Alpha peak frequency-based Brainmarker-I as a method to stratify to pharmacotherapy and brain stimulation treatments in depression

Helena T. S. Voetterl [1,2] ✉, Alexander T. Sack [2,3], Sebastian Olbrich [4], Sven Stuiver [5], Renee Rouwhorst[6], Amourie Prentice [1,2,7], Diego A. Pizzagalli [8,9], Nikita van der Vinne [1,2,7], Jeroen A. van Waarde [5], Martin Brunovsky [10,11], Iris van Oostrom[12], Ben Reitsma[13], Johan Fekkes[14], Hanneke van Dijk[1,2,7,15] & Martijn Arns [1,2,15] ✉

Biomarkers predicting treatment outcome in major depressive disorder could enhance clinical improvement. Here this observational and prospective accuracy study investigates whether an age- and sex-normalized electroencephalography biomarker, based on the individual alpha frequency (iAF), can successfully stratify patients to different interventions such as repetitive transcranial magnetic stimulation (rTMS) and electroconvulsive therapy (ECT). Differential iAF directions were explored for sertraline, as well as rTMS ($N = 196$) and ECT ($N = 41$). A blinded out-of-sample validation (EMBARC; $N = 240$) replicated the previously found association between low iAF and better sertraline response. The subgroup of patients with an iAF around 10 Hz had a higher remission rate following 10 Hz rTMS compared with the group level, while the high-iAF subgroup had highest remission to 1 Hz rTMS and the low-iAF subgroup to ECT. Blinded out-of-sample validations for 1 Hz ($N = 39$) and ECT ($N = 51$) corroborated these findings. The present study suggests a clinically actionable electroencephalography biomarker that can successfully stratify between various antidepressant treatments.

Major depressive disorder (MDD) is one of the most common and debilitating disorders worldwide[1]. The disorder's high level of heterogeneity (in both symptoms and neurophysiology) complicates adequate treatment prescription, which may limit treatment response[2–4]. 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[5], with response rates around 40–50% and remission rates around 30–40% (ref. [6]).

While targeting the patient's individual neurophysiology (for example, precision psychiatry) seems to be infeasible at present,

[1]Research Institute Brainclinics, Brainclinics Foundation, Nijmegen, the Netherlands. [2]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, the Netherlands. [3]Department of Psychiatry and Neuropsychology, School for Mental Health and Neuroscience, Brain+Nerve Centre, Maastricht University Medical Centre+, Maastricht, the Netherlands. [4]Department for Psychiatry, Psychotherapy and Psychosomatics, University Hospital of Psychiatry Zurich, Zurich, Switzerland. [5]Rijnstate Hospital, Arnhem, the Netherlands. [6]neurocare clinics, Amsterdam & Den Haag, The Hague, the Netherlands. [7]Synaeda Psycho Medisch Centrum, Leeuwarden, the Netherlands. [8]Department of Psychiatry, Harvard Medical School, Boston, MA, USA. [9]McLean Hospital, Belmont, MA, USA. [10]National Institute of Mental Health, Klecany, Czech Republic. [11]Department of Psychiatry and Medical Psychology, Third Faculty of Medicine, Charles University, Prague, Czech Republic. [12]neuroCare Clinics, Nijmegen, the Netherlands. [13]Brain Dynamics Groningen, Groningen, the Netherlands. [14]Brain Dynamics Hengelo, Hengelo, the Netherlands. [15]These authors jointly supervised this work: Hanneke van Dijk, Martijn Arns. ✉e-mail: helena@brainclinics.com; martijn@brainclinics.com
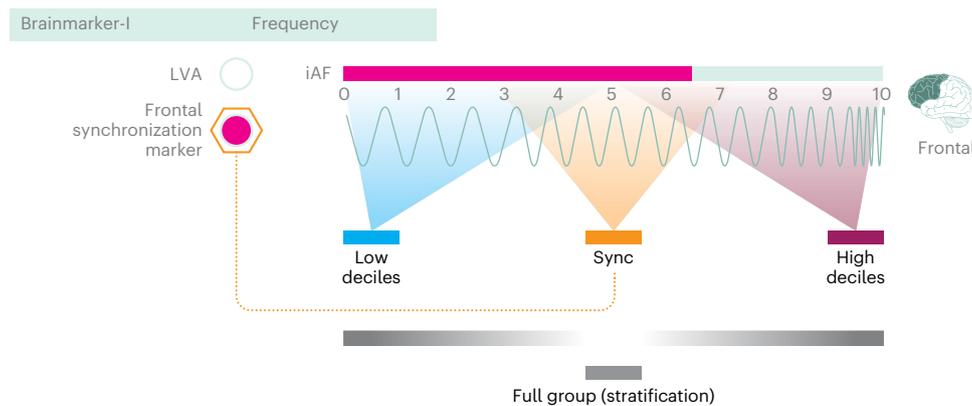
Fig. 1 | Visualization of the Brainmarker-I classification. A filled, pink dot on the left denotes either that the patient has low voltage alpha or that their iAF falls into the frontal synchronization range (9.6–10.4 Hz) (depicted above). The iAF is depicted in Brainmarker-I decile scores from 1 (relatively 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.

an implementable alternative is treatment stratification (for a discussion, see ref. 6), which reduces heterogeneity within a disorder by identifying subgroups of patients that preferentially respond to a certain treatment, using so-called biomarkers[7,8]. A nonrandomized, open-label study, based on resting-state electroencephalography (EEG) biomarkers, prospectively stratified between three antidepressants in MDD that resulted in better clinical outcomes relative to treatment-as-usual[7]. 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[8,9]. 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[10].

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[11–14], while slower iAF has been associated with higher symptom severity[15,16] and less favorable treatment outcome[17–19] and has been observed across many disorders such as Alzheimer's disease[20], burnout syndrome[21], mild cognitive impairment[22], psychosis[23], schizophrenia[23,24] and attention-deficit hyperactivity disorder (ADHD)[25], with this slowing potentially reflecting reduced thalamocortical information transfer.

In patients with ADHD, slower iAF has been related to worse treatment outcome to methylphenidate[26] and better treatment outcome to multimodal neurofeedback[27]. 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[28]. For antidepressant medication (amitriptyline and pirlindole), a slow iAF was shown to be predictive of nonresponse[29]. 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[30]. For repetitive transcranial magnetic stimulation (rTMS), a different association has been observed. Specifically, an iAF closer to 10 Hz (that is, the rTMS stimulation frequency) was associated with better improvement to 10-Hz left-dorsolateral prefrontal cortex (L-DLPFC) rTMS[31], which was independently replicated, while no association emerged between iAF and outcome of 1-Hz right-DLPFC (R-DLPFC) rTMS[32]. For electroconvulsive therapy (ECT), to our knowledge, iAF prediction of treatment outcome is unknown.

In this Article, following these promising findings, we aimed to extend Brainmarker-I, developed for ADHD treatment stratification[28], to treatments for MDD. We decided a priori to conduct statistical analyses in line with Voetterl et al.[28] and the hypotheses outlined below, focusing only on remission as our primary outcome, given its higher clinical relevance and to avoid multiple testing. We first conducted a blinded out-of-sample validation in the double-blind placebo-controlled Establishing Moderators and Biosignatures of Antidepressant Response for Clinical Care (EMBARC) dataset[33,34], 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 priori considered an indication for 10 Hz L-DLPFC rTMS, based on the above-mentioned replicated research[31,32]. 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.

Resting-state eyes-closed EEG data were preprocessed for all datasets, in line with previous preprocessing[35]. The iAF was calculated in accordance with Voetterl et al.[28], 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 Hz and 10.4 Hz at the F3 location was implemented (Fig. 1) to mark close proximity to 10 Hz, resulting in three distinct biomarker subgroups that were compared for the different treatments: synchronization, low deciles (decile score 1–5 without synchronization range) and high deciles (decile score 6–10 without synchronization range). Positive predictive values (PPVs) indicated the remission rate within each Brainmarker-I subgroup. A normalized PPV (nPPV) was calculated to be able to compare remission rates that differed between datasets. In short, the respective remission

Table 1 | Basic demographic information EMBARC dataset

Treatment arm	Sertraline	Placebo	Sertraline	Placebo	Sertraline–bupropion
Time point 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

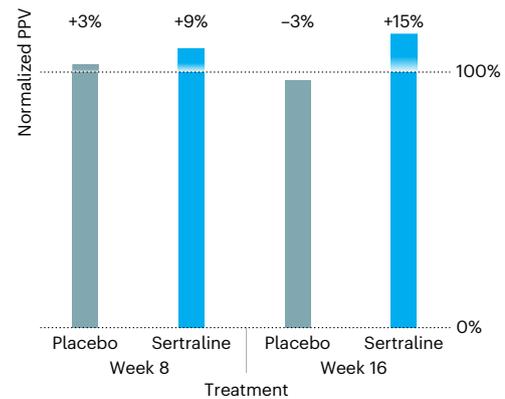Basic demographic information for different treatment arms in EMBARC Dataset-1.

Fig. 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.

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 with treating patients with the same active treatments but in a random fashion (not informed by the biomarker).

Results

Basic information about the different datasets is provided in Table 1 and Table 2. Remission rates of each dataset and treatment group are summarized in Supplementary Table 1.

Blinded sertraline replication

Results for the EMBARC dataset are visualized in Fig. 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 with 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 visualized in Fig. 3. Full results of analyses in all three biomarker subgroups can be found in Supplementary Table 2.

In line with previous evidence[31,32], the remission rate in the synchronization subgroup (iAF between 9.6 Hz and 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. Therefore, 10 Hz rTMS was 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 three different subgroups across all included rTMS and ECT datasets differed (Discussion). For low-decile, synchronization and high-decile subgroup, these were 47%, 30% and 23%, respectively.

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 nonstratified 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 visualized in Supplementary Fig. 1.

Confounding factors analyses

To ascertain the presented findings were not related to differences in depression severity, we conducted one-way analyses of variance
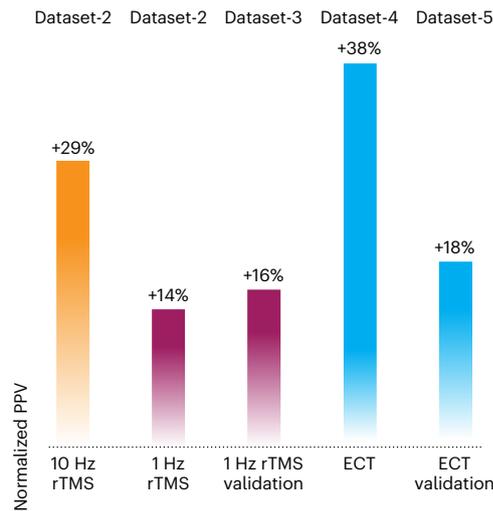
Table 2 | Basic demographic information all datasets

Full datasets	Dataset-2	Dataset-2	Dataset-3	Dataset-4	Dataset-5	Dataset-6	Dataset-7
	10 Hz L-DLPFC rTMS	1 Hz R-DLPFC rTMS	1 Hz R-DLPFC rTMS	ECT	ECT	Psychotherapy	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, (s.d.)	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)

Basic demographic information for all remaining datasets.

Fig. 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.

between the three 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 ($P > .147$).

Discussion

The present study successfully extends the previously introduced Brainmarker-I 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[30], 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[30] 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%), probably due to the high placebo remission rate of 29% that did not differ from the week-8 sertraline remission rate of 32%. It is known that placebo response can be substantial in antidepressant trials[36,37] and a diminished response to the active antidepressant treatment has been reported in studies that include a placebo arm[38]. 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[31] and replicated[32]. We quantified this finding by determining the Brainmarker-I synchronization subgroup in Dataset-2, which showed an increased normalized remission rate of +29% to 10 Hz rTMS compared with 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[31,39].

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

normalized 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[21–24,40,41], 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 versus 67%, respectively)[42]. 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.[42]. Interestingly, in a secondary analysis (Supplementary Discussion 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[6]. 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 with the alternative one-size-fits-all approach, no harm is done by using stratification (for a more in-depth explanation, see ref. 6).

In this manuscript, we present a stratification solution for difficult-to-treat depression, based partially on previous findings (for example, 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 nine patients need to be treated with the biomarker-recommended treatment to have one more patient remit compared with active treatment prescribed in an arbitrary way. This NNT is close to the effect of tricyclic antidepressant and SSRI monotherapy (minimum NNT 7)[43] compared with placebo. This is rather impressive, considering that the simulated stratified remission rate was not compared with 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 health care 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 cutoff 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.[44]. Similarly, EEG parameters and amplifiers differed across collection locations, resulting in a total of six 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 with 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 three stratification subgroups (Supplementary Table 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 9.8 Hz is already considered to fall in the upper alpha range, that is, fast alpha[45], making the synchronization marker (9.6–10.4 Hz) overlap more with the higher-decile subgroup.

One limitation linked to the use of an nPPV 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, additional 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

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, Dataset-3, Dataset-4 and Dataset-6 were recorded in a standardized manner in accordance with Brain Resource Ltd.[35]. 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-min resting-state recordings (2 min eyes open, 2 min eyes closed). Sampling frequency (FS) was 500 Hz, and a low-pass filter with an attenuation of 40 dB per decade above 100 Hz was applied before 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 kΩ for all electrodes.

Artifact rejection was performed with a fully automated, custom Python package[46–49].

In short, bipolar EOG was removed from the EEG signal using Gratton[50]. A band-pass filter between 0.5 Hz and 100 Hz was applied, and the notch frequency of 50 Hz was removed. The following artifacts 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 artifactual, it was repaired using a Euclidian distance weighted average of at least three neighboring channels. If neighboring channels were not available due to artifactual data, the channel was removed. Very artifactual data were excluded on the basis of visual inspection.

For full details on preprocessing, see van Dijk et al.[51]. 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 artifact rejection for Datasets-1, -5 and -7 were performed in Brain Vision Analyzer version 2.2.0 (Brain Products GmbH) by semi-automatic removal of epochs with signal amplitudes >150 mV.

For Dataset-1 (EMBARC)[33] 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 downsampled to the lowest FS (250 Hz), and electrodes were adjusted to match the 26 locations listed above.

ECT Dataset-5 (ref. 52) was treated accordingly, resulting in an FS of 200 Hz and 19 channel locations (FC3, FCz, FC4, CP3, CPz, CP4 and Oz missing).

Similarly, the ketamine Dataset-7 (ref. 53) combined three 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.[28], 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, Dataset-3 and Dataset-6), ≤7 on the HRSD-17 (for Dataset-1 and Dataset-4), ≤2 on the Clinical Global Impression ratings (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.[44], except for the HRSD-17 cutoff, which was based on the original sertraline study[30], 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[28]. The biomarker was developed in a large heterogeneous clinical dataset (TDBRAIN+; *N* = 4,249). A subset of the data, the open-access TDBRAIN dataset (N = 1,274; two decades brainclinics research archive for insights in neurophysiology), is freely available at http://www.brainclinics.com/resources ref. 53 after login, with all data recorded at Research Institute Brainclinics

(Brainclinics Foundation, Nijmegen, the Netherlands). In addition, the data are available on the data repository Synapse at ref. 54.

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.[28], 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 1,671 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 >4,000 individuals were sorted and divided into ten equal-sized bins that 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 2 and 3.

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 5 s and re-referenced to an average reference, based on previous literature[28,31].

The highest peak within the frequency range of 7–13 Hz was identified at the 10–20 EEG system locations F3 and Fz, in line with previous predictions[28,31,32]. 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 cutoff values determined in the large TDBRAIN+ dataset. Treatment predictions were made on the basis of low (decile 1–5) or high deciles (decile 6–10) in the Fz electrode. For an example, see Fig. 1.

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. 1). To determine the optimal range for this third biomarker subgroup, we tested different cutoff values that were equidistant from the 10 Hz frequency. Due to the frequency resolution of 0.2 Hz, 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 three distinct subgroups.

The automated algorithm described in Voetterl et al.[28] was used to calculate iAF and decile scores for individuals of all datasets (Fig. 1).

Statistics

PPVs indicate the remission rate within the subsample of patients that Brainmarker-I would have stratified to the respective treatment. An nPPV was calculated to be able to compare predicted remission rates of different datasets, using the formula $\left(\frac{m}{w}-1\right)\times100$ (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.

In addition, 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.

To test whether potential findings could be explained by differences in depression severity, we conducted one-way analysis of variance analyses 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. Biomarker calculation was conducted in Python, using modules numpy[55], pandas[49] and scipy[47]. 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 1 and Table 2. All studies were approved by their respective institutional review boards (with ethical approval numbers available in the primary publications of the studies).

Dataset-1: EMBARC Sertraline

The EMBARC data were precollected data that were specifically requested for secondary analyses (for information on ethical approval, CONSORT diagrams, study protocol and participant inclusion, we refer the reader to the relevant references)[33,34]. 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 29 July 2011 and 15 December 2015, outpatients were recruited at four sites: Columbia University, New York; Massachusetts General Hospital, Boston; University of Michigan, Ann Arbor; and University of Texas Southwestern Medical Center, Dallas. A total of 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 nonresponders to a different medication (sertraline for placebo nonresponders and bupropion for sertraline nonresponders) 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 outcome, we calculated nPPVs after prolonged sertraline or placebo administration (week 16).

Dataset-2 and Dataset-3: rTMS

Dataset-2 and Dataset-3 are open-label, clinical datasets composed of patient data collected at multiple outpatient mental health care clinics in the Netherlands (Brainclinics Treatment, 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 before collection of the EEG data. In rTMS Dataset-2, 196 patients with MDD received 10 Hz rTMS over the L-DLPFC or 1 Hz rTMS over the R-DLPFC (at 120% resting motor threshold, 1,500 or 1,200 pulses, respectively) concurrent with psychotherapy[32]. 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 Dataset-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 have been used before 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[56]) stimulation with stimulus dose relative to seizure threshold (SDRST; that is, 6 times seizure threshold (ST) in RUL and 2.5 times ST in BL ECT) and using 0.5 ms pulse width. Resting-state EEG data and HRSD-17-score were collected before ECT and 2 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 have 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 (pulse width 0.5 ms, SDRST 1.5 of ST)[52], with outcome analyzed by Clinical Global Impression ratings.

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[32], 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 and 10.4) who had received 10 Hz rTMS.

For 1 Hz and ECT datasets, all possible directions of effect were tested, that is, low decile score (1–5) excluding synchronization range, synchronization range, and high decile score (6–10) excluding 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 Dataset-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 before 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 before 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[53]).

Exploratory analyses were performed in Dataset-6 for psychotherapy[57], in Dataset-7 for ketamine treatment and in Dataset-1 for the subgroup of sertraline nonresponders, 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 4.

Clinical trials

Data from the following trials were used in this study: Establishing Moderators and Biosignatures of Antidepressant Response for Clinical Care for Depression (EMBARC[58], identifier NCT01407094), QEEG Cordance and EEG Connectivity Changes after Administration of Subanesthetic Ketamine Doses in Patients with Depressive Disorder[59], The Role of mTOR (Mammalian Target of Rapamycin) Signaling Pathway in the Antidepressive Effect of Ketamine in Patients with Depressive Disorder[60] and Clinical and Neurobiological Predictors of Response to Ketamine: towards Personalized Treatment of Depression[61].

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The TDBRAIN EEG data are 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/.

References

1. WHO. *The Global Burden of Disease: 2004 Update* (WHO Press, 2008).
2. Drysdale, A. T. et al. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat. Med.* **23**, 28–38 (2017).
3. GOLDBERG, D. The heterogeneity of 'major depression'. *World Psychiatry* **10**, 226–228 (2011).
4. Luo, Y., Weibman, D., Halperin, J. M. & Li, X. A review of heterogeneity in attention deficit/hyperactivity disorder (ADHD). *Front. Hum. Neurosci.* **13**, 42 (2019).
5. Rush, A. J. et al. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *Am. J. Psychiat.* **163**, 1905–1917 (2006).
6. Arns, M., Dijk, H., van, Luykx, J. J., Wingen, Gvan & Olbrich, S. Stratified psychiatry: tomorrow's precision psychiatry? *Eur. Neuropsychopharm.* **55**, 14–19 (2022).
7. Vinne, N. et al. EEG biomarker informed prescription of antidepressants in MDD: a feasibility trial. *Eur. Neuropsychopharm.* **44**, 14–22 (2021).

8. Olbrich, S., Dinteren, Rvan & Arns, M. Personalized medicine: review and perspectives of promising baseline EEG biomarkers in major depressive disorder and attention deficit hyperactivity disorder. *Neuropsychobiology* **72**, 229–240 (2015).

9. Olbrich, S. & Arns, M. EEG biomarkers in major depressive disorder: discriminative power and prediction of treatment response. *Int. Rev. Psychiatr.* **25**, 604–618 (2013).

10. Widge, A. S. et al. Electroencephalographic biomarkers for treatment response prediction in major depressive illness: a meta-analysis. *Am. J. Psychiatry* **176**, 44–56 (2019).

11. Clark, C. R. et al. Spontaneous alpha peak frequency predicts working memory performance across the age span. *Int. J. Psychophysiol.* **53**, 1–9 (2004).

12. Pahor, A. & Jaušovec, N. Making brains run faster: are they becoming smarter? *Span. J. Psychol.* **19**, E88 (2016).

13. Grandy, T. H. et al. Individual alpha peak frequency is related to latent factors of general cognitive abilities. *NeuroImage* **79**, 10–18 (2013).

14. Klimesch, W. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Rev.* **29**, 169–195 (1999).

15. Struve, F. A. & Boutros, N. N. Somatic implications of generalized and/or focal EEG slowing in psychiatric patients. *Clin. EEG Neurosci.* **36**, 171–175 (2005).

16. Boutros, N. N. Diffuse electroencephalogram slowing in psychiatric patients: a preliminary report. *J. Psychiatry Neurosci. Jpn* **21**, 259–263 (1996).

17. Arns, M., Gunkelman, J., Breteler, M. & Spronk, D. EEG phenotypes predict treatment outcome to stimulants in children with ADHD. *J. Integr. Neurosci.* **07**, 421–438 (2008).

18. Martijn, A., Desirée, S. & Fitzgerald, P. B. Potential differential effects of 9 Hz rTMS and 10 Hz rTMS in the treatment of depression. *Brain Stimul.* **3**, 124–126 (2010).

19. Arns, M. EEG-based personalized medicine in ADHD: individual alpha peak frequency as an endophenotype associated with nonresponse. *J. Neurother.* **16**, 123–141 (2012).

20. Rodriguez, G., Copello, F., Vitali, P., Perego, G. & Nobili, F. EEG spectral profile to stage Alzheimer's disease. *Clin. Neurophysiol.* **110**, 1831–1837 (1999).

21. Luijtelaar, G., van, Verbraak, M., Bunt, M., van den, Keijsers, G. & Arns, M. EEG findings in Burnout patients. *J. Neuropsychiatry Clin. Neurosci.* **22**, 208–217 (2010).

22. Garcés, P. et al. Brain-wide slowing of spontaneous alpha rhythms in mild cognitive impairment. *Front. Aging Neurosci.* **5**, 100–100 (2013).

23. Murphy, M. & Öngür, D. Decreased peak alpha frequency and impaired visual evoked potentials in first episode psychosis. *NeuroImage Clin.* **22**, 101693 (2019).

24. Yeum, T.-S. & Kang, U. G. Reduction in alpha peak frequency and coherence on quantitative electroencephalography in patients with schizophrenia. *J. Korean Med. Sci.* **33**, e179 (2018).

25. Bazanova, O. M., Auer, T. & Sapina, E. A. On the efficiency of individualized Theta/Beta ratio neurofeedback combined with forehead EMG training in ADHD children. *Front. Hum. Neurosci.* **12**, 3 (2018).

26. Arns, M. et al. Electroencephalographic biomarkers as predictors of methylphenidate response in attention-deficit/hyperactivity disorder. *Eur. Neuropsychopharm.* **28**, 881–891 (2018).

27. Krepel, N. et al. A multicenter effectiveness trial of QEEG-informed neurofeedback in ADHD: replication and treatment prediction. *NeuroImage Clin.* https://doi.org/10.1016/j.nicl.2020.102399 (2020).

28. Voetterl, H. et al. Brainmarker-I differentially predicts remission to various attention-deficit/hyperactivity disorder treatments: a blinded discovery, transfer and validation study. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* https://doi.org/10.1016/j.bpsc.2022.02.007 (2022).

29. Ulrich, G., Renfordt, E., Zeller, G. & Frick, K. Interrelation between changes in the EEG and psychopathology under pharmacotherapy for endogenous depression. A contribution to the predictor question. *Pharmacopsychiatry* **17**, 178–183 (1984).

30. Arns, M., Gordon, E. & Boutros, N. N. EEG abnormalities are associated with poorer depressive symptom outcomes with escitalopram and venlafaxine-XR, but not sertraline. *Clin. EEG Neurosci.* **48**, 33–40 (2016).

31. Corlier, J. et al. 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 (2019).

32. Roelofs, C. et al. Individual alpha frequency proximity associated with repetitive transcranial magnetic stimulation outcome: An independent replication study from the ICON-DB consortium. *Clin. Neurophysiol.* https://doi.org/10.1016/j.clinph.2020.10.017 (2020).

33. Pizzagalli, D. A. et al. Pretreatment rostral anterior cingulate cortex theta activity in relation to symptom improvement in depression: a randomized clinical trial. *JAMA Psychiatry* **75**, 547 (2018).

34. Trivedi, M. H. et al. Establishing moderators and biosignatures of antidepressant response in clinical care (EMBARC): rationale and design. *J. Psychiatr. Res.* **78**, 11–23 (2016).

35. Arns, M. et al. EEG alpha asymmetry as a gender-specific predictor of outcome to acute treatment with different antidepressant medications in the randomized iSPOT-D study. *Clin. Neurophysiol.* **127**, 509–519 (2016).

36. Walsh, B. T., Seidman, S. N., Sysko, R. & Gould, M. Placebo response in studies of major depression: variable, substantial, and growing. *JAMA* **287**, 1840–1847 (2002).

37. Brunoni, A. R., Lopes, M., Kaptchuk, T. J. & Fregni, F. Placebo response of non-pharmacological and pharmacological trials in major depression: a systematic review and meta-analysis. *PLoS ONE* **4**, e4824 (2009).

38. Sinyor, M. et al. Does inclusion of a placebo arm influence response to active antidepressant treatment in randomized controlled trials?: Results from pooled and meta-analyses. *J. Clin. Psychiatry* **71**, 270–279 (2010).

39. Huang, W. A. et al. Transcranial alternating current stimulation entrains alpha oscillations by preferential phase synchronization of fast-spiking cortical neurons to stimulation waveform. *Nat. Commun.* **12**, 3151 (2021).

40. Ramsay, I. S., Lynn, P. A., Schermitzler, B. & Sponheim, S. R. Individual alpha peak frequency is slower in schizophrenia and related to deficits in visual perception and cognition. *Sci Rep.* **11**, 17852 (2021).

41. Dickinson, A., DiStefano, C., Senturk, D. & Jeste, S. S. Peak alpha frequency is a neural marker of cognitive function across the autism spectrum. *Eur. J. Neurosci.* **47**, 643–651 (2018).

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

43. Arroll, B. et al. Antidepressants versus placebo for depression in primary care. *Cochrane Database Syst. Rev.* https://doi.org/10.1002/14651858.cd007954 (2009).

44. Riedel, M. et al. Response and remission criteria in major depression—a validation of current practice. *J. Psychiatr. Res.* **44**, 1063–1068 (2010).

45. Sauseng, P. et al. EEG alpha synchronization and functional coupling during top-down processing in a working memory task. *Hum. Brain Mapp.* **26**, 148–155 (2005).

46. Hunter, J. D. Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* **9**, 90–95 (2007).

47. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. Methods* **17**, 261–272 (2020).

48. Harris, C. R. et al. Array programming with NumPy. *Nature* **585**, 357–362 (2020).

49. pandas-dev/pandas: Pandas 1.0.3. *Zenodo* https://doi.org/10.5281/zenodo.3715232 (2020).

50. Gratton, G., Coles, M. G. H. & Donchin, E. A new method for off-line removal of ocular artifact. *Electroencephalogr. Clin. Neurophysiol.* **55**, 468–484 (1983).

51. Dijk, H. et al. The two decades brainclinics research archive for insights in neurophysiology (TDBRAIN) database. *Sci. Data* **9**, 333 (2022).

52. Kirsten, A., Seifritz, E. & Olbrich, S. Electroencephalogram source connectivity in the prediction of electroconvulsive therapy outcome in major depressive disorder. *Clin. EEG Neurosci.* **51**, 10–18 (2019).

53. Meyer, T. et al. Predictive value of heart rate in treatment of major depression with ketamine in two controlled trials. *Clin. Neurophysiol.* **132**, 1339–1346 (2021).

54. Two-Decades Brainclinics Archive for Insights in Neurophysiology (TD-BRAIN). *SYNAPSE* https://www.synapse.org/TDBRAIN (2022).

55. Gramfort, A. et al. MEG and EEG data analysis with MNE-Python. *Front. Neurosci.* **7**, 267 (2013).

56. D'Elia, G. in *Psychobiology of Convulsive Therapy* (ed. Fink, M., Kety, S., McGaugh, J. and Williams, T.) (VH Winston, 1974) pp.21–34.

57. Meijs, H. et al. A polygenic-informed approach to a predictive EEG signature empowers antidepressant treatment prediction: a proof-of-concept study. *Eur. Neuropsychopharm.* **62**, 49–60 (2022).

58. EMBARC clinical trial. *National Library of Medicine* http://clinicaltrials.gov/show/NCT01407094 (2011).

59. QEEG Cordance and EEG Connectivity Changes after Administration of Subanesthetic Ketamine Doses in Patients with Depressive Disorder. *EU Clinical Trials Register* https://www.clinicaltrialsregister.eu/ctr-search/trial/2009-010625-39/CZ (2010).

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

61. Clinical and Neurobiological Predictors of Response to Ketamine: towards Personalized Treatment of Depression. *EU Clinical Trials Register* https://www.clinicaltrialsregister.eu/ctr-search/trial/2018-001539-39/CZ (2018).

Acknowledgements

Author contributions

H.T.S.V. conceptualized the study, performed all statistical analyses, participated in developing the EEG processing pipelines and code, and drafted the manuscript. A.T.S. provided supervision, and substantively revised the manuscript. S.O., S.S., R.R., A.P., D.A.P., N.v.V., J.A.v.W., M.B., B.R. and J.F. were involved in collecting the data and reviewing the manuscript. I.v.O. was involved in conceptualizing the study and revised the manuscript. H.v.D. took part in conceptualizing the study, developed the EEG processing pipelines, substantively reviewed the manuscript, and provided supervision. M.A. conceptualized the study, participated in analyzing the data, substantively reviewed the manuscript and supervised all phases of the study.

Competing interests

The authors declare the following competing interests: A.T.S. 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). S.O. is cofounder of DeepPsy, a company for EEG-biomarker report systems. Over the past 3 years, D.A.P. 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 Pharmaceuticals, 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. M.B. 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'. M.A. 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 H.T.S.V., S.S., R.R., A.P., N.v.V., J.A.v.W., M.B., I.v.O., B.R., J.F. and H.v.D. report no conflicts of interest.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s44220-023-00160-7.

Correspondence and requests for materials should be addressed to Helena T. S. Voetterl or Martijn Arns.

Peer review information *Nature Mental Health* thanks Aron Hill, Albino Oliveira-Maia and the other, anonymous, reviewers for their contribution to the peer review of this work.

nature portfolio

Corresponding author(s): Helena Voetterl, Martijn Arns

Last updated by author(s): Oct 5, 2023

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	Neuroscan NuAmps amplifier (for data collection at Brainclinics Foundation)
Data analysis	A fully-automated, custom Python package was used for data cleaning and biomarker calculation (Python version 3.6; packages numpy version 1.19.2, pandas version 1.1.3, scipy version 1.5.2). This code is freely available for download at https://brainclinics.com/resources/. SPSS version 27; Brain Vision Analyzer version 2.2.0]

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:
- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

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 on request (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.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences   ☐ Behavioural & social sciences   ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Sample size: TD-BRAIN+ = 4249, Dataset-1 (EMBARC)= 240, Dataset-2 and -3= 187, Dataset-4 (SNOEP)= 41, Dataset-5 (University Hospital Zurich)=51, Dataset-6 (SYNAEDA) = 156, Dataset-7 (Prague Psychiatric Center)=81. Data from all samples used were pre-collected clinical samples of which all complete data was used.
Data exclusions	TD-BRAIN+: N=123 excluded (insufficient clean EEG data) Dataset-1 (EMBARC): N= 56 (insufficient clean EEG data, missing clinical data) Dataset-2 and -3: N= 9 (insufficient clean EEG data, missing clinical data) Dataset-4 (SNOEP): N= 31 (missing/insufficient clean EEG data, missing clinical data, baseline depression scores below inclusion criterion) Dataset-5 (University Hospital Zurich): N=0 Dataset-6 (SYNAEDA): N = 24 (insufficient clean EEG data, missing clinical data) Dataset-7 (Prague Psychiatric Center): N=0
Replication	The novel findings for 1 Hz rTMS and ECT were verified by blinded validation in 2 independent datasets (Dataset-3 and Dataset-5). The sertraline finding in Dataset-1 was a conceptual replication of a previous finding, however this time using iAF calculated with Brainmarker-1 as predictor. The 10 Hz rTMS finding had been tested and replicated before. A replication of the psychotherapy, bupropion and ketamine findings (Dataset-6, -1 and -7) was not attempted in the present manuscript; these findings are reported as preliminary and requiring replication.
Randomization	TD-BRAIN+: N/A (no experimental groups) Dataset-1 (EMBARC): patients were randomised to sertraline or placebo for 8 weeks, then treatment response was assessed and responders remained on the same drug regimen, while non-responders were switched to a different medication Dataset-2 and -3: patients were assigned to 10 Hz or 1 Hz rTMS according to clinical parameters and the current clinical practice at that time. Since the here presented biomarker was developed to predict outcome in a naturalistic, clinical setting, development in heterogeneous clinical datasets seemed justified. Dataset-4 (SNOEP): naturalistic, clinical dataset with no experimental groups, patients were assigned to different ECT protocols based on clinical judgment of the treating physician Dataset-5 (University Hospital Zurich): naturalistic, clinical dataset with no experimental groups Dataset-6 (SYNAEDA): naturalistic, clinical dataset, patients were assigned to different psychotherapy interventions based on clinical judgment of the treating physician Dataset-7 (Prague Psychiatric Center): patients were not randomized, all patients first received placebo then ketamine
Blinding	In the validation analyses (Dataset-1, -3 and -5), author HV who determined the prediction, was blinded to clinical outcome. Clinical outcome data was provided by the respective study site (authors SO, SS/JW, AP, DAG, MB). Accuracy of prediction was verified by author MA. Blinding to treatment allocation was not possible due to the potential opposite directionality of biomarker predictions to different treatments and guided hypotheses.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☐	☒ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	TD-BRAIN+: heterogeneous clinical database with diverse psychiatric diagnoses, comprising children and adults (ages 6-88) of both genders who received different treatments or no treatment (irrelevant to the present study because only baseline data was used) Dataset-1 (EMBARC): 240 female and male adult outpatients (mean age=35.4, age range =18-65, males=35.8%) with primary diagnosis of early onset (before age 30) non-psychotic recurrent or chronic major depressive disorder per SCID-I, with symptom severity of QIDS≥14 and with no previous failed antidepressant trials in the current episode Dataset-2: 187 female and male outpatients (mean age=43.6, age range 18-80, males=51%) who met clinical criteria for primary diagnosis of nonpsychotic MDD or dysthymia at treatment start Dataset -3: female and male outpatients (mean age=42.4, age range =18-80, males=46%) who met clinical criteria for primary diagnosis of nonpsychotic MDD or dysthymia at treatment start Dataset-4 (SNOEP): female and male adult patients (mean age=51, age range =18-80, males=37%) who met clinical criteria for diagnosis of MDD at treatment start Dataset-5 (University Hospital Zurich): female and male adult patients (mean age=51.2, age range =18-85, males=36%) who met ICD-10 criteria for diagnosis of MDD at treatment start Dataset-6 (SYNAEDA): female and male adult patients (mean age=37.2, age range =18-66, males=37%) who met DSM-IV criteria for primary diagnosis of nonpsychotic MDD or dysthymia and BDI≥14. 94 patients underwent cognitive behavioral therapy (CBT) and 81 underwent another form of psychotherapy. Dataset-7 (Prague Psychiatric Center): female and male adult participants (mean age=43.5, age range =18-65, males=46% who met DSM-IV criteria for primary diagnosis of non-psychotic MDD of medium severity (MADRS>20) at study entry.
Recruitment	TD-BRAIN+: no active recruitment, data collection was part of treatment-as-usual in clinical practices; all available data for which written informed consent was present, was used. Dataset-1 (EMBARC): Patients were consecutively recruited from 4 university hospitals: 634 patients were screened, 296 were randomized to receive sertraline hydrochloride or placebo. No self-selection biases could be identified. This data was acquired for secondary analysis; participant recruitment is detailed in the primary publications of this study. Dataset-2, -3, -4, -5, and -6 : no active recruitment, data collection was part of treatment-as-usual in clinical practices; all available data for which written informed consent was present, was used. Dataset-7 (Prague Psychiatric Center): Patients were recruited into three consecutive controlled trials and enrolled by the principal study investigator. No self-selection biases could be identified. This data was acquired for secondary analysis; participant recruitment is detailed in the primary publications of this study.
Ethics oversight	TD-BRAIN+: clinical data. This study was not reviewed by an independent ethics committee. All participants provided written informed consent. Dataset-1 (EMBARC): The original study was approved by the institutional review boards of all sites: University of Texas Southwestern Medical Center (Coordinating Center), Columbia University/Stony Brook (Data Center), Massachusetts General Hospital (MGH), 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; participants received financial compensation. Dataset-2, -3, -5, and -6 are clinical datasets collected in practices in clinics. These studies were not reviewed by an independent ethics committee. All participants provided written informed consent. Dataset-4: Rijnstate Hospital and the medisch-ethische toetsingscommissies (METC) Arnhem/Nijmegen approved the study protocol; all participants provided written consent for the original study from which the data has been used. Dataset-7 (Prague Psychiatric Center): 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; all participants provided written consent for the original study from which the data has been used.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about clinical studies

All manuscripts should comply with the ICMJE guidelines for publication of clinical research and a completed CONSORT checklist must be included with all submissions.

Clinical trial registration	TD-BRAIN+: N/A Dataset-1 (EMBARC): Establishing Moderators and Biosignatures of Antidepressant Response for Clinical Care for Depression (EMBARC). Identifier: NCT01407094. URL: http://clinicaltrials.gov/show/NCT01407094. Dataset-2, -3, -4, -5, and -6: N/A Dataset-7: 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
Study protocol	The full study protocols can be accessed in the respective primary publications of Dataset-1 (Pizzagalli, D. A., et al. Pretreatment Rostral Anterior Cingulate Cortex Theta Activity in Relation to Symptom Improvement in Depression: A Randomized Clinical Trial. Jama Psychiat 75, 547 (2018). and Trivedi, M. H., et al. Establishing moderators and biosignatures of antidepressant response in clinical care (EMBARC): Rationale and design. J Psychiatr Res 78, 11–23 (2016).) and Dataset-7 (Meyer, T., et al. Predictive value of heart rate in treatment of major depression with ketamine in two controlled trials. Clin Neurophysiol 132, 1339–1346 (2021).). Dataset-2, -3, -5, and -6 are clinical datasets collected in practices at different treatment sites and did not have a trial protocol.

Dataset-4 is an unpublished datasets of clinical data; the full trial protocol can be acquired via the principal investigator (author JAW)

Data collection

TD-BRAIN+: Multiple sites in the Netherlands: Research Institute Brainclinics (Brainclinics Foundation, Nijmegen), EPI-PIT clinics (Eindhoven & Tilburg), EEG resource (Nijmegen), Neuroscan (Dordrecht), neuroCare clinics (Hengelo; Groningen; Munich; Sydney)
Dataset-1 (EMBARC): 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
Dataset-2: All patients who received treatment at one of three outpatient mental health care clinics (neuroCare Clinic Nijmegen, neuroCare Clinic The Hague, and Psychologenpraktijk Timmers Oosterhout) between May 2007 and November 2016 and provided written informed consent were screened for inclusion.
Dataset-3: All patients who received treatment at one of two outpatient mental health care clinics (neuroCare Clinic Nijmegen, neuroCare Clinic The Hague) between December 2016 and June 2022 were screened for inclusion.
Dataset-4 (SNOEP): All patients who were referred for ECT treatment at Rijnstate hospital between August 2016 and June 2022 and provided written informed consent were screened for inclusion.
Dataset-5 (University Hospital Zurich): All patients who underwent ECT treatment at University Hospital Zurich between 2006 and 2015 and provided written informed consent were screened for inclusion.
Dataset-6 (SYNAEDA): All patients who were referred for treatment at one of three outpatient mental health care clinics (Synaeda Leeuwaarden Fonteinland, Synaeda Drachten, Synaeda Heerenveen) and provided written informed consent were screened for inclusion.
Dataset-7 (Prague Psychiatric Center): Between 2010 and 2022 outpatients were recruited for study participation at Prague Psychiatric Centre, Czech Republic.

Outcomes

The predefined clinical primary outcome was remission in accordance with the previous publication on iAF as a biomarker. Remission was defined in line with Riedel et al. (Response and remission criteria in major depression – A validation of current practice. J Psychiatr Res 2010;44:1063–8. https://doi.org/10.1016/j.jpsychires.2010.03.006.) as a score of ≤12 on the BDI-II (for Dataset-2, -3, and -6), ≤7 on the HRSD-17 (for Dataset-1 and -4), ≤2 on the CGI (Dataset-5), and ≤7 on the MADRS (for Dataset-7) dependent on the depression scale used in the different studies. No secondary clinical outcomes were taken into account.
Primary outcome of assessing biomarker performance was evaluated as the (normalized) positive predictive value in each biomarker subgroup and number-needed-to-treat.