Review

Evaluating Robustness of Brain Stimulation Biomarkers for Depression: A Systematic Review of Magnetic Resonance Imaging and Electroencephalography Studies

Debby Klooster, Helena Voetterl, Chris Baeken, and Martijn Arns

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.

https://doi.org/10.1016/j.biopsych.2023.09.009

The search for biomarkers of clinical response to noninvasive brain stimulation (NIBS) treatments has been a major focus of attention over the last decade. After the introduction of the DSM-5 in 2013, an even stronger focus on biomarker research was ignited by the launch of the National Institute of Mental Health Research Domain Criteria project. A few years later, Research Domain Criteria inclusion became mandatory for National Institute of Mental Health–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 iSPOT-D (International Study to Predict Optimized Treatment in Depression) (1), EMBARC (Establishing Moderators and Biosignatures of Antidepressant Response for Clinical Care) (2), and CAN-BIND (Canadian Biomarker Integration Network in Depression) (3). 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 or addiction, with currently more than 24 Food and Drug Administration device approvals (4), as well as transcranial electrical stimulation. Many NIBS studies have been complemented by imaging work (5–7). Because many NIBS applications have built upon neuroscientific knowledge (e.g., frontal asymmetry) given the focus on interventional

psychiatry and brain circuit therapeutics (8,9), identifying NIBS biomarkers is of great importance both to improve clinical outcomes and to validate hypothesized working mechanisms. Therefore, we aimed to systematically review the current state of biomarker-driven precision psychiatry for NIBS.

Several previous reviews and meta-analyses have investigated biomarkers for depression focused on electroencephalography (EEG) (10) or magnetic resonance imaging (MRI) (11), and a critical meta-analysis questioned the usefulness of EEG biomarkers for guiding antidepressant response (12). That meta-analysis raised valid concerns about biomarker studies, criticizing a lack of replications, particularly those that are out-of-sample, 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. Thus, this systematic review was focused on 1) adequately powered imaging studies and 2) studies that attempted to (out-of-sample) replicate earlier findings. Concretely, biomarkers were considered robust if they showed the ability to predict treatment response in an adequately powered study and/or were replicated in multiple studies. 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.* (12) was that EEG biomarker studies were limited by small 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 the Supplement). Because these pivotal trials yielded inconsistent sample size justifications, we conducted a power calculation in GPower 3.1 (13) to determine a minimum sample size to define robust studies. We used a categorical outcome measure that reflects the difference in biomarker presence between responders and nonresponders expressed as a medium effect size (Cohen's d = 0.5), with an alpha level of $p < .05$ and power of 0.7, resulting in a sample size of N = 88. Furthermore, studies with smaller sample sizes could be included on the condition that subsequent replication studies were reported in an independent sample. Studies that investigated pretreatment biomarkers of any NIBS modality and protocol were included. Studies that identified treatment-emergent biomarkers (biomarkers that reflect changes during treatment) were not taken into account in this review because 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 the Supplement. Figure 1 shows the inclusion/exclusion and final selection of studies for EEG and MRI.

RESULTS

Results are summarized below as well as in Figure 1, with biomarkers grouped thematically. Information on the included MRI and EEG studies is summarized in Figures 2 and 3, respectively. The systematic review only yielded rTMS studies because no studies on other NIBS/transcranial electrical stimulation modalities met our inclusion criteria. rTMS is a technique that can be used to noninvasively modulate brain activity and is based on the principles of electromagnetic induction (14). In the specific case of depression treatment, mostly the left dorsolateral prefrontal cortex (DLPFC) is stimulated (15). When a different stimulation location was used or the biomarker was protocol specific, this is explicitly stated. In addition, a detailed description of technical terms used in this section can be found in Figure 4.

MRI Biomarkers

Anatomical MRI: Cortical Thickness. In a first study, Boes *et al.* reported that thinner rostral anterior cingulate cortex (rACC) at baseline was associated with better clinical improvement (16). However, subsequent work failed to replicate this finding (17), although here accelerated intermittent theta burst stimulation (iTBS) was used, whereas Boes *et al.* used 10 Hz rTMS.

Functional MRI: DLPFC–Subgenual ACC Functional Connectivity. In an influential 2012 study, Fox *et al.* 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 (18). They demonstrated that the clinical benefit of rTMS for depression was related to the
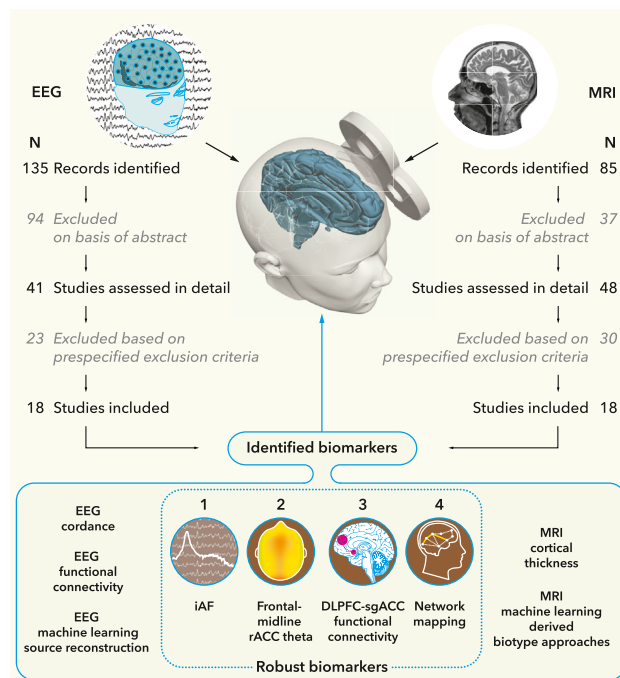
Figure 1. Flow diagram of total studies identified, excluded, and included in the systematic review for electroencephalography (EEG) biomarkers (left) and magnetic resonance imaging (MRI) biomarkers (right), as well as all biomarkers identified and the most robust biomarkers that emerged from this systematic 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.

intrinsic functional connectivity (FC) of the respective DLPFC target to the subgenual ACC (sgACC) (as part of the default mode network), as determined by resting-state functional MRI (rs-fMRI). This FC was indexed as "anticorrelation" of the sgACC to prefrontal cortical areas and is suggestive of a way to individualize prefrontal rTMS sites for MDD treatment by selecting the most sgACC-anticorrelated prefrontal site.

Several studies have attempted to replicate this finding, with successful conceptual replications by Weigand *et al.* (19), Cash *et al.* (20), Siddiqi *et al.* (21), and Elbau *et al.* (22). However, studies in which a whole-brain FC analysis was performed using the sgACC as seed region have shown no relationship between functional anticorrelations between the seed and stimulation targets in the left DLPFC and response (23–28). These nonreplication studies are all based on individual rs-fMRI data. Hopman *et al.* even suggested an inverse relationship, i.e., stronger connectivity between the sgACC and stimulation site was related to improved clinical response (25).

The studies by Fox *et al.* (18) and Weigand *et al.* (19) 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 improve TMS personalization (20). Besides replication of previous results

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)	10 Hz rTMS or 20 Hz rTMS	BDI	r = −0.55, p < .005	Better clinical response was related to higher anti-correlations between the stimulation site in the DLPFC and the sgACC	
		16 (12)	10 Hz rTMS	MADRS	r = −0.52, p < .05		
	Baeken, 2017	50 (44)	Accelerated iTBS	HDRS-17	p<.01		FC between sgACC and medial orbito-frontal 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)	10 Hz rTMS or iTBS	HDRS-17	AUC = 0.87, p<.001, r = −0.62, p<.001		Clinical response (post-treatment and 12 weeks after rTMS) was related to lower functional connectivity between the sgACC and the right DLPFC
		50 (32)*			AUC = 0.79, p<.001, r = −0.49, p=.001		
	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

Figure 2. Overview of study details on the included magnetic resonance imaging studies based on sample size ($N \geq 88$; 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
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 3. 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

based on the normative connectome, this study showed that the relationship between functional anticorrelation 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 (29). 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 the advantages of using individual rs-fMRI data. Siddiqi et al. (21) 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 to date (N = 295) focusing on the potential of sgACC connectivity to infer TMS coil positions (22). Although an association between FC between the sgACC and left DLPFC target and clinical response was observed, this association was much weaker (r = −0.16) than that observed in previous studies [e.g., r = −0.355 (18)], 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. (30) was used to derive the time series of the sgACC, a weak but robust correlation was found. Notably, this relationship was stronger in a subgroup of patients with strong global signal fluctuations due to burst breathing patterns (22). It was suggested that this weaker relationship could potentially be attributed to the relatively low resolution of the rs-fMRI data (voxel size 5 × 5 × 5 mm) (31). Indeed, better data quality could lead to better predictions, and currently, more sophisticated scanning sequences such as multiecho and multiband sequences are available (32). Moreover, studies that showed stronger relationships between anticorrelations 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 in the future may become a robust MRI-derived biomarker. However, optimal methodology to compute FC needs further investigation, and future prospective studies are warranted to demonstrate the 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

(33), have also been related to clinical responses. 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, representing different sources of causal effects (8). 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 patients with MDD using co-ordinate network mapping of spatially heterogeneous coordinates (34). It is worth noting that this emotional network resembles the depression network by Siddiqi et al. (ρ = 0.47, p = .00) (8). Closer proximity between the stimulation target and the emotional network–derived personalized targets was associated with better clinical response (34).

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: Machine Learning-Derived Biotype Approach. Using FC as input to machine learning (ML) approaches, Drysdale et al. (35) identified 4 clusters, called biotypes, which showed differential sensitivity to response to rTMS over the dorsomedial PFC in a subsequent validation. 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). Notably, partial response was defined as a >25% reduction in Hamilton Depression Rating Scale scores, although the results were similar when using the more traditional >50% cutoff for response, but predicted full response was lower (e.g. ∼ 63% for biotype 1).

Later work by Dinga et al. (36) failed to replicate these findings in a more heterogeneous sample of 187 patients with depression and anxiety. Their analysis led to 3 instead of 4 clusters. Neither the canonical correlates nor the clusters were statistically significant. Overfitting of the nonregularized canonical correlation analysis and arbitrary definitions of the subtypes are potential methodological explanations for this nonreplication (37). Variations in clinical sample characteristics may also explain the nonreplication (38).

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.

Global Impressions Scale (Improvement); DASS, Depression Anxiety Stress Scale; DLPFC, dorsolateral prefrontal cortex; EC, eyes closed; EO, eyes open; 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.

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 (30), 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. (22), 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-stimsite 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 (33). Network mapping has since been expanded to contain other (causal) sources of information such as TMS stimulation sites (TMS network mapping) (8) or coordinates related to abnormal brain functioning (coordinate network mapping) (34).

The emotional network, identified by Cash et al. (34), 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. (35) 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 (53).

Cordance is an EEG measure, originally developed by Leuchter and colleagues (55) 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.
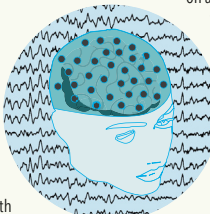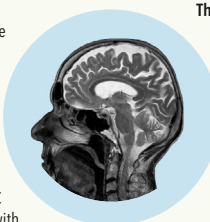
Figure 4. Glossary of terms used throughout the article. 3D, 3-dimensional; DLPFC, dorsolateral prefrontal cortex; EEG, electroencephalography; FC, functional connectivity; MRI, magnetic resonance imaging; rs-fMRI, resting-state functional MRI; sgACC, subgenual anterior cingulate cortex; TMS, transcranial magnetic stimulation.

An early study by Arns *et al.* (39) conducted with 90 patients with MDD found that high frontocentral theta power, low prefrontal delta and beta cordance, and high P300 amplitude at baseline were associated with nonresponse to 10 Hz rTMS over the DLPFC. However, the findings for theta and P300 could not be replicated by the same group (40).

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 (41).

The findings were replicated in an independent sample, and it was also shown that the obtained predictor was specific to 10 Hz rTMS because it could not predict response to iTBS treatment (42). 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 study and *n* = 35 per treatment arm in the replication study); however, the concept of independent-sample replication strengthens the findings, and the differential prediction for

iTBS versus 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 (43). When this alpha signature of response to rTMS was prospectively tested, it predicted change on the anxiety subscale of the Depression Anxiety Stress Scale after 1 Hz TMS 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, because no effects for depressive symptoms were reported (neither Beck Depression Inventory nor Depression Anxiety Stress Scale–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 receiver operating characteristic curve, model performance was rather weak with an area under the receiver operating characteristic curve = 0.67 [for a detailed critique about the methodology, see (44)].

A novel approach that conceptually resembles the previously mentioned rs-fMRI biotype analyses is 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 (45). Subsequent application of this signature to new samples yielded an association with response to both antidepressants and rTMS in men, but not women. Because selecting EEG biomarkers using genetic data is a novel technique, this study should 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 (46–48). Initial findings for iAF were mixed. Some studies reported an association between slow alpha and nonresponse to DLPFC rTMS (39,49), which could not be replicated by the same group (40) or by Widge *et al.* (50). On the other hand, adding iAF to a predictive model of nonlinear EEG features of the alpha band improved model prediction, although in a rather small group of non-responders ($N = 20$) (51).

More recent work has 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 right DLPFC rTMS) and could only be found using an average reference (indexing more focal activity than the linked-ears montage that was used in the studies mentioned above) (52). 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 (6). These results were successfully replicated (52) in the same sample by Krepel *et al.* (40), where previous findings (using linked-ears reference) initially could not be replicated. This emphasizes the importance of exact methodological replications and a uniform way to preprocess and analyze EEG data (53). These differential predictive results for iAF in 10 Hz and 1 Hz rTMS were recently further validated in a treatment stratification approach using iAF-based Brainmarker-I including multiple blinded out-of-sample validations for 1 Hz rTMS and electroconvulsive therapy (54).

EEG Cordance. A study investigating prefrontal theta cordance, developed by Leuchter *et al.* (55), found that baseline cordance could predict response to 1 Hz rTMS with high accuracy (56), although this could not be replicated in another study in which only 1-week change in theta cordance at central electrode sites predicted differences in response but not baseline or prefrontal cordance (57).

Two ML studies investigated pretreatment frontal cordance to predict outcome to 25 Hz rTMS in the same dataset of 147 subjects using artificial neural networks (58,59). High classification accuracies were obtained, although in the first study, only a 6-fold cross validation was conducted; however, models were not tested in an external validation set, which is considered necessary to prevent model overfitting (60). The second study in 2016 included a separate sample of 36 subjects for external validation, and high accuracy was achieved (area under the receiver operating characteristic curve = 0.807–0.918). 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 nonresponders (61). 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 that could distinguish 2 clinical subtypes that responded differentially to psychotherapy in posttraumatic stress disorder and selective serotonin reuptake inhibitor treatment in MDD (62). However, no such differences between subtypes were found for rTMS, suggesting little relevance for rTMS prediction but possible relevance for stratification between selective serotonin reuptake inhibitor and rTMS treatment. Another ML model, which was 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%) (63), which could not be replicated in an independent sample (64). The discovery analysis was based on only 12 responders, compared with 128 responders in the replication sample. One important caveat of the replication analysis was the strong differences in EEG processing, which can lead to different results (52).

Thus, findings regarding FC are inconclusive, with different processing and modeling approaches hampering robust findings.

DISCUSSION

The aim of this systematic review was to assess the progress regarding EEG and MRI biomarkers for NIBS techniques. To be responsive to previous criticisms, particularly the lack of replications as highlighted by Widge *et al.* (12), we focused on robustness in this review. To achieve this, we only included studies with a sample size of $N \geq 88$ or studies that attempted to replicate biomarkers in independent samples 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 1). All studies focused on rTMS while no relevant imaging biomarker studies were found for transcranial electrical stimulation.

MRI Biomarkers

The most robust rs-fMRI–based metric predicting clinical response supported by a large sample ($N = 295$) (22) as well as several independent replications is the anticorrelation between the stimulation target (within the left DLPFC) and the sgACC (18). This anticorrelation 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 (22), potentially suggesting reduced utility in clinical practice. Thus, prospective studies that target the personalized location in the DLPFC with the highest anticorrelation with the sgACC should demonstrate whether 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 that used data from 151 TMS stimulation sites (4 merged TMS datasets) (8) and was independently replicated (32), 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 (65)], which is supported by the finding that an rACC-engaging task can elicit this frequency (41,42). Interestingly, rACC activation has been found to be predictive across imaging modalities, including EEG and fMRI (65). However, this was true for both sertraline and placebo response (66). Thus, despite successful replication, future studies should further investigate whether this finding is specific to 10 Hz rTMS versus iTBS or should instead be considered a nonspecific predictor of response, including placebo.

The iAF finding emerged from 2 well-powered studies ($N = 143$, $N = 153$) conducted by 2 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. Although 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 been shown to be highly influential because Elbau *et al.* (22) only found a relationship between the FC between the stimulation site and the 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 that were included in this review demonstrated the clinical value of using individual rs-fMRI data compared with group connectome data (21,22,34,67). Future research needs to compare biomarkers derived from these different connectomes and answer the question of whether baseline individualized rs-fMRI data collection should be added to treatment protocols.

In the case of the iAF EEG biomarker, initial findings were mixed, despite the fact that several well-powered studies were used to examine the effect [e.g., $N = 180$ (48) or $N = 90–106$ (39,40)], and replication analyses were conducted. Later work actually led to consistent and robust findings (6,52), showing that the crucial factors were 1) use of the correct EEG montage; initial studies used the less focal linked-ears reference, while Roelofs *et al.* (52) 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 that 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 depends on the preprocessing pipelines used. Future research is necessary to investigate whether the content of these metrics is related to core brain mechanisms or reflects other sources of signal fluctuations such as respiration or cardiac patterns.

Artificial Intelligence and ML

The current review reveals limited biomarker potential for artificial intelligence 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 (58), which are needed to prevent overfitting (60). In addition, some out-of-sample validation results were only significant for measures other than the discovery measure (e.g., anxiety instead of depression) (43), and some could not be replicated, possibly due to overfitting (64). Finally, it is important to use consistent definitions of

response and remission (e.g., not "partial response") 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 may explain at least in part some of the unsuccessful replication attempts. Moreover, especially for biomarkers with weaker effects, the cost/benefit ratio needs to be assessed.

Although predicting NIBS outcome in other disorders would be highly relevant, we only discussed robust biomarkers for MDD in the current article. Future research is needed to determine whether these are also predictive of treatment response in other disorders.

Finally, prospective studies, similar to the one conducted by van der Vinne et al. (68), will be necessary to test treatment individualization in daily clinical practice.

CONCLUSIONS

This systematic review has identified 4 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 [e.g., stratification between 10 Hz vs. 1 Hz rTMS protocols using iAF (54)], which may be a more realistic and feasible approach for clinical practice than precision psychiatry (69).

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 4 robust biomarkers over the past decade presents a promising outlook and justifies large trials, similar to iSPOT-D and EMBARC for antidepressant medication, but aimed at rTMS and NIBS instead.

ARTICLE INFORMATION

From the Ghent Experimental Psychiatry Laboratory, Department of Head and Skin, Faculty of Medicine and Health Sciences, Ghent University, Ghent, Belgium (DK, CB); 4BRAIN Team, Department of Head and Skin, Faculty of Medicine and Health Sciences, Ghent University, Ghent, Belgium (DK); Center for Care and Cure, Department of Electrical Engineering, Eindhoven University of Technology, Eindhoven, the Netherlands (DK, CB); Research Institute Brainclinics, Brainclinics Foundation, Nijmegen, the Netherlands (HV, MA); Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, the Netherlands (HV,

MA); and Vrije Universiteit Brussel, Universitair Ziekenhuis Brussel, Department of Psychiatry, Brussels, Belgium (CB).

DK and HV contributed equally to this work as joint first authors.

Address correspondence to Martijn Arns, Ph.D., at martijn@brainclinics.com.

REFERENCES

1. Williams LM, Rush AJ, 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 12:4.

2. Trivedi MH, McGrath PJ, Fava M, Parsey RV, Kurian BT, Phillips ML, et al. (2016): Establishing moderators and biosignatures of antidepressant response in clinical care (EMBARC): Rationale and design. J Psychiatr Res 78:11–23.

3. 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.

4. Cohen SL, Bikson M, Badran BW, George MS (2022): A visual and narrative timeline of US FDA milestones for Transcranial Magnetic Stimulation (TMS) devices. Brain Stimul 15:73–75.

5. Blumberger DM, Vila-Rodriguez F, Thorpe KE, Feffer K, Noda Y, Giacobbe P, et al. (2018): Effectiveness of theta burst versus high-frequency repetitive transcranial magnetic stimulation in patients with depression (THREE-D): A randomised non-inferiority trial. Lancet 391:1683–1692.

6. 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 12:1572–1578.

7. 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 (TDBRAIN) database. Sci Data 9:333.

8. 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 5:1707–1716.

9. Spellman T, Liston C (2020): Toward circuit mechanisms of pathophysiology in depression. Am J Psychiatry 177:381–390.

10. 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 72:229–240.

11. Cohen SE, Zantvoord JB, Wezenberg BN, Bockting CLH, van Wingen GA (2021): Magnetic resonance imaging for individual prediction of treatment response in major depressive disorder: A systematic review and meta-analysis. Transl Psychiatry 11:168.

12. Widge AS, Bilge MT, Montana R, Chang W, Rodriguez CI, Deckersbach T, et al. (2019): Electroencephalographic biomarkers for treatment response prediction in major depressive illness: A meta-analysis. Am J Psychiatry 176:44–56.

13. Faul F, Erdfelder E, Lang AG, Buchner A (2007): G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. Behav Res Methods 39:175–191.

14. Baeken C, Brem AK, Arns M, Brunoni AR, Filipčić I, Ganho-Ávila A, et al. (2019): Repetitive transcranial magnetic stimulation treatment for depressive disorders: Current knowledge and future directions. Curr Opin Psychiatry 32:409–415.

15. O'Reardon JP, Solvason HB, Janicak PG, Sampson S, Isenberg KE, Nahas Z, et al. (2007): Efficacy and safety of transcranial magnetic stimulation in the acute treatment of major

depression: A multisite randomized controlled trial. Biol Psychiatry 62:1208–1216.

16. Boes AD, Uitermarkt BD, Albazron FM, Lan MJ, Liston C, Pascual-Leone A, et al. (2018): Rostral anterior cingulate cortex is a structural correlate of repetitive TMS treatment response in depression. Brain Stimul 11:575–581.

17. Baeken C, van Beek V, Vanderhasselt MA, 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 24:938–949.

18. Fox MD, Buckner RL, White MP, Greicius MD, Pascual-Leone A (2012): Efficacy of transcranial magnetic stimulation targets for depression is related to intrinsic functional connectivity with the subgenual cingulate. Biol Psychiatry 72:595–603.

19. Weigand A, Horn A, Caballero R, Cooke D, Stern AP, Taylor SF, et al. (2018): Prospective validation that subgenual connectivity predicts antidepressant efficacy of transcranial magnetic stimulation sites. Biol Psychiatry 84:28–37.

20. Cash RFH, Zalesky A, Thomson RH, Tian Y, Cocchi L, Fitzgerald PB (2019): Subgenual functional connectivity predicts antidepressant treatment response to transcranial magnetic stimulation: Independent validation and evaluation of personalization. Biol Psychiatry 86:e5–e7.

21. Siddiqi SH, Weigand A, Pascual-Leone A, Fox MD (2021): Identification of personalized transcranial magnetic stimulation targets based on subgenual cingulate connectivity: An independent replication. Biol Psychiatry 90:e55–e56.

22. 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 Psychiatry 180:230–240.

23. Baeken C, Marinazzo D, Wu GR, Van Schuerbeek P, De Mey J, Marchetti I, et al. (2014): Accelerated HF-rTMS in treatment-resistant unipolar depression: Insights from subgenual anterior cingulate functional connectivity. World J Biol Psychiatry 15:286–297.

24. Baeken C, Duprat R, Wu GR, De Raedt RD, 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? Biol Psychiatry Cogn Neurosci Neuroimaging 2:556–565.

25. Hopman HJ, Chan SMS, Chu WCW, Lu H, Tse CY, Chau SWH, et al. (2021): Personalized prediction of transcranial magnetic stimulation clinical response in patients with treatment-refractory depression using neuroimaging biomarkers and machine learning. J Affect Disord 290:261–271.

26. 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. Behav Brain Res 394:112834.

27. Ge R, Downar J, Blumberger DM, Daskalakis ZJ, 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 Stimul 13:206–214.

28. Salomons TV, Dunlop K, Kennedy SH, Flint A, Geraci J, Giacobbe P, Downar J (2014): Resting-state cortico-thalamic-striatal connectivity predicts response to dorsomedial prefrontal rTMS in major depressive disorder. Neuropsychopharmacology 39:488–498.

29. Cash RFH, Cocchi L, Lv J, Fitzgerald PB, Zalesky A (2021): Functional magnetic resonance imaging–guided personalization of transcranial magnetic stimulation treatment for depression. JAMA Psychiatry 78:337–339.

30. 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.

31. Siddiqi SH, Philip NS (2023): Hitting the target of image-guided psychiatry? Am J Psychiatry 180:185–187.

32. Lynch CJ, Power JD, Scult MA, Dubin M, Gunning FM, Liston C (2020): Rapid precision functional mapping of individuals using multi-echo fMRI. Cell Rep 33:108540.

33. Boes AD, Prasad S, Liu H, Liu Q, Pascual-Leone A, Caviness Jr VS, Fox MD (2015): Network localization of neurological symptoms from focal brain lesions. Brain 138:3061–3075.

34. Cash RFH, Müller VI, Fitzgerald PB, Eickhoff SB, Zalesky A (2023): Altered brain activity in unipolar depression unveiled using connectomics. Nat Mental Health 1:174–185.

35. 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.

36. Dinga R, Schmaal L, Penninx BWJH, van Tol MJ, Veltman DJ, van Velzen L, et al. (2019): Evaluating the evidence for biotypes of depression: Methodological replication and extension of Drysdale et al. NeuroImage Clin 22:101796.

37. Dinga R, Schmaal L, Marquand AF (2020): A Closer Look at Depression Biotypes: Correspondence Relating to Grosenick et al. (2019). Biol Psychiatry Cogn Neurosci Neuroimaging 5:554–555.

38. Grosenick L, Liston C (2020): Reply to: A closer look at depression biotypes: Correspondence relating to Grosenick et al. (2019). Biol Psychiatry Cogn Neurosci Neuroimaging 5:556.

39. Arns M, Drinkenburg WH, Fitzgerald PB, Kenemans JL (2012): Neurophysiological predictors of non-response to rTMS in depression. Brain Stimul 5:569–576.

40. 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.

41. Li CT, Hsieh JC, Huang HH, Chen MH, Juan CH, Tu PC, et al. (2016): Cognition-modulated frontal activity in prediction and augmentation of antidepressant efficacy: A randomized controlled pilot study. Cereb Cortex 26:202–210.

42. Li CT, Cheng CM, Juan CH, Tsai YC, Chen MH, Bai YM, et al. (2021): Task-modulated brain activity predicts antidepressant responses of prefrontal repetitive transcranial magnetic stimulation: A randomized sham-control study. Chronic Stress (Thousand Oaks) 5: 24705470211006855.

43. Wu W, Zhang Y, Jiang J, Lucas MV, Fonzo GA, Rolle CE, et al. (2020): An electroencephalographic signature predicts antidepressant response in major depression. Nat Biotechnol 38:439–447.

44. Nilsonne G, Harrell FE (2021): EEG-based model and antidepressant response. Nat Biotechnol 39:27.

45. Meijs H, Prentice A, Lin BD, De Wilde B, Van Hecke J, Niemegeers P, et al. (2022): A polygenic-informed approach to a predictive EEG signature empowers antidepressant treatment prediction: A proof-of-concept study. Eur Neuropsychopharmacol 62:49–60.

46. Posthuma D, Neale MC, Boomsma DI, de Geus EJC (2001): Are smarter brains running faster? Heritability of alpha peak frequency, IQ, and their interrelation. Behav Genet 31:567–579.

47. Smit DJA, Posthuma D, Boomsma DI, de Geus EJC (2005): Heritability of background EEG across the power spectrum. Psychophysiology 42:691–697.

48. van Beijsterveldt CEM, van Baal GCM (2002): Twin and family studies of the human electroencephalogram: A review and a meta-analysis. Biol Psychol 61:111–138.

49. Arns M, Spronk D, Fitzgerald PB (2010): Potential differential effects of 9 Hz rTMS and 10 Hz rTMS in the treatment of depression. Brain Stimul 3:124–126.

50. Widge AS, Avery DH, Zarkowski P (2013): Baseline and treatment-emergent EEG biomarkers of antidepressant medication response do not predict response to repetitive transcranial magnetic stimulation. Brain Stimul 6:929–931.

51. Arns M, Cerquera A, Gutiérrez RM, Hasselman F, Freund JA (2014): Non-linear EEG analyses predict non-response to rTMS treatment in major depressive disorder. Clin Neurophysiol 125:1392–1399.

52. Roelofs CL, Krepel N, Corlier J, Carpenter LL, Fitzgerald PB, Daskalakis ZJ, et al. (2021): Individual alpha frequency proximity associated with repetitive transcranial magnetic stimulation outcome: An independent replication study from the ICON-DB consortium. Clin Neurophysiol 132:643–649.

53. Voetterl H, van Wingen G, Michelini G, Griffiths KR, Gordon E, DeBeus R, et al. (2023): Brainmarker-I differentially predicts remission

to various attention-deficit/hyperactivity disorder treatments: A discovery, transfer, and blinded validation study. Biol Psychiatry Cogn Neurosci Neuroimaging 8(1):52–60.

54. Voetterl HTS, Sack A, Olbrich S, Stuiver S, Rouwhorst R, Prentice A, et al. (2023): Alpha peak frequency-based Brainmarker-I as a method to stratify to pharmacotherapy and brain stimulation treatments in depression. Nature Mental Health. https://doi.org/10.1038/s44220-023-00160-7.

55. 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.

56. 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. Clin EEG Neurosci 46:73–80.

57. Hunter AM, Nghiem TX, Cook IA, Krantz DE, Minzenberg MJ, Leuchter AF (2018): Change in quantitative EEG theta cordance as a potential predictor of repetitive transcranial magnetic stimulation clinical outcome in major depressive disorder. Clin EEG Neurosci 49:306–315.

58. Erguzel TT, Ozekes S, Gultekin S, Tarhan N, Hizli Sayar GH, Bayram A (2015): Neural network based response prediction of rTMS in major depressive disorder using QEEG cordance. Psychiatry Investig 12:61–65.

59. Erguzel TT, Tarhan N (2016): Machine learning approaches to predict repetitive transcranial magnetic stimulation treatment response in major depressive disorder. In: Bi Y, Kapoor S, Bhatia R, editors. Proceedings of SAI Intelligent Systems Conference (IntelliSys) 2016. Cham: Springer, 391–401.

60. Ho SY, Phua K, Wong L, Bin Goh WWB (2020): Extensions of the external validation for checking learned model interpretability and generalizability. Patterns (N Y) 1:100129.

61. Hasanzadeh F, Mohebbi M, Rostami R (2019): Prediction of rTMS treatment response in major depressive disorder using machine learning techniques and nonlinear features of EEG signal. J Affect Disord 256:132–142.

62. Zhang Y, Wu W, Toll RT, Naparstek S, Maron-Katz A, Watts M, et al. (2021): Identification of psychiatric disorder subtypes from functional connectivity patterns in resting-state electroencephalography. Nat Biomed Eng 5:309–323.

63. Bailey NW, Hoy KE, Rogasch NC, Thomson RH, McQueen S, Elliot D, et al. (2019): Differentiating responders and non-responders to rTMS treatment for depression after one week using resting EEG connectivity measures. J Affect Disord 242:68–79.

64. Bailey NW, Krepel N, van Dijk H, Leuchter AF, Vila-Rodriguez F, Blumberger DM, et al. (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. Clin Neurophysiol 132:650–659.

65. Pizzagalli DA (2011): Frontocingulate dysfunction in depression: Toward biomarkers of treatment response. Neuropsychopharmacology 36:183–206.

66. 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 Psychiatry 75:547–554.

67. Cash RFH, Weigand A, Zalesky A, Siddiqi SH, Downar J, Fitzgerald PB, Fox MD (2021): Using brain imaging to improve spatial targeting of transcranial magnetic stimulation for depression. Biol Psychiatry 90:689–700.

68. van der Vinne N, Vollebregt MA, Rush AJ, Eebes M, van Putten MJAM, Arns M (2021): EEG biomarker informed prescription of antidepressants in MDD: A feasibility trial. Eur Neuropsychopharmacol 44:14–22.

69. Arns M, van Dijk H, Luykx JJ, van Wingen G, Olbrich S (2022): Stratified psychiatry: Tomorrow's precision psychiatry? Eur Neuropsychopharmacol 55:14–19.