 13*, 528396. https://doi.org/10.3389/fnmol.2020.528396

Thut, G., Schyns, P. G. & Gross, J. (2011). Entrainment of Perceptually Relevant Brain Oscillations by Non-Invasive Rhythmic Stimulation of the Human Brain. *Frontiers in Psychology, 2*, 170. https://doi.org/10.3389/fpsyg.2011.00170

Thut, G., Veniero, D., Romei, V., Miniussi, C., Schyns, P. & Gross, J. (2011). Rhythmic TMS Causes Local Entrainment of Natural Oscillatory Signatures. *Current Biology, 21*(14), 1176–1185. https://doi.org/10.1016/j.cub.2011.05.049

van Tricht, M. J., Ruhrmann, S., Arns, M., Müller, R., Bodatsch, M., Velthorst, E., Koelman, J. H. T. M., Bour, L. J., Zurek, K., Schultze-Lutter, F., Klosterkötter, J., Linszen, D. H., Haan, L. de, Brockhaus-Dumke, A. & Nieman, D. H. (2014). Can quantitative

EEG measures predict clinical outcome in subjects at Clinical High Risk for psychosis? A prospective multicenter study. *Schizophrenia Research, 153*(1–3), 42–47. https://doi.org/10.1016/j.schres.2014.01.019

Trivedi, M. H., McGrath, P. J., Fava, M., Parsey, R. V., Kurian, B. T., Phillips, M. L., Oquendo, M. A., Bruder, G., Pizzagalli, D., Toups, M., Cooper, C., Adams, P., Weyandt, S., Morris, D. W., Grannemann, B. D., Ogden, R. T., Buckner, R., McInnis, M., Kraemer, H. C., Petkova, E., Carmody, T.J.& Weissman, M. M. (2016). Establishing moderators and biosignatures of antidepressant response in clinical care (EMBARC): Rationale and design. *Journal of Psychiatric Research, 78*, 11–23. https://doi.org/10.1016/j.jpsychires.2016.03.001

Ulrich, G., Renfordt, E., Zeller, G. & Frick, K. (1984). Interrelation between Changes in the EEG and Psychopathology under Pharmacotherapy for Endogenous Depression. A contribution to the predictor question. *Pharmacopsychiatry, 17*(06), 178–183. https://doi.org/10.1055/s-2007-1017433

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*(Biol. Psychol. 67 1–2 2004), 79–87. https://doi.org/10.1016/j.nicl.2017.07.006

van der Vinne, N., Vollebregt, M. A., Rush, A. J., Eebes, M., van Putten, M. J. A. M. & Arns, M. (2021). EEG biomarker informed prescription of antidepressants in MDD: a feasibility trial. *European Neuropsychopharmacology, 44*, 14–22. https://doi.org/10.1016/j.euroneuro.2020.12.005

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C.J., Polat, I., Feng, Y., Moore, E.W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E.A., Harris, C.R., Archibald, A.M., Ribeiro, A.H., Pedregosa, F., Paul van Mulbregt & SciPy 1. 0 Contributors. (2020). SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nature Methods, 17*, 261–272.

Voetterl, H. T. S., Sack, A. T., Olbrich, S., Stuiver, S., Rouwhorst, R., Prentice, A., Pizzagalli, D. A., van der Vinne, N., Waarde, J. A. van, Brunovsky, M., Oostrom, I. van, Reitsma, B., Fekkes, J., van Dijk, H. & Arns, M. (2023). Alpha peak frequency-based Brainmarker-I as a method to stratify to pharmacotherapy and brain stimulation treatments in depression. *Nature Mental Health*, 1–10. https://doi.org/10.1038/s44220-023-00160-7

Voetterl, H., van Wingen, G., Michelini, G., Griffiths, K. R., Gordon, E., DeBeus, R., Korgaonkar, M. S., Loo, S. K., Palmer, D., Breteler, R., Denys, D., Arnold, L. E., Jour, P. du, van Ruth, R., Jansen, J., van Dijk, H. & Arns, M. (2022). Brainmarker-I differentially predicts remission to various attention-deficit/hyperactivity disorder treatments: a blinded discovery, transfer and validation study. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging.* https://doi.org/10.1016/j.bpsc.2022.02.007

Wahbeh, H. & Oken, B. S. (2013). Peak High-Frequency HRV and Peak Alpha Frequency Higher in PTSD. *Applied Psychophysiology and Biofeedback, 38*(1), 57–69. https://doi.org/10.1007/s10484-012-9208-z

Walsh, B. T., Seidman, S. N., Sysko, R. & Gould, M. (2002). Placebo Response in Studies of Major Depression: Variable, Substantial, and Growing. *JAMA, 287*(14), 1840–1847. https://doi.org/10.1001/jama.287.14.1840

Webb, C. A., Trivedi, M. H., Cohen, Z. D., Dillon, D. G., Fournier, J. C., Goer, F., Fava, M., McGrath, P. J., Weissman, M., Parsey, R., Adams, P., Trombello, J. M., Cooper, C., Deldin, P., Oquendo, M. A., McInnis, M. G., Huys, Q., Bruder, G., Kurian, B. T., Jha, M., DeRubeis, R.J. & Pizzagalli, D. A. (2019). Personalized prediction of antidepressant v. placebo response: evidence from the EMBARC study. *Psychological Medicine, 49*(7), 1118–1127. https://doi.org/10.1017/s0033291718001708

Weigand, A., Horn, A., Caballero, R., Cooke, D., Stern, A. P., Taylor, S. F., Press, D., Pascual-Leone, A. & Fox, M. D. (2018). Prospective Validation That Subgenual Connectivity Predicts Antidepressant Efficacy of Transcranial Magnetic Stimulation Sites. *Biological Psychiatry, 84*(1), 28–37. https://doi.org/10.1016/j.biopsych.2017.10.028

World Health Organization. (2017). Depression and other common mental disorders: global health estimates. Retrieved June 15, 2022, from https://apps.who.int/iris/handle/10665/254610

World Health Organization. (2021). Suicide worldwide in 2019: global health estimates. Licence: CC BY-NC-SA 3.0 IGO.

World Health Organization. (2008). The global burden of disease: 2004 update. WHO Press.

Widge, A. S., Avery, D. H. & Zarkowski, P. (2013). Baseline and Treatment-Emergent EEG Biomarkers of Antidepressant Medication Response Do Not Predict Response to Repetitive Transcranial Magnetic Stimulation. *Brain Stimulation, 6*(6), 929–931. https://doi.org/10.1016/j.brs.2013.05.001

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

Wilhelm, K. (2019). Judging a book by its cover: Changing 'treatment resistant' to 'difficult-to-treat' depression. *Australian & New Zealand Journal of Psychiatry, 53*(2), 101–103. https://doi.org/10.1177/0004867419827644

Williams, L. M. (2016). Precision psychiatry: a neural circuit taxonomy for depression and anxiety. *The Lancet Psychiatry, 3*(5), 472–480. https://doi.org/10.1016/s2215-0366(15)00579-9

Williams, L. M., Carpenter, W. T., Carretta, C., Papanastasiou, E. & Vaidyanathan, U. (2023). Precision psychiatry and Research Domain Criteria: Implications for clinical trials and future practice. *CNS Spectrums*, 1–14. https://doi.org/10.1017/s1092852923002420

Williams, L. M. & Hack, L. M. (2020). A precision medicine–based, 'fast-fail' approach for psychiatry. *Nature Medicine, 26*(5), 653–654. https://doi.org/10.1038/s41591-020-0854-z

Williams, L. M., Rush, A. John., Koslow, S. H., Wisniewski, S. R., Cooper, N. J., Nemeroff, C. B., Schatzberg, A. F. & Gordon, E. (2011). International Study to Predict Optimized Treatment for Depression (iSPOT-D), a randomized clinical trial: Rationale and protocol. *Trials, 12*, 1–17. https://doi.org/10.1186/1745-6215-12-4

van Wingen, G., Bockting, C., Zantvoord, J., Wezenberg, B. & Cohen, S. (2020). Magnetic resonance imaging for individual prediction of treatment response in major depressive disorder: a systematic review and meta-analysis. *MedRxiv*, 2020.06.27.20141465. https://doi.org/10.1101/2020.06.27.20141465

Wu, W., Zhang, Y., Jiang, J., Lucas, M. V., Fonzo, G. A., Rolle, C. E., Cooper, C., Chin-Fatt, C., Krepel, N., Cornelssen, C. A., Wright, R., Toll, R. T., Trivedi, H. M., Monuszko, K., Caudle, T. L., Sarhadi, K., Jha, M. K., Trombello, J. M., Deckersbach, T., Adams, P., McGrath, P.J., Weissman, M.M., Fava, M., Pizzagalli, D.A., Arns, M., Trivedi, M.H. & Etkin, A. (2020). An electroencephalographic signature predicts antidepressant response in major depression. *Nature Biotechnology, 38*(4), 439–447. https://doi.org/10.1038/s41587-019-0397-3

Yao, D., Qin, Y., Hu, S., Dong, L., Vega, M. L. B. & Sosa, P. A. V. (2019). Which Reference Should We Use for EEG and ERP practice? *Brain Topography, 32*(4), 530–549. https://doi.org/10.1007/s10548-019-00707-x

Yeum, T.-S. & Kang, U. G. (2018). Reduction in Alpha Peak Frequency and Coherence on Quantitative Electroencephalography in Patients with Schizophrenia. *Journal of Korean Medical Science, 33*(26). https://doi.org/10.3346/jkms.2018.33.e179

Zangen, A., Zibman, S., Tendler, A., Barnea-Ygael, N., Alyagon, U., Blumberger, D. M., Grammer, G., Shalev, H., Gulevsky, T., Vapnik, T., Bystritsky, A., Filipčić, I., Feifel, D., Stein, A., Deutsch, F., Roth, Y. & George, M. S. (2023). Pursuing personalized medicine for depression by targeting lateral or medial prefrontal cortex with deep TMS. *JCI Insight, 8*(4), e165271. https://doi.org/10.1172/jci.insight.165271

Zhang, T., Huang, Y., Jin, Y., Ma, X. & Liu, Z. (2021). Treatment for Major Depressive Disorder by Repetitive Transcranial Magnetic Stimulation in Different Parameters: A Randomized Double-Blinded Controlled Trial. *Frontiers in Psychiatry, 12*, 623765. https://doi.org/10.3389/fpsyt.2021.623765

Zhang, Y., Wu, W., Toll, R. T., Naparstek, S., Maron-Katz, A., Watts, M., Gordon, J., Jeong, J., Astolfi, L., Shpigel, E., Longwell, P., Sarhadi, K., El-Said, D., Li, Y., Cooper, C., Chin-Fatt, C., Arns, M., Goodkind, M. S., Trivedi, M. H., Marmar, C.R. & Etkin, A. (2020). Identification of psychiatric disorder subtypes from functional connectivity patterns in resting-state electroencephalography. *Nature Biomedical Engineering, 5*(4), 309–323. https://doi.org/10.1038/s41551-020-00614-8

Zibman, S., Pell, G. S., Barnea-Ygael, N., Roth, Y. & Zangen, A. (2021). Application of transcranial magnetic stimulation for major depression: Coil design and neuroanatomical variability considerations. *European Neuropsychopharmacology, 45*, 73–88. https://doi.org/10.1016/j.euroneuro.2019.06.009

Zmeykina, E., Mittner, M., Paulus, W. & Turi, Z. (2020). Weak rTMS-induced electric fields produce neural entrainment in humans. *Scientific Reports, 10*(1), 11994. https://doi.org/10.1038/s41598-020-68687-8

Zrenner, C., Desideri, D., Belardinelli, P. & Ziemann, U. (2018). Real-time EEG-defined excitability states determine efficacy of TMS-induced plasticity in human motor cortex. *Brain Stimulation, 11*(2), 374–389. https://doi.org/10.1016/j.brs.2017.11.016

Zrenner, C. & Ziemann, U. (2023). Closed-loop stimulation. *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2023.09.014

ENGLISH SUMMARY

Depression is a severe and debilitating disorder that is highly prevalent and, especially when unsuccessfully treated, can have detrimental effects on the lives of those affected, one of the most severe effects being ending one's own life. Finding the best treatment for the individual early on in the course of the illness is therefore crucial. Despite overall effectiveness of different approved depression treatments, a substantial number of patients experiences no or only little symptom improvement. With increasing failed treatment attempts, the likelihood of achieving remission diminishes, and after multiple futile treatment courses a patient is considered difficult-to-treat. These patients often move on to receive noninvasive brain stimulation, such as rTMS, which shows good efficacy at the group level, but still results in a considerable number of non-remitters. One strategy to enhance treatment response is individualizing treatment to the patient, so called precision psychiatry. Within the context of precision psychiatry, stratified psychiatry can aid in identifying the best out of multiple approved treatments for a given disorder, for instance by means of imaging biomarkers. Throughout the years, a multitude of MRI and EEG markers has been examined, however valid criticism was raised regarding the quality of EEG biomarker research,

such as publication bias and the lack of direct and out-of-sample validations, concluding that EEG biomarkers are not yet ready for clinical implementation.

The present work first introduces the burden of depression and details the need for reliable biomarkers for treatment outcome in Chapter 1. It summarizes the scientific evidence for repetitive transcranial magnetic stimulation (rTMS) as an effective treatment for depression, especially for difficult-to-treat cases, and the concepts of precision psychiatry, and more specifically stratified psychiatry are explained. Neuroimaging modalities EEG and fMRI are shortly introduced, leading up towards Chapter 2 that reviews the literature of robust EEG and MRI biomarkers for TMS treatment in depression.

In Chapter 2 the current issues with biomarker research are shortly summarized. Only MRI and EEG biomarker studies with sample sizes of ≥ 88 patients or studies that attempted external replication were included in this systematic review, ensuring robustness of potential identified biomarkers. In total 2 MRI and 2 EEG biomarkers were identified as robust: network mapping and DLPFC-sgACC functional connectivity for MRI, and frontal-midline rACC theta and iAF for EEG.

The subsequent chapters focus on one of these identified robust biomarkers, the iAF, which had already been investigated as predictor for both ADHD and MDD albeit with somewhat inconsistent results. One potential reason for these discrepancies in findings could be the differences in methodology, particularly data cleaning, utilized across these studies. Thus, in Chapter 3, EEG processing was first optimized in a large clinical dataset of over 4000 people for optimal detection of the alpha peak. The resulting biomarker, Brainmarker-I was then normalized for age and sex and tested for treatment stratification in ADHD.
It was demonstrated that Brainmarker-I can successfully stratify boys with a relatively faster iAF to stimulant medication and those with a slower iAF to multimodal neurofeedback treatment. These findings were subsequently replicated in unseen independent datasets through blinded prediction of treatment outcome, thereby lending credibility to the presented findings and improving on criticism pertaining prior biomarker research.

In Chapter 4, the same biomarker was shown to extend to depression treatments, indicating stratification potential of patients with a first-line depression to sertraline and those with a difficult-to-treat depression to either 1 Hz or 10 Hz rTMS or electroconvulsive therapy (ECT). In addition to replicating the previous finding that patients with an iAF close to the stimulation frequency of 10 Hz are more likely to remit to 10 Hz rTMS, higher remission rates were found in a fast-iAF subgroup for 1 Hz rTMS and for a slow-iAF group for ECT. These new findings were once again replicated in independent, out-of-sample validations in a blinded fashion. This complements the third chapter, showing that Brainmarker-I represents a transdiagnostic treatment stratification biomarker for ADHD and MDD. This marker is both easy to collect and to interpret and, thanks to the stratification approach, can be applied without risk, which aids in straightforward realization in clinical practice.

In Chapter 5, the predictive potential of Brainmarker-I for depression was examined for higher frequency deep TMS (18 Hz), indicating that it could represent an alternative to ECT treatment in the slow iAF subgroup. A successful replication in 20 Hz TMS data that administered stimulation with a figure-of-eight coil suggested a higher relevance of stimulation frequency than coil shape. These findings enable stratification between TMS protocols only, which is particularly valuable since many patients seeking TMS treatment would opt for TMS rather than ECT at this stage of their treatment history, for fear of side effects. Furthermore, it enables full implementation of Brainmarker-I at TMS clinics that have no means to provide ECT treatment. The same study, additionally, discussed the indication that stimulation with a frequency of 18 Hz might not lead to entrainment of endogenous oscillations at the harmonic frequency of 9 Hz, and suggested that it might in fact lead to an unbalancing of the oscillators.

Thanks to rigorous testing in multiple, diverse clinical datasets and validation of effects in independent samples, the present work presents a robust transdiagnostic stratification biomarker that is ready for implementation in clinical practice.

NEDERLANDSE SAMENVATTING

Depressie is een ernstige en slopende stoornis die veel voorkomt en die, vooral indien deze niet succesvol behandeld wordt, nadelige gevolgen kan hebben voor het leven van de getroffenen, met als een van de ernstigste gevolgen het beëindigen van het eigen leven. Het vinden van de beste behandeling voor het individu in een vroeg stadium van de ziekte is daarom cruciaal. Ondanks de algemene effectiviteit van verschillende goedgekeurde depressiebehandelingen, ervaart een aanzienlijk aantal patiënten echter geen of slechts gedeeltelijke symptoomverbetering. Met elke onsuccesvolle behandelingspoging neemt de kans op het bereiken van remissie af en na meerdere onsuccesvolle behandelingskuren wordt een patiënt als moeilijk te behandelen gezien. Deze patiënten stappen vaak over op niet-invasieve hersenstimulatie, zoals rTMS, die een goede werkzaamheid op groepsniveau laat zien, maar nog steeds resulteert in een substantieel aantal niet-remitters. Een strategie om behandelrespons te verbeteren is het individualiseren van de behandeling naar de patiënt, de zogenaamde precisiepsychiatrie. Binnen de context van precisiepsychiatrie kan gestratificeerde psychiatrie helpen bij het identificeren van de beste van meerdere goedgekeurde

behandelingen voor een bepaalde stoornis, bijvoorbeeld door middel van neuroimaging biomarkers. In de loop der jaren is een veelheid aan MRI- en EEG-markers onderzocht, maar er is terechte kritiek geuit op de kwaliteit van EEG biomarkeronderzoek, zoals het ontbreken van directe en out-of-sample validaties, met als conclusie dat deze nog niet klaar zijn voor klinische implementatie.

Dit proefschrift introduceert eerst in hoofdstuk 1 de last van depressie en beschrijft de behoefte aan betrouwbare biomarkers voor het voorspellen van behandeluitkomst. Het wetenschappelijk bewijs voor repetitieve transcraniële magnetische stimulatie (rTMS) als effectieve behandeling voor depressie, met name voor moeilijk te behandelen gevallen, wordt samengevat en de concepten van precisiepsychiatrie en gestratificeerde psychiatrie worden toegelicht. EEG en fMRI als neuroimaging modaliteiten worden kort geïntroduceerd, wat leidt tot hoofdstuk 2 waarin de literatuur over robuuste EEG en MRI biomarkers voor TMS-behandeling bij depressie wordt besproken. Hier worden eerst de huidige problemen met biomarkeronderzoek kort samengevat. Alleen MRI en EEG biomarker studies met een steekproefgrootte van ≥ 88 patiënten of studies die probeerden hun bevindingen te repliceren in externe data, werden opgenomen in deze systematische review, om de robuustheid van elke potentieel geïdentificeerde biomarker te garanderen. In totaal werden 4 biomarkers geïdentificeerd als robuust: network mapping en DLPFC-sgACC functionele connectiviteit voor MRI en frontale-midlijn rACC theta en individuele alfa piek frequentie (iAF) voor EEG.

In de volgende hoofdstukken wordt specifiek ingegaan op één van deze geïdentificeerde robuuste biomarkers, de iAF, die al eerder was onderzocht als voorspeller voor zowel ADHD als MDD, alhoewel met enigszins inconsistente resultaten. Een mogelijke reden voor deze discrepanties in bevindingen zou kunnen liggen in de verschillen in methodologie, met name het opschonen van data, die in deze studies is gebruikt. Zo werd in hoofdstuk 3 in een grote klinische dataset van meer dan 4000 mensen de EEG-verwerking eerst geoptimaliseerd voor optimale detectie van de alfapiek. De resulterende biomarker, Brainmarker-I, werd vervolgens genormaliseerd voor leeftijd en geslacht en getest voor stratificatie van behandeling bij ADHD.

Er werd aangetoond dat Brainmarker-I gebruikt kan worden door jongens met een relatief snellere iAF te stratificeren naar stimulantia en jongens met een langzamere iAF naar multimodale neurofeedbackbehandeling. Deze bevindingen werden vervolgens gerepliceerd in onafhankelijke datasets door middel van geblindeerde voorspelling van behandeluitkomst, waardoor betrouwbaarheid werd gegeven aan de gepresenteerde bevindingen en eerdere kritiek op biomarkeronderzoek ondervangen werd.

In hoofdstuk 4 werd beschreven hoe deze zelfde biomarker, Brainmarker-I, ook toegepast kan worden bij depressie, waarbij patiënten met een eerstelijns depressie konden worden gestratificeerd naar sertraline en patiënten met een moeilijk te behandelen depressie naar 1 Hz, 10 Hz rTMS of elektroconvulsietherapie (ECT). Naast het repliceren van de eerdere bevinding dat patiënten met een iAF rond de 10 Hz stimulatiefrequentie meer kans hebben op remissie na 10 Hz rTMS, werden hogere remissiepercentages gevonden in een snelle iAF-subgroep voor 1 Hz rTMS en voor een langzame iAF-groep voor ECT. Deze nieuwe bevindingen werden opnieuw blind gerepliceerd in onafhankelijke, out-of-sample validaties. Dit laat ter aanvulling op het tweede hoofdstuk zien dat Brainmarker-I een transdiagnostische biomarker voor behandelstratificatie bij ADHD en MDD is. Deze marker is zowel eenvoudig te verzamelen als te interpreteren en kan, dankzij de stratificatiemethode, zonder risico worden toegepast, wat bijdraagt aan een directe realisatie in de klinische praktijk.

In hoofdstuk 5 werd Brainmarker-I voor depressie uitgebreid met diepe-TMS die een hogere stimulatiefrequentie (18 Hz) gebruikt met een helmspoel en die een alternatief kan bieden voor ECT-behandeling in de trage iAF-subgroep. Een succesvolle replicatie in 20 Hz TMS data waarbij stimulatie werd toegediend met een gebruikelijke figure-of-eight spoel suggereerde dat dit effect direct gerelateerd was aan de stimulatiefrequentie en niet aan de spoelvorm. Deze bevindingen maken stratificatie tussen uitsluitend TMS-protocollen mogelijk, wat bijzonder waardevol is omdat veel patiënten die TMS-behandeling wensen in dit stadium van hun behandelgeschiedenis eerder voor TMS dan voor ECT zouden kiezen uit angst voor bijwerkingen. Daarnaast maakt het de volledige implementatie van Brainmarker-I

mogelijk in TMS-klinieken die geen ECT kunnen aanbieden. Dezelfde studie bespreekt bovendien de indicatie dat stimulatie met een frequentie van 18 Hz mogelijk niet leidt tot entrainment van endogene oscillaties op de harmonische frequentie van 9 Hz, en suggereert dat het zelfs zou kunnen leiden tot een onbalans van de oscillatoren. Dankzij rigoureuze tests in meerdere, diverse klinische datasets en validatie van effecten in onafhankelijke datasets, levert het huidige werk een robuuste transdiagnostische stratificatiebiomarker die klaar is voor implementatie in de klinische praktijk.

IMPACT OF RESEARCH

As summarized throughout this thesis, depression is a life-threatening disorder. When not successfully treated, symptoms can become more severe and longer disease course and higher symptom severity complicate treating the disorder effectively and elevate the risk of relapse. In addition, symptom severity and time until full remission are positively correlated with risk of suicide attempts and death by suicide. Prescribing an effective treatment (for the individual) as early as possible is therefore crucial and can potentially be lifesaving.

Stratified psychiatry can help in achieving remission earlier relative to standard treatment prescription, which usually follows a trial-and-error approach regarding the order of evidence-based therapy attempts, by means of stratification biomarkers such as the one identified and validated in this thesis. Besides being potentially lifesaving, successful stratification can reduce the amount of disability-adjusted life years in people suffering from depression, enable improved functioning in society and diminish the use of clinical resources.
Likewise, faster symptom improvement in patients with ADHD can improve daily functioning, and thereby for instance educational per-

formance in children with ADHD and mitigate the burden on the mental health system.

The stratification biomarker introduced in this thesis was thoroughly investigated and identified as highly valuable for the prediction of treatment outcome in two common disorders, MDD and ADHD. Crucially for stratification, the biomarker predicted outcome differentially for various established and effective types of therapy, which is necessary to be able to match every patient to their respective best treatment early in the therapy process. These findings provide the basis for implementation of these stratification principles in the clinic, replacing the current trial-and-error approach and thereby eliminating the negative consequences of multiple failed treatment attempts. Thorough optimization of pre-analysis data processing steps and, importantly, blinded out-of-sample validations utilizing the same data processing in unseen heterogeneous clinical data add to the value of Brainmarker-I.

Crucially, the stratification approach and particularly the blinded, out-of-sample validations in independent datasets that were conducted in the presented studies can inform and provide an example for future biomarker research. As discussed throughout this work, previous neuroimaging biomarker research suffered from methodological issues such as small sample sizes, underreporting of null findings and lack of direct and particularly out-of-sample replications. These aspects have been taken into account in the development of Brainmarker-I, lending credibility and robustness to the presented results and assuring generalizability of the proposed method to clinical data acquired with differing equipment.
Thanks to this robust approach of biomarker development, the presented conclusions can be assumed to be reliable and to provide concrete improvements in remission rates when employed prospectively. We are proud to report that the here developed and introduced biomarker advice has indeed by now been implemented in regular clinical care for depression at a large mental healthcare clinic with multiple locations throughout the Netherlands. Naturally, the precise clinical value and applicability of Brainmarker-I remains to be determined and validated again by prospective feasibility trials,

now possible in these clinics. However, thanks to the nature of the stratification approach compared to a precision or personalized approach and due to the robust validation procedure applied, the risk of harming the patient through selection of a wrong treatment prescription is minimized, since patients will always be prescribed an evidence-based treatment. Our biomarker-based stratification in this sense simply determines for each individual which of the multiple available evidence-based therapies to begin with.

Although this concrete clinical implementation of our work represents a clear knowledge utilization with an impact on patients' lives already today, our research, like other research, is the basis for future scientific developments to further optimize the here described stratification principles. For example, as the transdiagnostic value of iAF has already been successfully demonstrated in our work, Brainmarker-I could represent a valuable predictor for treatment outcome in disorders other than depression and ADHD.

Moreover, the findings in Chapter 5 showing that a harmonic frequency of the iAF might not entrain oscillations but even lead to worse treatment outcome might inform future research that investigates the question whether stimulation at the individualized frequency can improve effectiveness. That MDD patients stimulated with a harmonic of their individual frequency showed diminished treatment effect might suggest that the individual brain state at the time of stimulation, such as the phase of the oscillation, could be more relevant than the frequency itself.

In summary, this thesis provides perhaps the first clinically actionable biomarker that can be - and in fact is already being - implemented in mental health care practices. It, moreover, illustrates how treatment stratification can act as a steppingstone to fully individualized precision psychiatry and can enhance remission rates and patients' lives in the here and now.

CURRICULUM VITAE

Helena Theresa Sofia Voetterl was born on 24th of August 1993 in Berlin, Germany.

After finishing secondary school, she went backpacking in Australia and Southeast Asia before returning to Germany to start her higher education. In 2017, she received her bachelor's degree in Cognitive Science after a semester abroad at the Behavioural Neurogenetics Lab at Victoria University of Wellington in New Zealand, studying autism spectrum disorder by means of animal models. For her master studies, Helena moved to the Netherlands to follow the research program Cognitive and Clinical Neuroscience with the specialization Neuropsychology. She conducted her master project at Toronto University, working at the rTMS clinic at UHN, Toronto Western Hospital, where she set up a TMS research lab, treated depression patients with accelerated rTMS and conducted TMS-EEG research. The resulting master thesis was published shortly after.

When returning from Canada, Helena started her doctoral research at Research Institute Brainclinics in Nijmegen, the Netherlands, where she collected EEGs for over 4 years and became a certified

TMS technician and supervisor in the Netherlands. She presented her research at multiple national and international conferences and published several papers, which are partially compiled in this thesis.

LIST OF PUBLICATIONS

PUBLISHED

Meijs, H., Voetterl, H., Sack, A. T., van Dijk, H., Wilde, B. D., Hecke, J. V., Niemegeers, P., Gordon, E., Luykx, J. J. & Arns, M. (2024). A posterior-alpha ageing network is differentially associated with antidepressant effects of venlafaxine and rTMS. *European Neuropsychopharmacology, 79*, 7–16. https://doi.org/10.1016/j.euroneuro.2023.11.002

Voetterl, H. T. S., Sack, A. T., Olbrich, S., Stuiver, S., Rouwhorst, R., Prentice, A., Pizzagalli, D. A., van der Vinne, N., Waarde, J. A. van, Brunovsky, M., Oostrom, I. van, Reitsma, B., Fekkes, J., van Dijk, H. & Arns, M. (2023). Alpha peak frequency-based Brainmarker-I as a method to stratify to pharmacotherapy and brain stimulation treatments in depression. *Nature Mental Health*, 1–10. https://doi.org/10.1038/s44220-023-00160-7

Klooster, D.*, Voetterl, H.*, Baeken, C. & Arns, M. (2023). Evaluating Robustness of Brain Stimulation Biomarkers for depression: A Systematic Review of MRI and EEG Studies. *Biological Psychiatry*. https://doi.org/10.1016/j.biopsych.2023.09.009
*contributed equally to this work as joint first authors

Voetterl, H., Wingen, G. van, Michelini, G., Griffiths, K. R., Gordon, E., DeBeus, R., Korgaonkar, M. S., Loo, S. K., Palmer, D., Breteler, R., Denys, D., Arnold, L. E., Jour, P. du, Ruth, R. van, Jansen, J., van Dijk, H. & Arns, M. (2022). Brainmarker-I differentially predicts remission to various attention-deficit/hyperactivity disorder treatments: a blinded discovery, transfer and validation study. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*. https://doi.org/10.1016/j.bpsc.2022.02.007

Luykx, J. J. & Voetterl, H. T. S. (2022). Toward Ecologically Sustainable Mental Health Care—A Call for Action From Within Psychiatry. *JAMA Psychiatry, 79*(6), 524–525. https://doi.org/10.1001/jamapsychiatry.2022.0594

Sheen, J. Z., Miron, J.-P., Mansouri, F., Dunlop, K., Russell, T., Zhou, R., Hyde, M., Fox, L., Voetterl, H., Daskalakis, Z. J., Griffiths, J. D., Blumberger, D. M. & Downar, J. (2022). Cardiovascular biomarkers of response to accelerated low frequency repetitive transcranial magnetic stimulation in major depression. *Journal of Affective Disorders, 318*, 167–174. https://doi.org/10.1016/j.jad.2022.08.105

Voetterl, H., Miron, J.-P., Mansouri, F., Fox, L., Hyde, M., Blumberger, D. M., Daskalakis, Z. J., Vila-Rodriguez, F., Sack, A. T. & Downar, J. (2021). Investigating EEG biomarkers of clinical response to low frequency rTMS in depression. *Journal of Affective Disorders Reports, 6*, 100250. https://doi.org/10.1016/j.jadr.2021.100250

Arns, M. & Voetterl, H. (2021). Neurophysiological effects of rTMS: Revisiting the role of the N100 as a clinically useful marker in depression. *Clinical Neurophysiology, 132*(9), 2259–2260. https://doi.org/10.1016/j.clinph.2021.06.010

Miron, J.-P., Voetterl, H., Fox, L., Hyde, M., Mansouri, F., Dees, S., Zhou, R., Sheen, J., Jodoin, V. D., Mir-Moghtadaei, A., Blumberger, D. M., Daskalakis, Z. J., Vila-Rodriguez, F. & Downar, J. (2021). Optimized repetitive transcranial magnetic stimulation techniques for the treatment of major depression: A proof of concept study. *Psychiatry Research, 298*, 113790. https://doi.org/10.1016/j.psychres.2021.113790

Miron, J.-P., Sheen, J., Panzarella, T., Hyde, M., Mansouri, F., Fox, L., Voetterl, H., Jodo-

in, V. D., Lespérance, P., Longpré-Poirier, C., Juster, R.-P., Daskalakis, Z. J., Blumberger, D. M. & Downar, J. (2021). The relationship between pre-treatment heart rate variability and response to low-frequency accelerated repetitive transcranial magnetic stimulation in major depression. *Journal of Affective Disorders Reports, 6*, 100270. https://doi.org/10.1016/j.jadr.2021.100270

Miron, J.-P., Voetterl, H., Mansouri, F., Blumberger, D. M., Daskalakis, Z. J. & Downar, J. (2019). A case series of a novel 1 Hz right-sided dorsolateral prefrontal cortex rTMS protocol in major depression. *Brain Stimulation, 13*(2). https://doi.org/10.1016/j.brs.2019.11.006

UNDER REVIEW

Voetterl, H.T.S., Alyagon, U., Middleton, V.J., Downar, J., Zangen, A., Sack, A.T., van Dijk, H., Halloran, A., Donachie, N. & Arns, M. Does 18 Hz deep TMS benefit a different subgroup of depressed patients relative to 10 Hz rTMS? The role of the individual alpha frequency (under review). European Neuropsychopharmacology

Boxum, M.*, Voetterl, H.*, van Dijk, H., Gordon, E., DeBeus, R., Arnold, L.E. & Arns, M. Challenging the Diagnostic Value of TBR: Insights from an EEG subtyping meta-analytical approach in ADHD (under review). Applied Psychophysiology and Biofeedback

* Indicates shared first authorship.

CONFERENCE PROCEEDINGS

ORAL PRESENTATIONS

Does 18 Hz deep TMS benefit a different subgroup of depressed patients relative to 10 Hz rTMS? The role of the individual alpha frequency. *6th European Conference on Brain Stimulation in Mental Health*, 2024, Lisbon, Portugal

Does 18 Hz deep TMS benefit a different subgroup of depressed patients relative to 10 Hz rTMS? The role of the individual alpha frequency. *BeNe Brain Stimulation Symposium*, 2023, Nijmegen, the Netherlands

Brainmarker-I: Een EEG-biomarker voor het individualiseren van behandeling voor hardnekkige depressive. *Nederlandse Vereniging voor Psychiatrie (NVvP) Voorjaarscongres* 2023, Maastricht, the Netherlands

Facilitating rTMS treatment prescription with Brainmarker-I: a transdiagnostic EEG stratification biomarker. *5th European Conference on Brain Stimulation in Psychiatry*, 2022, Zagreb, Croatia

Brainmarker-I differentially predicts remission to 1 Hz rTMS, 10 Hz rTMS and ECT: a stratified psychiatry approach. *BeNe Brain Stimulation Symposium*, 2022, Hasselt, Belgium

Taking the Guesswork Out of Stepped-Care: Biomarker Based Stratification to rTMS and Antidepressants. *International Pharmaco-EEG Society*, 2021, online

Neurodevelopmentally inspired EEG biomarker for transdiagnostic treatment stratification. *4th European Conference on Brain Stimulation in Psychiatry*, 2020, Nijmegen, the Netherlands

POSTER PRESENTATIONS

Brainmarker-I for depression: EEG-biomarker stratification between ECT, 10 Hz and 1 Hz rTMS. *5th International Brain Stimulation Conference*, 2023, Lisbon, Portugal

Brainmarker-I: a transdiagnostic EEG biomarker predicting remission to various ADHD and depression treatments. *4th International Brain Stimulation Conference*, 2021, Charleston, SC, US.

ACKNOWLEDGEMENT

I first and foremost would like to thank my team of supervisors Martijn Arns, Hanneke van Dijk and Alexander Sack who have been by my side throughout my PhD journey, lending a wonderful mix of support, critical feedback and creative ideas. Martijn, your enthusiasm and flood of ideas are unstoppable. Thank you for all the intriguing discussions and for giving me the opportunity to prove myself. Hanneke, if it hadn't been for you, I would never have managed to code in Python. With much fondness I look back on many hours of quietly sitting next to you, observing the ease with which you make the code bend to your will. Alex, your calm, thoughtful way of bringing up insightful ideas, whether they were related to research or raised at a dinner party, always inspired me.

This thesis would not have been the same without the motivational attitude and creative artfulness of Mark Koppenberg who is almost solely responsible for the unique design of this book and the figures throughout the chapters. Mark, your ingenuity and wit are unrivalled.

A heartfelt thank you goes to all my colleagues and former colleagues Noralie, Nikita, Lauren, Amourie, Eva, Hannah, Renee and Katrien who made the office lively and work much more enjoyable, and to the interns I had the honor to train and attempt to guide along the way: Priya, Ylka, Rand, Marit. I would also like to thank Nicole for all the help and the pleasant interaction throughout the years, and the many neurocare colleagues I worked with, trained, and raised a glass with many times.

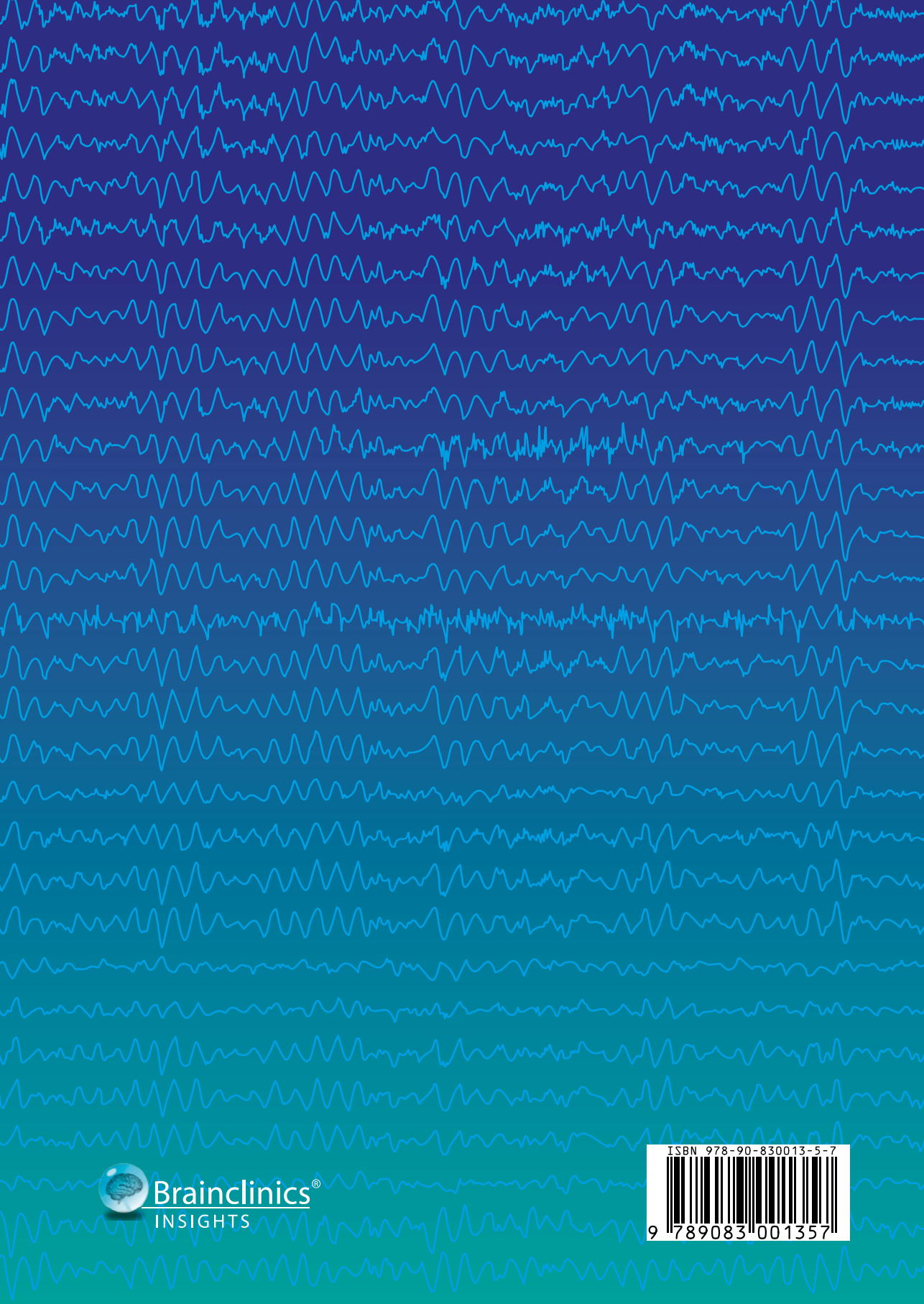
In order to effectively realize a future of personalized psychiatry, data sharing is of utmost importance. I would therefore like to extend a special thank you to all collaborators who kindly shared their clinical and imaging data with me which allowed for the inclusion of the broad range of treatments presented in this book. I am very much looking forward to further successful collaborations in the future.

Last but definitely not least, I want to thank all my close ones, who supported me throughout the whole, or part of this, journey with their bottomless belief in me and my skills, particularly in times when I wasn't sure of them myself.
I particularly want to thank my family. Especially my wonderful sister Viky, for her inexhaustible curiosity, wise counsel and for always asking the right questions. Please never stop cheering for me! And my loving mother: thank you, Mum, for giving me everything you have and for always believing in me.

I'd also like to thank Cor for looking at me as if there is nothing more fascinating than my lengthy digressions about the brain. Thank you for celebrating every little success with me as if it was your own. Many thanks to Carrie, for your critical input in both professional and personal matters and for always having my back. Jesse, thank you for warmly welcoming me to the Netherlands, for fostering my English skills and for the fun distractions from PhD life. Alejandro, Carolina, Casper, Fernando, Greg, Mariana, Nishant, Raimon, Tea: thank you for all the wonderful times we had and for showing me that a PhD can be so much more stressful. ;) Bart, Aafke, Sem, Amarens: thank you for your courage and for being the best companions for sedentary activities. Christine, Justin, Laura, Pia, Zamantha,

and everyone else I met on the way: thank you for suffering with me through the hard and the fun times of science.

I'd also like to thank everyone else who hasn't been named yet who supported me in any way, whether financially, emotionally or otherwise.



Brainclinics®
INSIGHTS

ISBN 978-90-830013-5-7