



# An electroencephalographic signature predicts antidepressant response in major depression

Wei Wu<sup>1,2,3,4</sup>, Yu Zhang<sup>2,3,4</sup>, Jing Jiang<sup>2,3,4</sup>, Molly V. Lucas<sup>2,3,4</sup>, Gregory A. Fonzo<sup>2,3,4</sup>, Camarin E. Rolle<sup>2,3,4</sup>, Crystal Cooper<sup>5,6</sup>, Cherise Chin-Fatt<sup>5,6</sup>, Noralie Krepel<sup>7,8</sup>, Carena A. Cornelissen<sup>2,3,4</sup>, Rachael Wright<sup>2,3,4</sup>, Russell T. Toll<sup>2,3,4</sup>, Hersh M. Trivedi<sup>2,3,4</sup>, Karen Monuszko<sup>2,3,4</sup>, Trevor L. Caudle<sup>2,3,4</sup>, Kamron Sarhadi<sup>2,3,4</sup>, Manish K. Jha<sup>5</sup>, Joseph M. Trombello<sup>5,6</sup>, Thilo Deckersbach<sup>9</sup>, Phil Adams<sup>10</sup>, Patrick J. McGrath<sup>10</sup>, Myrna M. Weissman<sup>10</sup>, Maurizio Fava<sup>9</sup>, Diego A. Pizzagalli<sup>9</sup>, Martijn Arns<sup>7,11,12</sup>, Madhukar H. Trivedi<sup>5,6,13</sup> and Amit Etkin<sup>2,3,4,13</sup> ✉

**Antidepressants are widely prescribed, but their efficacy relative to placebo is modest, in part because the clinical diagnosis of major depression encompasses biologically heterogeneous conditions. Here, we sought to identify a neurobiological signature of response to antidepressant treatment as compared to placebo. We designed a latent-space machine-learning algorithm tailored for resting-state electroencephalography (EEG) and applied it to data from the largest imaging-coupled, placebo-controlled antidepressant study ( $n = 309$ ). Symptom improvement was robustly predicted in a manner both specific for the antidepressant sertraline (versus placebo) and generalizable across different study sites and EEG equipment. This sertraline-predictive EEG signature generalized to two depression samples, wherein it reflected general antidepressant medication responsivity and related differentially to a repetitive transcranial magnetic stimulation treatment outcome. Furthermore, we found that the sertraline resting-state EEG signature indexed prefrontal neural responsivity, as measured by concurrent transcranial magnetic stimulation and EEG. Our findings advance the neurobiological understanding of antidepressant treatment through an EEG-tailored computational model and provide a clinical avenue for personalized treatment of depression.**

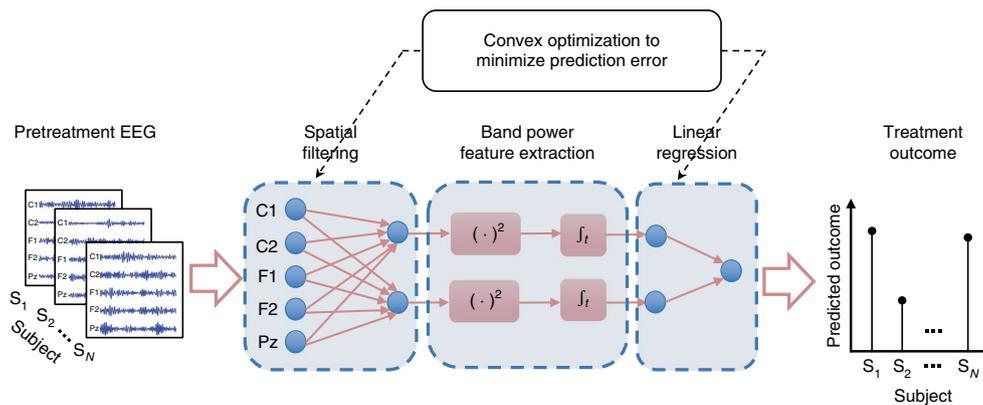
Major depression is currently defined on the basis of clinical criteria and encompasses a heterogeneous mix of neurobiological phenotypes<sup>1</sup>. This heterogeneity may account for the modest superiority of antidepressant medication over placebo (Cohen's  $d$  of  $\sim 0.3$ )<sup>2–6</sup>. Work over the past two decades has suggested that resting-state EEG (rsEEG) may be able to identify treatment-predictive heterogeneity in depression<sup>7–10</sup>. Specific attention has been paid to prefrontal and parietal signals carried by the  $\theta$  (4–7 Hz) and  $\alpha$  (8–12 Hz) frequency bands<sup>7–13</sup>. However, due to lack of cross-validation and small sample sizes, previous studies have either identified nonspecific predictors that do not differentiate between response to drug versus placebo, such as rostral anterior cingulate  $\theta$  current density<sup>11–13</sup> or failed to yield robust (that is, generalizable) and reproducible neural signatures that are predictive at the individual patient level<sup>8</sup>. As such, we still lack a robust neurobiological signature for an antidepressant-responsive phenotype that could identify which patients will derive a large benefit from medication. Delineating such a signature would advance both a neurobiological understanding of treatment response and yield important clinical implications.

To identify a robust antidepressant-responsive depression phenotype, machine learning can be used to combine across the

complex multivariate relationships existing within rsEEG data. An effective predictive rsEEG computational model, however, must deal with three critical challenges. The first is a smearing of signal and noise resulting from volume conduction due to each electrode picking up neural signals from multiple sources, and adjacent electrodes detecting neural signals from the same sources<sup>14</sup>. Second, there is a risk of overfitting the model given the high spatiotemporal dimensionality and noisiness of EEG data<sup>15,16</sup>. Third, there are challenges in simultaneously optimizing feature identification and fitting of predictive regression models due to the nonlinearity of the error function with respect to the model parameters<sup>17</sup>.

To address each of these challenges, we developed a machine-learning algorithm that we called Sparse EEG Latent Space Regression (SELSER). The data were derived from four studies. First, we established the rsEEG predictive signature by training SELSER on data derived from the largest neuroimaging-coupled placebo-controlled randomized clinical study of antidepressant efficacy ( $n = 309$ ). We then used three additional datasets to validate our findings. Together, these efforts aimed to reveal a treatment-responsive phenotype in depression, dissociate between medication and placebo response, establish its mechanistic significance and

<sup>1</sup>School of Automation Science and Engineering, South China University of Technology, Guangzhou, Guangdong, China. <sup>2</sup>Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA. <sup>3</sup>Wu Tsai Neuroscience Institute, Stanford University, Stanford, CA, USA. <sup>4</sup>Veterans Affairs Palo Alto Healthcare System and the Sierra Pacific Mental Illness, Research, Education and Clinical Center (MIRECC), Palo Alto, CA, USA. <sup>5</sup>Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, TX, USA. <sup>6</sup>O'Donnell Brain Institute, University of Texas Southwestern Medical Center, Dallas, TX, USA. <sup>7</sup>Research Institute Brainclinics, Brainclinics Foundation, Nijmegen, the Netherlands. <sup>8</sup>Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, the Netherlands. <sup>9</sup>Department of Psychiatry, Harvard Medical School and McLean Hospital, Belmont, MA, USA. <sup>10</sup>New York State Psychiatric Institute & Department of Psychiatry, College of Physicians and Surgeons of Columbia University, New York, NY, USA. <sup>11</sup>Department of Experimental Psychology, Utrecht University, Utrecht, the Netherlands. <sup>12</sup>neuroCare Group Netherlands, Nijmegen, the Netherlands. <sup>13</sup>These authors contributed equally: Madhukar H. Trivedi, Amit Etkin. ✉e-mail: [amitetkin@stanford.edu](mailto:amitetkin@stanford.edu)



**Fig. 1 | End-to-end prediction of the treatment outcome with a latent-space model.** The model consists of three stages: (1) spatial filtering that linearly transforms the EEG signals to the latent signals; (2) band power feature extraction that computes the band power of each latent signal; and (3) linear regression that uses the band powers to predict the treatment outcome. By solving a convex optimization problem, all the unknown parameters (spatial filters and linear regression weight coefficients) are optimized in conjunction under a unified objective function that trades off between the prediction error and dimensionality of the latent signals.  $S_1$ ,  $S_2$  and  $S_N$  refer to Subject 1, Subject 2 and the  $N$ th Subject, respectively. C1, C2, F1, F2 and Pz refer to example electrode locations according to the 10/10 international system.  $(\cdot)^2$  denotes the square operator and  $\int_t$  denotes the average of a time series over time.

provide initial evidence for the potential for treatment selection on the basis of an rsEEG signature.

## Results

**Development of SELSER.** We developed SELSER (Fig. 1 and Methods) to address the challenges mentioned above and identify a robust signature from EEG data that could predict response to antidepressants. Signal identification and mitigation of volume conduction are accomplished by amplifying signal-to-noise ratio through use of spatial filters<sup>18</sup>. Each spatial filter transforms the multichannel EEG data into a single latent signal, the power of which is used as a feature for the machine-learning algorithm. Because model fitting is performed under a sparse constraint on the number of spatial filters, this also serves to reduce the dimension of the underlying latent signals and thus, decrease the chance of overfitting. Finally, we mathematically formulated the outcome prediction framework as solving a convex problem that related EEG time series to the treatment outcome directly, yielding a single and globally optimal solution<sup>19</sup>. This approach can be contrasted, for example, with regression models being applied to channel-level rsEEG power measures, which does little to mitigate volume conduction. Likewise, conventional latent-space methods, such as independent-component analysis (ICA)<sup>20</sup> and principal-component analysis (PCA) are not optimal, in that they are unsupervised approaches not directly related to optimizing the model fit to the treatment outcome prediction target. In light of previous rsEEG work<sup>7–9</sup>, we predicted that SELSER-established neural signals drawn from  $\vartheta$  and  $\alpha$  frequency bands would most strongly predict treatment outcome.

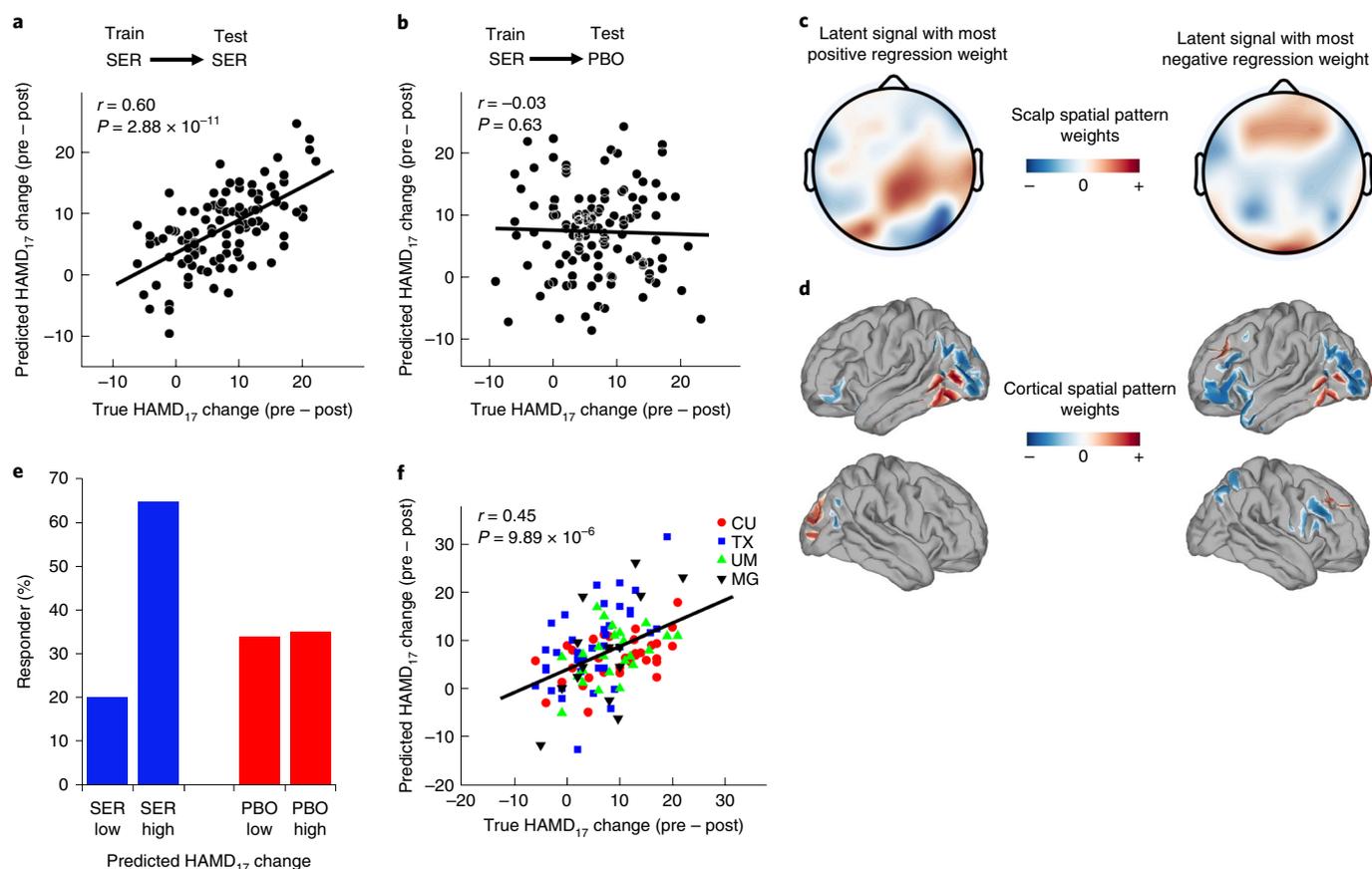
Data were drawn from four studies. Establishment of the treatment-predictive rsEEG signature was accomplished with data from the Establishing Moderators and Biosignatures of Antidepressant Response in Clinic Care (EMBARC) study<sup>21</sup>. EMBARC is the largest neuroimaging-coupled placebo-controlled randomized clinical trial (RCT) in depression to date and involved randomization of 309 medication-free depressed outpatients ( $n = 228$  with high quality rsEEG data) to receive either the selective serotonin reuptake inhibitor sertraline or placebo for 8 weeks (Supplementary Fig. 1). Eyes-open and eyes-closed EEG data were collected before randomization at four sites in the United States, each of which used a different high-density EEG system and/or electrode montage (60 or more electrodes). Clinical outcome was assessed on the 17-item clinician-administered Hamilton Depression Rating Scale (HAMD<sub>17</sub>).

The generalizability of the antidepressant-predictive signature was then tested in a second independent sample of depressed patients ( $n = 72$ ), for whom we had historical information about treatment response during the current depressive episode, as well as rsEEG data. A third independent sample of depressed patients ( $n = 24$ ) was used to assess two features of the treatment-predictive rsEEG signature: convergent validity and neurobiological significance. Specifically, we tested whether expression of our rsEEG signature correlated with another machine-learning signature we developed based on task-based functional magnetic resonance imaging (fMRI) activation in EMBARC<sup>22</sup>, as this would provide further convergent validation of the rsEEG signature identified here. We also tested in the third sample whether regions that were prominent in the rsEEG signature reflected individual differences in cortical responsivity, as directly assessed through single-pulse transcranial magnetic stimulation (TMS) during concurrent EEG recording.

Finally, in a fourth depressed sample ( $n = 152$ ) that was treated with either 10-Hz left dorsolateral prefrontal repetitive TMS (rTMS) or 1-Hz right dorsolateral prefrontal rTMS (both with concurrent psychotherapy), we tested whether the strength of the EMBARC-trained rsEEG signature predicted outcome with an antidepressant treatment that has a putatively different mechanism of action. This allowed us to test the generalizability of our results and open up the potential for treatment selection by defining the neural predictors of antidepressant response.

**Treatment prediction from pretreatment resting EEG using SELSER.** We built prediction models using pretreatment rsEEG by applying SELSER to each of four canonical EEG frequency bands ( $\vartheta$ , 4–7 Hz;  $\alpha$ , 8–12 Hz;  $\beta$ , 13–30 Hz;  $\gamma$ , 31–50 Hz) in each resting condition (two 2-min blocks each of eyes open or eyes closed). Treatment outcome was quantified as the pre- minus post-treatment difference in HAMD<sub>17</sub> scores, with missing endpoint values imputed to maintain an intent-to-treat framework. Model performance was tested using tenfold cross-validation (Fig. 1 and Supplementary Fig. 2; Methods).

For the sertraline arm, only  $\alpha$  signals from the resting eyes open (REO) condition were significantly predictive of the observed treatment score changes during cross-validation (Fig. 2a; Pearson's  $r = 0.60$ , r.m.s. error (r.m.s.e.) = 5.68, Bonferroni-corrected  $P = 2.88 \times 10^{-11}$ ; permutation test-verified using 1,000 permutations,  $P < 10^{-3}$ ). When the sertraline-trained model was applied to the placebo arm, however, outcome could not be predicted (Fig. 2b;



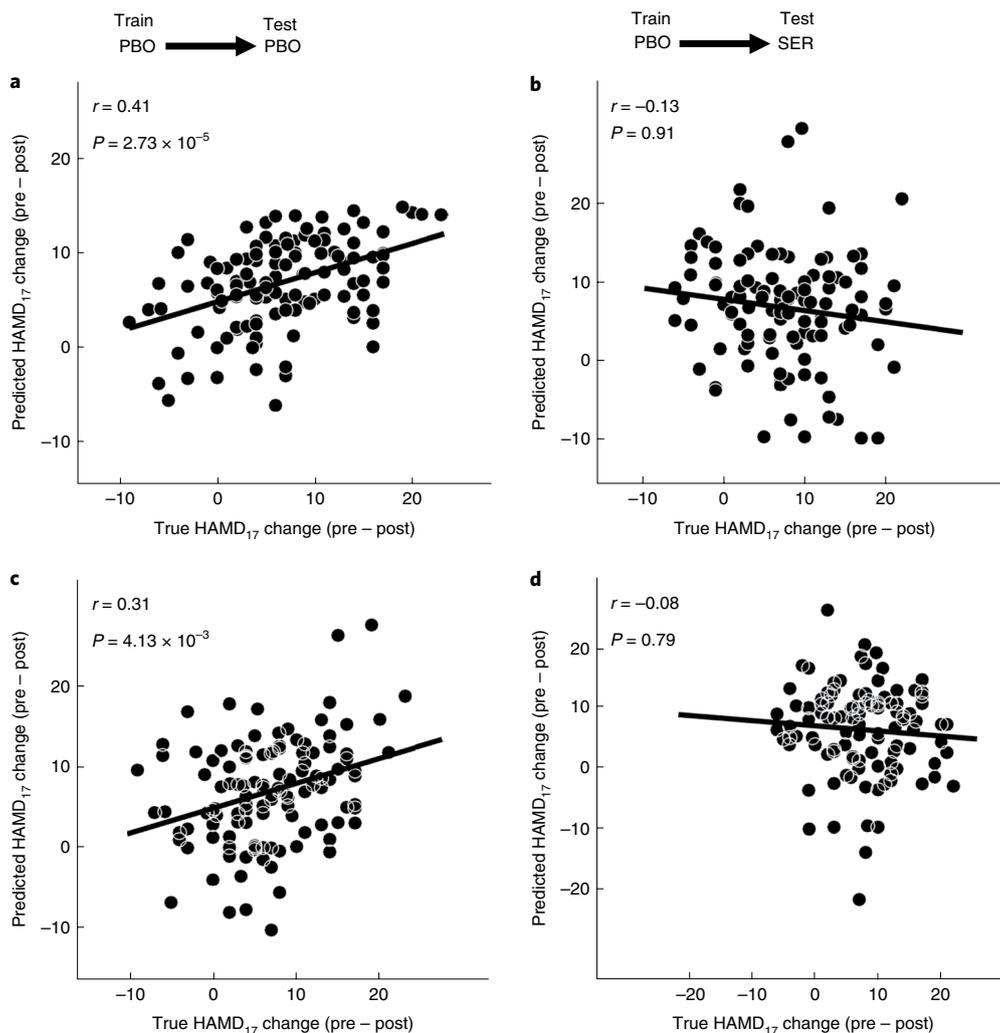
**Fig. 2 | Prediction of outcome specific to sertraline using SELSER on REO  $\alpha$ -frequency range data.** **a**, A  $10 \times 10$  stratified cross-validation prediction of HAMD<sub>17</sub> change in the sertraline arm ( $n=109$ ) using SELSER. Pearson's  $r=0.60$ , Bonferroni-corrected  $P=2.88 \times 10^{-11}$  based on the one-sided test against the alternative hypothesis that  $r>0$ . **b**, Application of the sertraline-trained model to the placebo arm ( $n=119$ ) failed to predict outcome, demonstrating the specificity of the model for sertraline prediction. Pearson's  $r=-0.03$ ,  $P=0.63$  based on the one-sided test against the alternative hypothesis that  $r>0$ . **c**, Scalp spatial patterns of the SELSER latent signals, with the most positive ( $\beta=759.31$ ; left) and negative ( $\beta=-853.13$ ; right) regression weights, respectively ( $n=109$ ). **d**, Cortical spatial patterns of the SELSER latent signals, with the most positive ( $\beta=759.31$ ; left) and negative ( $\beta=-853.13$ ; right) regression weights, respectively ( $n=109$ ). **e**, Purely for the purpose of visualizing the utility of the rsEEG predictive signature, patients in each arm were partitioned into the low and high groups by applying a median split on the cross-validated predicted HAMD<sub>17</sub> score changes for sertraline response. The response rate was then calculated for each group (defined as a 50% or greater decrease in symptoms from baseline). SER, sertraline (blue); PBO, placebo (red). **f**, Treatment prediction across study sites in a leave-study-site-out cross-validation on the  $\alpha$  REO sertraline model ( $n=109$ ). CU, Columbia University; TX, University of Texas Southwestern Medical Center; UM, University of Michigan; MG, Massachusetts General Hospital. Site effect was corrected for by removing the mean of the covariance matrix from each study site before the SELSER analysis. Pearson's  $r=0.45$ , Bonferroni-corrected  $P=9.89 \times 10^{-6}$  based on the one-sided test against the alternative hypothesis that  $r>0$ .

Pearson's  $r=-0.03$ , r.m.s.e. = 9.77,  $P=0.63$ ), thus demonstrating the specificity of this model for sertraline efficacy prediction (Fisher's  $z$ -test:  $z=4.94$ ,  $P=8 \times 10^{-7}$ ). Application of SELSER to  $\alpha$ -frequency REO rsEEG signals could not, however, predict baseline HAMD<sub>17</sub> scores (Pearson's  $r=0.06$ , r.m.s.e. = 6.10,  $P=0.27$ ), thus illustrating that the treatment-predictive model was not related to the severity of baseline depression.

As a result of the algorithm-enforced low-dimensionality constraint on the latent signals in SELSER, only a few latent signals were obtained in each model (Supplementary Fig. 3). For the sertraline  $\alpha$  REO model, the scalp and cortical spatial maps of the two latent signals with the most positive and negative regression weights are shown in Fig. 2c,d, respectively. The spatial pattern of the latent signal with the most positive regression weight was mainly centered around the right parietal-occipital regions, in line with previous work<sup>23</sup>. In contrast, the spatial pattern of the latent signal with the most negative regression weight was heavily concentrated in both the lateral prefrontal and parieto-occipital regions.

We also tested the effects of the amount of data on model fitting and found that performance began degrading when fewer than two blocks of 1.5 min each per patient were used (Supplementary Figs. 4 and 5).

For the placebo arm, both  $\alpha$  signals from the REO and resting eyes closed (REC) conditions significantly predicted the HAMD<sub>17</sub> score change (REO: Fig. 3a, Pearson's  $r=0.41$ , r.m.s.e. = 6.34, Bonferroni-corrected  $P=2.73 \times 10^{-5}$ ; and REC: Fig. 3c, Pearson's  $r=0.31$ , r.m.s.e. = 7.60, Bonferroni-corrected  $P=4.13 \times 10^{-3}$ ; permutation test-validated  $P<10^{-3}$ ). The spatial maps of the two latent signals with the most positive and negative regression weights are shown in Supplementary Fig. 6 for the REO and REC conditions. For the REO condition, the spatial patterns of the latent signals were predominantly in the temporal and occipital regions, whereas for the REC condition, frontal-parietal and frontal regions were the most prominent. When applied to the sertraline arm, both regression models failed to predict outcome (Fig. 3b,d; Fisher's  $z$ -test,  $z>2.98$ ,  $P<3 \times 10^{-3}$ ), demonstrating the specificity of these models for placebo outcome prediction and distinction from the sertraline-predictive model above.



**Fig. 3 | Prediction of outcome specific to placebo using SELSER on  $\alpha$ -frequency range data. a,c,** A  $10 \times 10$  stratified cross-validation prediction of HAMD<sub>17</sub> change in the placebo arm ( $n=119$ ) using SELSER on REO (Pearson's  $r=0.41$ , Bonferroni-corrected  $P=2.73 \times 10^{-5}$  based on the one-sided test against the alternative hypothesis that  $r>0$ ) (a) and REC (Pearson's  $r=0.31$ , Bonferroni-corrected  $P=4.13 \times 10^{-3}$ , based on the one-sided test against the alternative hypothesis that  $r>0$ ) (c)  $\alpha$ -frequency range data. **b,d,** Application of the REO (Pearson's  $r=-0.13$ ,  $P=0.91$ , based on the one-sided test against the alternative hypothesis that  $r>0$ ) (b) and REC (Pearson's  $r=-0.08$ ,  $P=0.79$ , based on the one-sided test against the alternative hypothesis that  $r>0$ ) (d) placebo-trained models to the sertraline arm ( $n=109$ ) failed to predict outcome, demonstrating specificity of the model for placebo prediction.

To visualize how the SELSER predictions in Fig. 2a could be used for treatment stratification, we partitioned the patients in each arm by applying an arbitrary median split on the cross-validated predicted HAMD<sub>17</sub> score changes derived from the sertraline  $\alpha$  REO model. We then calculated the rates of treatment response ( $\geq 50\%$  reduction in symptoms) for the portion of patients above the median ('high') and below the median ('low'), respectively, based on the model-predicted HAMD<sub>17</sub> change scores (Fig. 2e and Supplementary Fig. 7). For the sertraline arm, the high group reached a response rate of 65%, which more than tripled the response rate (20%) in the low group and was considerably higher than the response rates in the placebo arm (35% and 34% for the high and low groups, respectively).

**Treatment prediction across study sites.** To further assess the generalizability of the prediction models to unseen data collected with different EEG amplifiers, a leave-study-site-out analysis was performed by iteratively using three study sites' data to train the model, and the fourth site's data for testing the model. Since the four study sites' EEG data were acquired with different EEG

amplifiers and/or electrode montages (Supplementary Table 2), marked variability of prediction performance was observed across study sites (Supplementary Fig. 8). To mitigate this site effect, the mean of the covariance matrix was removed from each study site before the SELSER analysis. For the sertraline arm, only the  $\alpha$  REO model was significantly predictive of the treatment outcome when performing leave-study-site-out cross-validation (Fig. 2f; Pearson's  $r=0.45$ , r.m.s.e.=7.02, Bonferroni-corrected  $P=9.89 \times 10^{-6}$ ; permutation test-validated using 1,000 permutations,  $P<10^{-3}$ ). This demonstrates the robustness of our model for unseen data from a different EEG amplifier (arguably a worst-case scenario with respect to testing through cross-validation). An equivalent fourfold cross-validation model sampling across all sites remained strongly predictive (Pearson's  $r=0.58$ , r.m.s.e.=5.63,  $P=1.59 \times 10^{-11}$ ; permutation test-validated using 1,000 permutations,  $P<10^{-3}$ ). Further restricting the sample size via twofold cross-validation across all sites yielded a lower yet still highly significant predictive performance (Pearson's  $r=0.38$ , r.m.s.e.=6.32,  $P=2.25 \times 10^{-5}$ ). For the placebo arm, none of the REO and REC models was predictive of the treatment outcome when cross-validating between study sites (Pearson's

$r < 0.22$ , r.m.s.e.  $> 7.90$ , Bonferroni-corrected  $P > 0.07$ ). The sertraline leave-study-site-out  $\alpha$  REO predictive model also still showed significantly greater specificity in predicting sertraline over placebo response (Fisher's  $z$ -test,  $z = 3.83$ ,  $P = 10^{-4}$ ).

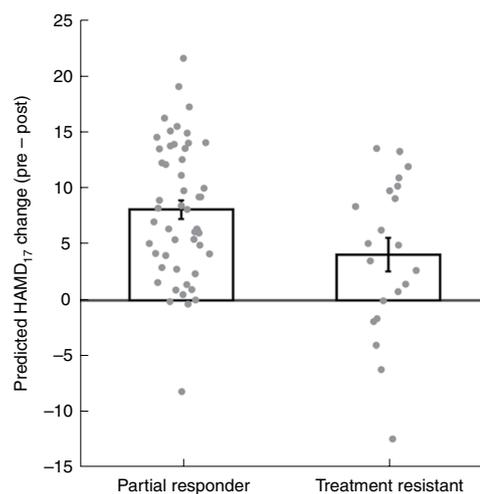
#### Comparison of SELSER predictions to previous methods.

To benchmark SELSER against conventional machine-learning approaches that do not use latent-space modeling, we also trained linear regression models on eyes open rsEEG data by using a relevance vector machine (RVM)<sup>24</sup> on channel-level  $\alpha$  band power<sup>25,26</sup>,  $\theta$  band power<sup>27,28</sup> and  $\theta$  cordance<sup>13,23</sup> (Supplementary Figs. 9 and 10). However, none of the models was predictive of the treatment outcome on cross-validation (Pearson's  $r < -0.06$ , r.m.s.e.  $> 8.20$ ,  $P > 0.74$ ). Moreover, to demonstrate the improvement of SELSER over conventional latent-space modeling approaches, we also trained the RVM on  $\alpha$  band power of the latent signals extracted with PCA<sup>29</sup> (Supplementary Fig. 9d) or ICA<sup>25</sup> (Supplementary Fig. 9e), which are among the most popular unsupervised methods to derive spatial filters from EEG (Supplementary Fig. 10). Here also, both models failed to predict the treatment outcome (Pearson's  $r < 0.15$ , r.m.s.e.  $> 7.17$ ,  $P > 0.09$ ).

**Treatment prediction from symptoms.** Assessing brain activation for defining an individual's sertraline-responsive phenotype may not be relevant in practice if lower-cost measures, such as clinical severity scores, demographic variables or historical factors like childhood trauma exposure could usefully predict outcome. This did not prove to be the case, however, as RVM trained with all of these was only modestly predictive (Supplementary Fig. 11; Pearson's  $r = 0.26$ , r.m.s.e. = 7.93,  $P = 3 \times 10^{-3}$  for sertraline, and Pearson's  $r = 0.16$ , r.m.s.e. = 9.56,  $P = 0.05$  for placebo) and worse if using only the Quick Inventory of Depressive Symptomatology (QIDS) scale (Pearson's  $r = 0.12$ , r.m.s.e. = 6.85,  $P = 0.12$  for sertraline and Pearson's  $r = 0.06$ , r.m.s.e. = 6.72,  $P = 0.26$  for placebo). Sertraline outcome prediction with clinical measures was also significantly weaker than the  $\alpha$  REO rsEEG model above (Fisher's  $z$ -test,  $z = 3.11$ ,  $P = 0.0019$ ).

#### Testing the generalization of the rsEEG sertraline-predictive signature.

We next tested the generalizability of the SELSER rsEEG sertraline-predictive signature from EMBARC in a second independent cohort of patients with depression. This cohort of patients were drawn from a naturalistic, longitudinal depression study in which rsEEG data were recorded at the baseline visit<sup>30</sup>. Patients also completed the Antidepressant Treatment Response Questionnaire (ATRQ), which provided historical information about the number of adequate antidepressant medication trials in the current episode, as well as whether patients responded to them or not. Following conventional groupings, patients were categorized as either treatment resistant (two or more failed antidepressant trials;  $n = 21$ ) or as partial responders (partial response to at least one medication;  $n = 51$ ). The mean-removal site correction procedure was performed for the rsEEG data, as for the leave-study-site-out analysis of the EMBARC study. We then applied the EMBARC-trained sertraline  $\alpha$ -band rsEEG model to each patient, yielding a predicted HAMD<sub>17</sub> change for each individual, which reflects their strength of expression of the sertraline-predictive rsEEG signature. As expected, the predicted HAMD<sub>17</sub> change was higher for the partial-responder group than the treatment-resistant group (Fig. 4), demonstrating the generalizability of the EMBARC rsEEG sertraline signature to the broader construct of treatment responsiveness and resistance to antidepressant medication. Moreover, information on the number of within-episode failed antidepressant trials was available for 45 of the 72 patients. Also as expected, we found a negative correlation between the number of failed trials and the magnitude of the rsEEG sertraline signature-predicted HAMD<sub>17</sub> improvement (Pearson's  $r = -0.34$ ,  $P = 0.023$ ).



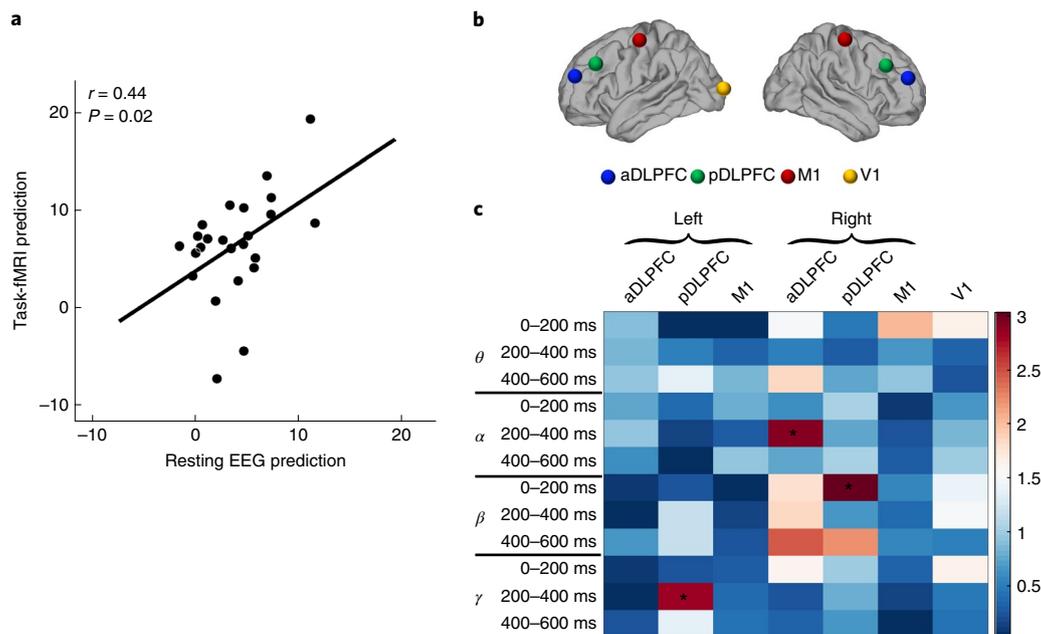
**Fig. 4 | Prediction of treatment outcome by the EMBARC-trained sertraline rsEEG model, applying to baseline eyes open rsEEG of the second depression study cohort.** The plot shows the predicted HAMD<sub>17</sub> change for patients who are partial responders ( $n = 51$ ) or resistant to treatment ( $n = 21$ ). These data demonstrate that the predicted HAMD<sub>17</sub> change is significantly larger in patients who are partial responders than in those who are treatment resistant (two-sample and two-sided Student's  $t$ -test  $P = 0.016$ ). Error bars depict s.e.m.

#### Convergence between rsEEG- and task-fMRI-derived machine-learning predictions.

To test the convergent validity of the sertraline rsEEG model from EMBARC, we examined a separate dataset of 24 patients with depression who were assessed in a cross-sectional manner (without treatment) using both rsEEG and task-based fMRI with the emotional conflict task<sup>22</sup>. The reason for doing so is that we could test whether the predicted HAMD<sub>17</sub> change based on the EMBARC-trained sertraline rsEEG model correlated with the predicted HAMD<sub>17</sub> change based on an fMRI emotional conflict task-based machine-learning model that we established in a separate analysis of EMBARC data<sup>22</sup> (Methods). Since the EEG data were recorded with yet another amplifier distinct from those used in EMBARC, the mean-removal site correction procedure was performed. The rsEEG and task-fMRI predictions were significantly correlated with each other in these independent data (Fig. 5a; Pearson's  $r = 0.44$ ,  $P = 0.02$ ). This finding provides convergent support, across fMRI and EEG, for the existence of a treatment-responsive neurobiological phenotype in major depressive disorder across populations and across assessment modalities.

#### TMS and EEG correlates of rsEEG phenotype.

Next, to provide further insight into the neural signals driving our sertraline-predictive rsEEG-defined phenotype, we analyzed concurrent single-pulse TMS and EEG (spTMS and EEG)<sup>31</sup> data from the third independent depression sample as used above. Specifically, we sought to test whether cortical responsivity, as assessed by direct stimulation using spTMS and EEG to regions either prominent or minimal within the spatial patterns of the rsEEG latent signals, induced neural responses that correlated with the rsEEG-defined treatment-predictive phenotype. The stimulated regions were the bilateral posterior dorsolateral prefrontal cortices (pDLPFC), anterior DLPFC (aDLPFC), along with primary visual cortex (V1) and bilateral primary motor cortices (M1) as the control regions (Fig. 5b). We localized pDLPFC and aDLPFC using neuronavigation on the basis of their being nodes within the frontoparietal and salience resting-state networks, respectively, as we have conducted in our previous work<sup>32</sup>. To quantify the correlation between the spTMS and EEG responses and the sertraline-predictive rsEEG phenotype, we again



**Fig. 5 | Alignment of predicted  $HAMD_{17}$  change calculated by the rsEEG model and predicted  $HAMD_{17}$  change calculated by a machine-learning model trained on task-based fMRI activation from a separate analysis on EMBARC data, as well as neural responsivity assessed through concurrent spTMS and EEG. **a**, The EMBARC-trained rsEEG and task-fMRI models were applied to an independent major depressive disorder dataset that had both data types and the ensuing predicted  $HAMD_{17}$  changes from both models were correlated with each other. Pearson's  $r = 0.44$ ,  $P = 0.02$  based on the one-sided test against the alternative hypothesis that  $r > 0$  ( $n = 24$ ). **b**, Concurrent spTMS and EEG correlates of the rsEEG phenotype in the independent depressed dataset, analyzed using SELSER. TMS was delivered to pDLPFC (part of the frontoparietal control network), aDLPFC (part of the ventral attention network), M1 and V1. These sites were identified on the basis of ICAs on resting-state fMRI data from a separate cohort. aDLPFC, anterior dorsolateral prefrontal cortex; pDLPFC, posterior dorsolateral prefrontal cortex. **c**, A significance plot of the correlation between the concurrent spTMS and EEG responses and sertraline-predictive rsEEG SELSER signature, as indexed by the leave-one-out cross-validated Pearson's correlation coefficients between the predicted rsEEG phenotype (using concurrent spTMS and EEG responses) and the true rsEEG phenotype, for each of the stimulation sites ( $n = 24$ ). The SELSER analysis was performed separately for the same set of frequency bands as used in the rsEEG prediction analysis and for three time windows relative to the TMS pulse (0–200 ms, 200–400 ms, 400–600 ms), followed by an FDR correction across all tests. Only right aDLPFC stimulation ( $\alpha$  band, 200–400 ms: Pearson's  $r = 0.60$ ,  $P = 5.5 \times 10^{-4}$ , based on the one-sided test against the alternative hypothesis that  $r > 0$ ), left pDLPFC stimulation ( $\gamma$  band, 200–400 ms: Pearson's  $r = 0.58$ , FDR-corrected  $P = 8 \times 10^{-4}$ ) and right pDLPFC stimulation ( $\beta$  band, 0–200 ms: Pearson's  $r = 0.60$ ,  $P = 4.6 \times 10^{-4}$ , based on the one-sided test against the alternative hypothesis that  $r > 0$ ) survived FDR correction (denoted by asterisks). The plot shows  $-\log_{10}(P)$  of the correlation of the SELSER-predicted rsEEG phenotype with the true rsEEG phenotype.**

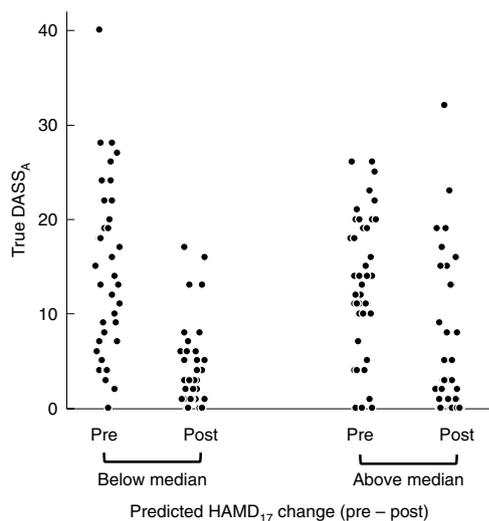
employed SELSER to relate spTMS and EEG data to the sertraline rsEEG signature (Fig. 5c). Correlations between the sertraline rsEEG phenotype and spTMS and EEG responses to stimulation observed at three of four prefrontal cortical regions survived correction for multiple comparisons. This included right aDLPFC stimulation ( $\alpha$  band, 200–400 ms, Pearson's  $r = 0.60$ , false discovery rate (FDR)-corrected  $P = 5.5 \times 10^{-4}$ ), left pDLPFC stimulation ( $\gamma$  band, 200–400 ms, Pearson's  $r = 0.58$ , FDR-corrected  $P = 8 \times 10^{-4}$ ) and right pDLPFC stimulation ( $\beta$  band, 0–200 ms, Pearson's  $r = 0.60$ , FDR-corrected  $P = 4.6 \times 10^{-4}$ ). Correlation of responses to stimulation of primary motor or visual cortices did not survive correction.

**Sertraline signature assessment in a combined repetitive TMS and psychotherapy treatment study.** In light of the sertraline rsEEG phenotype indexing cortical responsivity to stimulation at several DLPFC locations, we next considered whether the strength of this phenotype could predict outcome with rTMS treatment in depression. Analyses were performed on a fourth previously reported dataset of patients with depression and pretreatment rsEEG recordings, who received at least ten sessions of simultaneous rTMS and psychotherapy and were on a stable medication regimen<sup>33,34</sup>. Treatment involved rTMS applied with either a 10-Hz protocol over the left DLPFC ( $n = 64$ ) or a 1-Hz protocol over the right DLPFC ( $n = 88$ ; Methods). We computed each patient's expression of the

EMBARC-trained sertraline rsEEG model (expressed as predicted  $HAMD_{17}$  change) using the same mean site removal procedure as above. Symptoms were assessed with the Beck Depression Inventory (BDI) and the three subscales of the Depression, Anxiety and Stress Scale (DASS), separately by an rTMS protocol, using linear mixed models and a Bonferroni correction for eight comparisons (two frequencies and four outcome measures). One relationship survived, wherein less rsEEG-predicted  $HAMD_{17}$  change with sertraline was associated with greater response to 1-Hz rTMS on the DASS (rsEEG-predicted  $HAMD_{17}$  sertraline change  $\times$  time interaction:  $F(1,128) = 9.02$ ,  $P = 4 \times 10^{-3}$ ; Fig. 6). This suggests that patients who fail to respond to sertraline may be more amenable to 1-Hz right DLPFC rTMS, providing a potential evidence-based treatment selection approach for depression. This relationship was also specific to 1-Hz right DLPFC rTMS, as we found a treatment protocol  $\times$  predicted  $HAMD_{17}$  change  $\times$  time interaction when including both arms in a linear regression ( $F(1,126) = 4.54$ ,  $P = 0.035$ ).

## Discussion

Here we developed an rsEEG-optimized latent space computational model, called SELSER, with which we obtained robust prediction of antidepressant outcome and moderation (differential prediction) between outcome with an antidepressant versus placebo in a large placebo-controlled study. The antidepressant-predictive



**Fig. 6 | Prediction of treatment outcome with right DLPFC 1-Hz rTMS treatment by the EMBARC-trained sertraline rsEEG model, applying to pre-rTMS eyes open rsEEG.** The scatterplot shows the pre and post-treatment scores for patients on the Anxiety subscale of the DASS (DASS<sub>A</sub>). To visualize the linear mixed model relating rsEEG-predicted HAMD<sub>17</sub> change to observed changes in DASS<sub>A</sub> scores, shown here is a median split of the predicted HAMD<sub>17</sub> change values. These data demonstrate that the degree of pre-to-post change in DASS<sub>A</sub> symptoms due to 1-Hz rTMS treatment is greater in those patients with lower expected HAMD<sub>17</sub> change scores using the EMBARC-trained sertraline rsEEG model.

signature identified using SELSER on  $\alpha$  frequency range eyes-open rsEEG data was superior to conventional machine-learning models or latent modeling methods, such as ICA or PCA. This signature was furthermore superior to a model trained on clinical data alone, was able to predict outcome on rsEEG data acquired at a study site not included in the model training set and which used a different EEG amplifier and/or electrode montage and related to general antidepressant responsiveness and resistance in a separate depression sample. The attributes of our SELSER model support its potential utility in the context of real-world clinical care and with regard to stratification of future depression studies based on the expected antidepressant-specific treatment outcome for that individual. Critically, though promising and highly influential signals for antidepressant prediction with rsEEG have been reported for the past two decades<sup>7–9,35</sup>, previous signals neither moderated between outcome with an antidepressant versus placebo, nor provided robust individual patient-level prediction. Our present work is thus distinguished from previous findings on both fronts.

We also found evidence of multimodal convergent validity for our rsEEG antidepressant-response signature in a third depression dataset by virtue of its correlation with expression of a task-based fMRI signature that we recently identified using EMBARC data<sup>22</sup>. The strength of our rsEEG signature also correlated with prefrontal neural responsivity, as indexed by direct stimulation with spTMS and EEG. This led us to test the relationship between the sertraline rsEEG model and treatment outcome with rTMS in a fourth sample. There, we found an opposite relationship between sertraline-predicted improvement and observed treatment outcome with 1-Hz right DLPFC rTMS. This finding also opens an exciting avenue for neural signature-driven treatment selection in depression.

From a neural mechanism perspective, we note that the sertraline SELSER model revealed both positively and negatively weighted posterior cortical eyes open rsEEG signals, but only heavily negatively

weighted prefrontal signals. Considering suggestions that resting  $\alpha$  power reflects inhibitory tone in a brain region<sup>36,37</sup>, the negative weighting of prefrontal  $\alpha$  in our sertraline model suggests that the prefrontal cortices of better treatment responders are more active or excitable than those of poor responders. The positive and negative weights for posterior signals suggests that optimizing the balance of distinct posterior predictive signals may be what is critical to establishing a robust computational model. Prediction analysis using electrodes exclusively from posterior regions corroborates this claim (Supplementary Fig. 12). Our results are broadly in line with previous reports of better outcomes being associated with greater posterior cortical  $\alpha$  power<sup>23</sup>. However, a large-scale study that lacked a placebo control arm failed to replicate these findings<sup>38</sup>. Thus, as our RVM model trained on channel-level  $\alpha$  power failed to predict outcome, the critical element in attaining individual-level robust outcome prediction may be the use of a latent-space computational model.

Perhaps somewhat surprising is that only eyes open  $\alpha$  rsEEG, but not eyes closed, was predictive of the treatment outcome, given that the  $\alpha$  rhythm is more strongly present in the eyes closed condition. One explanation is that the increasing  $\alpha$  rhythm during the eyes closed condition is indicative of cortical areas being deactivated<sup>39</sup> and thus may contribute to background noise rather than predictive signal. This view is consistent with motor intention decoding work, wherein the goal has been to enhance the difference between the  $\mu$  motor rhythm, while suppressing the more broadly distributed  $\alpha$  rhythm<sup>40,41</sup>.

The present work does not directly inform the cognitive and emotional information processing relevance of the antidepressant-response rsEEG signature. However, the relationship between the rsEEG signature and one derived from task-based fMRI data from EMBARC<sup>22</sup> suggests that individuals with stronger expression of the rsEEG signature may be better able to regulate emotional conflict.

We utilized a range of cross-validation methods (including on data from entirely unseen study sites and collected on different EEG equipment), compared prediction of outcome with sertraline versus placebo across cross-validation methods and tested for generalization of the signals across a number of complimentary datasets. As such, our data provide neurobiological evidence that an antidepressant-responsive phenotype exists within the biological heterogeneity characteristic of the broader clinical diagnosis of depression. These findings thereby not only advance a ‘personalized’ approach to depression<sup>1,5,42</sup>, but also demonstrate that antidepressants only seem to be modestly more effective than placebo because they are typically given to an unselected sample of patients with depression.

If replication and extension of this phenotype to other antidepressant medications is successful in progressively larger and more diverse datasets, the rsEEG signature we identified may be helpful in deciding whether a patient should continue further medication trials after an initial failure with an antidepressant medication or switch to treatments with putatively different mechanisms of action (such as rTMS, electroconvulsive therapy (ECT) or psychotherapy). Indeed, patients with depression often undergo many medication trials before advancing to other treatments, such as rTMS<sup>43,44</sup>, which are effective for some medication-resistant patients<sup>45,46</sup>. This could result in potentially avoidable morbidity and economic cost if they are switched to another intervention earlier on the basis of evidence of little expected benefit with an antidepressant using our rsEEG signature.

Our finding of an opposite relationship between predicted change with sertraline, and treatment outcome with 1-Hz rTMS and concurrent psychotherapy, directly supports the potential for these findings to guide treatment selection in depression, pending further replication. Furthermore, as the sertraline-predictive rsEEG signature did not predict outcome with 10-Hz rTMS during concurrent psychotherapy, this suggests that it is the effect of the specific rTMS protocol that is being predicted rather than the

effects of psychotherapy, though further investigation with rTMS treatment without psychotherapy is needed for this to be conclusive. Of note, this opposite direction prediction between antidepressants and rTMS is consistent with our previous work as well as that of others. More intact default-mode network connectivity in the iSPOT-D study predicted better treatment outcome with antidepressant treatment<sup>47</sup>, whereas more disrupted default-mode network connectivity predicted outcome with rTMS in a clinic-based cohort<sup>48</sup>.

There are also several limitations to consider. First, our specific prediction findings remain to be replicated in an entirely independent sample and it is not known whether the sertraline signature predicts outcome with a broader class of selective serotonin reuptake inhibitors or antidepressant medication more generally. Second, EMBARC assessed all patients while they were medication-free at baseline. Thus, while this removes potential confounders from the experimental design, it also limits generalizability to typical outpatients, who are often already on an antidepressant. Finally, SELSER is a static latent-space modeling approach that does not consider resting EEG dynamics. Dynamic latent-space models have been recently developed to predict mood from multisite intracortical human brain signals<sup>49</sup>.

In summary, we developed a rsEEG-optimized latent-space computation model that was capable of robustly predicting treatment outcome with the antidepressant sertraline and distinguishing between response to sertraline versus placebo at the individual patient level and which may furthermore support treatment selection between medication and rTMS. Together, these findings ground in individual-level neurobiology a treatment-responsive phenotype obscured within the broader clinical diagnosis of depression and its associated biological heterogeneity and lay a path toward machine-learning-driven personalized approaches to treatment in depression.

### Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41587-019-0397-3>.

Received: 8 December 2018; Accepted: 17 December 2019;  
Published online: 10 February 2020

### References

- Drysdale, A. T. et al. Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat. Med.* **23**, 28–38 (2017).
- Cipriani, A. et al. Comparative efficacy and acceptability of 21 antidepressant drugs for the acute treatment of adults with major depressive disorder: a systematic review and network meta-analysis. *Lancet* **391**, 1357–1366 (2018).
- Fournier, J. C. et al. Antidepressant drug effects and depression severity: a patient-level meta-analysis. *JAMA* **303**, 47–53 (2010).
- Khan, A. & Brown, W. A. Antidepressants versus placebo in major depression: an overview. *World Psychiatry* **14**, 294–300 (2015).
- Kirsch, I. *The Emperor's New Drugs: Exploding the Antidepressant Myth* (Random House, 2009).
- Kirsch, I. et al. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the food and drug administration. *PLoS Med.* **5**, e45 (2008).
- Wade, E. C. & Iosifescu, D. V. Using electroencephalography for treatment guidance in major depressive disorder. *Biol. Psychiatry Cogn. Neurosci. Neuroimaging* **1**, 411–422 (2016).
- Widge, A. S. et al. Electroencephalographic biomarkers for treatment response prediction in major depressive illness: a meta-analysis. *Am. J. Psychiatry* **2018**, 17121358 (2018).
- Olbrich, S. & Arns, M. EEG biomarkers in major depressive disorder: discriminative power and prediction of treatment response. *Int. Rev. Psychiatry* **25**, 604–618 (2013).
- Jaworska, N., de la Salle, S., Ibrahim, M.-H., Blier, P. & Knott, V. Leveraging machine-learning approaches for predicting antidepressant treatment response using electroencephalography (EEG) and clinical data. *Front. Psychiatry* **9**, 768 (2019).
- 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**, 547–554 (2018).
- Korb, A. S., Hunter, A. M., Cook, I. A. & Leuchter, A. F. Rostral anterior cingulate cortex theta current density and response to antidepressants and placebo in major depression. *Clin. Neurophysiol.* **120**, 1313–1319 (2009).
- Leuchter, A. F., Cook, I. A., Witte, E. A., Morgan, M. & Abrams, M. Changes in brain function of depressed subjects during treatment with placebo. *Am. J. Psychiatry* **159**, 122–129 (2002).
- Nunez, P.L. & Srinivasan, R. *Electric Fields of the Brain: The Neurophysics of EEG* (Oxford Univ. Press, 2006).
- Müller, K.-R. et al. Machine learning for real-time single-trial EEG-analysis: from brain-computer interfacing to mental state monitoring. *J. Neurosci. Methods* **167**, 82–90 (2008).
- Wu, W., Nagarajan, S. & Chen, Z. Bayesian machine learning: EEG/MEG signal processing measurements. *IEEE Signal Process. Mag.* **33**, 14–36 (2016).
- Schirrmester, R. T. et al. Deep learning with convolutional neural networks for EEG decoding and visualization. *Hum. Brain Mapp.* **38**, 5391–5420 (2017).
- Haufe, S. et al. On the interpretation of weight vectors of linear models in multivariate neuroimaging. *Neuroimage* **87**, 96–110 (2014).
- Boyd, S. & Vandenberghe, L. *Convex Optimization* (Cambridge Univ. Press, 2004).
- Bell, A. J. & Sejnowski, T. J. An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* **7**, 1129–1159 (1995).
- 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).
- Fonzo, G. A. et al. Brain regulation of emotional conflict differentiates response to antidepressants versus placebo in depression. *Nat. Hum. Behav.* **3**, 1319–1331 (2019).
- Bruder, G. E. et al. Electroencephalographic  $\alpha$  measures predict therapeutic response to a selective serotonin reuptake inhibitor antidepressant: pre- and post-treatment findings. *Biol. Psychiatry* **63**, 1171–1177 (2008).
- Tipping, M. E. Sparse Bayesian learning and the relevance vector machine. *J. Mach. Learn. Res.* **1**, 211–244 (2001).
- Grin-Yatsenko, V. A., Baas, I., Ponomarev, V. A. & Kropotov, J. D. Independent component approach to the analysis of EEG recordings at early stages of depressive disorders. *Clin. Neurophysiol.* **121**, 281–289 (2010).
- Pozzi, D., Golimstock, A., Petracchi, M., García, H. & Starkstein, S. Quantified electroencephalographic changes in depressed patients with and without dementia. *Biol. Psychiatry* **38**, 677–683 (1995).
- Iosifescu, D. V. et al. Frontal EEG predictors of treatment outcome in major depressive disorders. *Eur. Neuropsychopharmacol.* **19**, 772–777 (2009).
- Arns, M., Drinkenburg, W. H., Fitzgerald, P. B. & Kenemans, J. L. Neurophysiological predictors of non-response to rTMS in depression. *Brain Stimul.* **5**, 569–576 (2012).
- Tipping, M. E. & Bishop, C. M. Probabilistic principal component analysis. *J. R. Stat. Soc. B* **61**, 611–622 (1999).
- Trivedi, M. H. et al. Comprehensive phenotyping of depression disease trajectory and risk: rationale and design of Texas Resilience Against Depression study (T-RAD). *J. Psychiatric Res.* **122**, 22–32 (2020).
- Hill, A. T., Rogasch, N. C., Fitzgerald, P. B. & Hoy, K. E. TMS-EEG: a window into the neurophysiological effects of transcranial electrical stimulation in non-motor brain regions. *Neurosci. Biobehav. Rev.* **64**, 175–184 (2016).
- Chen, A. C. et al. Causal interactions between fronto-parietal central executive and default-mode networks in humans. *Proc. Natl Acad. Sci. USA* **110**, 19944–19949 (2013).
- Donse, L., Padberg, F., Sack, A. T., Rush, A. J. & Arns, M. Simultaneous rTMS and psychotherapy in major depressive disorder: clinical outcomes and predictors from a large naturalistic study. *Brain Stimul.* **11**, 337–345 (2018).
- Krepel, N. et al. Non-replication of neurophysiological predictors of non-response to rTMS in depression and neurophysiological data-sharing proposal. *Brain Stimul.* **11**, 639–641 (2018).
- Leuchter, A. F. et al. Comparative effectiveness of biomarkers and clinical indicators for predicting outcomes of SSRI treatment in major depressive disorder: results of the BRITE-MD study. *Psychiatry Res.* **169**, 124–131 (2009).
- Klimesch, W., Sauseng, P. & Hanslmayr, S. EEG  $\alpha$  oscillations: the inhibition-timing hypothesis. *Brain Res. Rev.* **53**, 63–88 (2007).
- Jensen, O. & Mazaheri, A. Shaping functional architecture by oscillatory  $\alpha$  activity: gating by inhibition. *Front. Hum. Neurosci.* **4**, 186 (2010).
- 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).
- Lehtonen, J. & Lehtinen, I.  $\alpha$  rhythm and uniform visual field in man. *Electroencephalogr. Clin. Neurophysiol.* **32**, 139–147 (1972).

40. Hari, R. & Salmelin, R. Human cortical oscillations: a neuromagnetic view through the skull. *Trends Neurosci.* **20**, 44–49 (1997).
41. Ramoser, H., Müller-Gerking, J. & Pfurtscheller, G. Optimal spatial filtering of single trial EEG during imagined hand movement. *IEEE Trans. Rehabil. Eng.* **8**, 441–446 (2000).
42. Kraemer, H. C. Messages for clinicians: moderators and mediators of treatment outcome in randomized clinical trials. *Am. J. Psychiatry* **173**, 672–679 (2016).
43. Nguyen, K. H. & Gordon, L. G. Cost-effectiveness of repetitive transcranial magnetic stimulation versus antidepressant therapy for treatment-resistant depression. *Value Health* **18**, 597–604 (2015).
44. Voigt, J., Carpenter, L. & Leuchter, A. Cost effectiveness analysis comparing repetitive transcranial magnetic stimulation to antidepressant medications after a first treatment failure for major depressive disorder in newly diagnosed patients: a lifetime analysis. *PLoS ONE* **12**, e0186950 (2017).
45. O'Reardon, J. P. et al. Efficacy and safety of transcranial magnetic stimulation in the acute treatment of major depression: a multisite randomized controlled trial. *Biol. Psychiatry* **62**, 1208–1216 (2007).
46. George, M. S. et al. Daily left prefrontal transcranial magnetic stimulation therapy for major depressive disorder: a sham-controlled randomized trial. *Arch. Gen. Psychiatry* **67**, 507–516 (2010).
47. Williams, L. M., DeBattista, C., Duchemin, A., Schatzberg, A. & Nemeroff, C. Childhood trauma predicts antidepressant response in adults with major depression: data from the randomized international study to predict optimized treatment for depression. *Transl. Psychiatry* **6**, e799 (2016).
48. Liston, C. et al. Default mode network mechanisms of transcranial magnetic stimulation in depression. *Biol. Psychiatry* **76**, 517–526 (2014).
49. Sani, O. G. et al. Mood variations decoded from multi-site intracranial human brain activity. *Nat. Biotechnol.* **36**, 954–961 (2018).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature America, Inc. 2020

## Methods

**EMBARC study.** *Trial registration.* Establishing Moderators and Biosignatures of Antidepressant Response for Clinical Care for Depression (EMBARC) is registered with ClinicalTrials.gov (identifier: NCT 01407094).

*Participants and treatment.* Written informed consent was obtained from each participant under institutional review board (IRB)-approved protocols at each of the four study sites: University of Texas Southwestern Medical Center, Massachusetts General Hospital, Columbia University and University of Michigan. Data reported here are based on EMBARC participants who were randomly assigned to sertraline or placebo during stage 1 of the trial ( $N=309$  total). Key eligibility for the study included the following: being 18–65 years old; having major depression as a primary diagnosis by the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID); at least moderate depression severity with a score  $\geq 14$  on the QIDS Self Report scale at screening and randomization; a major depressive episode beginning before age 30; either a chronic recurrent episode (duration  $\geq 2$  years) or recurrent major depressive disorder (at least two lifetime episodes); no antidepressant failure during the current episode. Exclusion criteria included the following: current pregnancy, breastfeeding or no use of contraception; lifetime history of psychosis or bipolar disorder; substance dependence in the past 6 months or substance abuse in the past 2 months; unstable psychiatric or general medical conditions requiring hospitalization; study medication contraindication; clinically significant laboratory abnormalities; history of epilepsy or condition requiring an anticonvulsant; ECT, vagal nerve stimulation, TMS or other somatic treatments in the current episode; medications (including but not limited to antipsychotics and mood stabilizers); current psychotherapy; high suicide risk; or failure to respond to any antidepressant at an adequate dose and duration in the current episode.

*Clinical trial.* EMBARC used a double-blind design, wherein participants were randomized to an 8-week course of sertraline or placebo. Randomization was stratified by site, depression severity and chronicity using a block randomization procedure. Sertraline dosing began at 50 mg using 50 mg capsules and was increased as tolerated if the patient did not respond until a maximum dose of 200 mg. A similar dosing approach was used for placebo capsules.

*Clinical outcome measure.* Our primary outcome was the HAMD<sub>17</sub>. For participants lacking an endpoint HAMD<sub>17</sub>, multiple imputation by chained equations was conducted in R using the package mice<sup>50</sup>. The following observed variables were utilized to impute endpoint HAMD<sub>17</sub> values for missing data via Bayesian regression: baseline HAMD<sub>17</sub>, week 1 HAMD<sub>17</sub>, week 2 HAMD<sub>17</sub>, week 3 HAMD<sub>17</sub>, week 4 HAMD<sub>17</sub>, week 6 HAMD<sub>17</sub>, baseline QIDS total score, baseline Mood and Symptom Questionnaire subscale scores for Anxious Arousal, Anhedonic Depression and General Distress, Snaith–Hamilton Pleasure Scale total score, age, years of education, sex and Wechsler Abbreviated Scale of Intelligence  $t$ -scores for Vocabulary and Matrix Reasoning.

*Resting-state EEG acquisition.* An rsEEG was recorded from each of the four study sites. The EEG amplifier settings are summarized in Supplementary Table 1. At all study sites, amplifier calibrations were performed. Experimenters were certified by the Columbia EEG team after demonstrating accurate EEG cap placement and delivery of task instructions via video conference and then submitting satisfactory EEG data from a pilot subject.

An rsEEG was recorded during four 2-min blocks (two blocks for eyes-closed and two blocks for eyes open) in a counterbalanced order. Participants were instructed to remain still and minimize blinks or eye movements and to fixate on a centrally presented cross during the eyes-open condition.

*Resting-state EEG preprocessing.* The recorded rsEEG data were cleaned offline with our in-house fully automated artifact rejection pipeline, thereby minimizing the biases in preprocessing possible with manual rejection of artifacts. The steps are briefly described as follows: (1) EEG data were resampled to 250 Hz; (2) the 60-Hz AC line noise artifact was removed using CleanLine ([www.nitrc.org/projects/cleanline](http://www.nitrc.org/projects/cleanline)); (3) nonphysiological slow drifts in the EEG recordings were removed using a 0.01-Hz high-pass filter; (4) the spectrally filtered EEG data were then re-referenced to the common average; (5) bad epochs were rejected by thresholding the magnitude of each epoch. Bad channels were rejected based on thresholding the spatial correlations among channels. Subjects with more than 20% bad channels were discarded. The rejected bad channels were then interpolated from the EEG of adjacent channels via the spherical spline interpolation; (6) remaining artifacts were removed using ICA. Independent components (ICs) related to the scalp muscle artifact, ocular artifact and ECG artifact, were automatically rejected using a pattern classifier trained on expert-labeled ICs from another independent EEG dataset; and (7) EEG data were re-referenced to the common average. After artifact rejection, 54 EEG channels common to all four study sites were identified and extracted for each subject. Subjects whose total powers across all the channels were beyond three s.d. of the mean total power were discarded. Consequently, of the 266 patients with pretreatment EEG recordings, 228 had usable EEG data for analyses. The baseline sociodemographic and clinical information of these 228 patients is provided in Supplementary Table 1. The 38 patients with unusable EEG

recordings mainly had too many bad EEG channels and exceedingly large total power across channels.

**Second depression study cohort (validating rsEEG antidepressant-predictive signature).** *Participants and treatment.* The second depression study was carried out at University of Texas Southwestern Medical Center, which is one of the four study sites in EMBARC<sup>50</sup>. Written informed consent was obtained under an IRB-approved protocol at University of Texas Southwestern Medical Center. Individuals were eligible for the study if they were aged 10 years or older and could speak, read and understand English. To be included, participants needed to have a lifetime or current diagnosis of a mood disorder (major depressive disorder, persistent depressive disorder, bipolar I/II/NOS, bipolar/mood disorder with psychotic features or depressive disorder otherwise specified (subthreshold)) based upon a semi-structured diagnostic interview.

Exclusion criteria included the following: history of schizophrenia, schizoaffective disorders or chronic psychotic disorders based upon a semi-structured diagnostic interview; inability to provide a stable home address and contact information; having had any condition for which, in the opinion of the investigator or designee, study participation would not be in their best interest (including but not limited to marked cognitive impairment, unstable general medical condition, intoxication or active psychosis) or that could prevent, limit, or confound the protocol-specified assessments; or requirement for immediate hospitalization for psychiatric disorder or suicidal risk as assessed by a licensed study clinician.

*Screening and baseline assessments.* Whenever possible, screening and baseline assessments took place on the same day and began after informed consent was obtained. A potential participant's eligibility was determined following review of the inclusion and exclusion criteria and assessment with the Mini-International Neuropsychiatric Interview. In addition to a number of self-reported symptom measures, which are outside the scope of the current analysis, patients completed the ATRQ at the initial visit, which is a self-rated scale for determining treatment resistance in major depressive disorders, within the current depressive episode.

Once eligibility requirements were met, baseline procedures were completed over two visits depending on the needs of the individual participant. In the event of a split visit, neuroimaging and EEG procedures could be performed on separate days. This study captured a range of information including sociodemographics, general clinical data, physical exam, blood and stool samples, behavioral testing, neuroimaging and EEG, though here we focus only on the EEG data, given the scope of the present study.

*Resting-state EEG acquisition.* EEG signals were acquired with two EEG amplifiers, each for a different portion of the participants. The first amplifier was the same 62-channel NeuroScan SynAmps amplifier (NeuroScan) used in EMBARC with identical acquisition parameters. The second amplifier was the Net Amps 300 amplifier with the high-density 256-channel HydroCel Geodesic Sensor Net (Electrical Geodesic). Cz was used as the reference electrode and the sampling rate was set at 1,000 Hz. Electrode impedances were kept below 50 k $\Omega$ . A total of 35 and 37 participants' EEG data were collected using the first and second amplifiers, respectively. Participants were seated on a comfortable reclining chair and were instructed to remain awake and let their mind naturally wander in the eyes-closed paradigm and then fixate a given point in the eyes-open paradigm, with each paradigm in two 2-min blocks.

*Resting-state EEG preprocessing and analysis.* The recorded rsEEG data were cleaned offline with the identical fully automated artifact rejection pipeline as used in EMBARC. After artifact rejection, 54 EEG channels common to all four study sites in EMBARC were identified and extracted for each subject. Subjects whose total powers across all the channels were beyond three s.d. of the mean total power were discarded.

**Third depression study cohort (correlation between rsEEG antidepressant-predictive signature and fMRI signature and with TMS and EEG response to probe stimulation).** *Participants.* Written informed consent was obtained from each participant under an IRB-approved protocol at Stanford University. Participants in the study underwent several assessments during a clinical intake interview to determine eligibility and classification for the study. The diagnosis of depression (and comorbid conditions) was assessed by a clinician using the SCID, as for EMBARC. Key eligibility for the study included the following: age 18–50 years old; no current psychotherapy; free of metal or ferrous implant; good English comprehension and nonimpaired intellectual abilities to ensure understanding of task instructions; no history of neurological disorders, brain surgery, electroconvulsive or radiation treatment, brain hemorrhage or tumor, stroke, epilepsy, hypo- or hyperthyroidism; no daily use of as-needed benzodiazepines or opiates (maximum 3 $\times$  per week) or daily thyroid medications and no antidepressant, anticonvulsant or antipsychotic medications for >2 weeks (fluoxetine >6 weeks). Exclusion criteria included the following: left-handed; did not graduate from an English-speaking high school and English was not their first language; psychiatric medications (including but not limited to antipsychotics

and mood stabilizers) and hormonal and/or cancer medications; current psychotherapy; current rTMS treatment or ECT treatment; loss of consciousness greater than 30 min and/or a loss of memory greater than 24 h; lifetime evidence of psychosis, mania, hypomania or bipolar disorders and/or manic episodes on the SCID; diagnosis of substance dependence within the past 3 months (but not abuse). Overall, 24 subjects for whom resting EEG, spTMS and EEG and task-fMRI data were all acquired were considered in the subsequent analyses.

**Resting EEG acquisition.** EEG recordings were acquired with a BrainAmp DC amplifier (sampling rate: 5 kHz; measurement range:  $\pm 16.384$  mV; cutoff frequencies of the analog high-pass and low-pass filters: 0 and 1 kHz) and the Easy EEG cap with 64 extra-flat, freely rotatable, sintered Ag–AgCl electrodes (Brain Products). The electrode montage followed an equidistant arrangement extending from below the cheekbone back to below theinion. Electrode impedances were kept below 5 k $\Omega$ . An electrode attached to the tip of the nose was used as the reference. Participants were seated on a comfortable reclining chair and were instructed to remain awake and let their mind wander in the eyes-closed paradigm and then fixate a given point in the eyes-open paradigm, each for 3 min. Recordings were immediately assessed for quality using a custom MATLAB (R2014b, The Mathworks) script and rerun if necessary.

**Resting-state EEG preprocessing.** The recorded rsEEG data were cleaned offline with the identical fully automated artifact rejection pipeline as used in EMBARC.

**Emotional conflict task.** This well-characterized paradigm assesses both emotional conflict and emotional conflict regulation<sup>51</sup>. Each trial involved presentation of an emotional face with either a fearful or happy expression, drawn from the set of Ekman and Friesen, with an overlaid emotion word ('FEAR' or 'HAPPY'). Participants were instructed to identify the facial emotion with a key press, while trying to ignore the emotion word. The task consisted of 148 trials, with stimuli presented for 1,000 ms in a fast event-related design. Interstimulus intervals were 3,000–5,000 ms in a pseudo-randomized order, counterbalanced for facial expression, sex, word and response button. Stimuli were either congruent (such as, fearful face with 'FEAR') or incongruent (such as, fearful face with 'HAPPY') and stimuli were furthermore balanced to achieve an equal fraction of current and previous trial congruency, while ensuring no direct stimulus repetitions. Before performance of the task during neuroimaging, all participants underwent a practice version to ensure task proficiency was reached (minimum 80% accuracy) and the task instructions were understood. The neuroimaging task lasted 13 min and 14 s.

Regulation in the emotional conflict task occurs via an implicit process when conflict trials are preceded by other conflict trials<sup>51,52</sup>. That is, while emotional conflict results in slowing of reaction times, this effect can be mitigated in incongruent trials that follow incongruent trials (II trials), compared to incongruent trials that follow congruent trials (CI trials). This trial-to-trial adaptive regulation of emotional conflict reflects an active process by which the brain increases emotional control in response to previous trial conflict, which then benefits regulation of emotional conflict on the subsequent trial (captured by the II–CI contrast). This regulation effect, captured through the same contrast, has also been extensively described for nonemotional conflict stimuli<sup>53</sup>. Critically, this contrast between post-incongruent incongruent and post-congruent incongruent trials compares brain responses to physically identical stimuli (incongruent trials) that differ only on the relative emotional conflict regulatory context in which they come due to previous trial congruency and is furthermore independent of the incongruent versus congruent trial (I–C) conflict response contrast. Neuroimaging acquisition parameters are shown in Supplementary Table 3.

**fMRI preprocessing and first-level modeling.** FSL tools were used to preprocess imaging data<sup>54</sup>. Functional images were first realigned to structural images using an affine registration matrix and boundary-based registration on the basis of tissue segmentation as implemented in FSL's FLIRT, which was concatenated with a nonlinear normalization of each participant's T1 image to the Montreal Neurological Institute (MNI) 152-person 1-mm<sup>3</sup> T1 template using FNIRT from FSL v.5.0 to result in a single transformation step from individual native functional space to a structurally aligned and spatially normalized template space. Functional images were realigned to the middle volume of the run. Nuisance signals corresponding to segmented white matter and cerebrospinal fluid were regressed out of motion-corrected functional images. A 6-mm full-width half maximum isotropic smoothing kernel was then applied to preprocessed time series images to account for individual anatomical variability. To ensure the quality of imaging measures, we instituted cutoffs for absolute level of motion (r.m.s. of the absolute level of movement <4 mm across the mean of the squared maximum displacements in each of the six translational and rotational parameters estimated during realignment). In addition, to ensure brain activation measures reflect task-relevant metrics, we also instituted a minimum level of behavioral accuracy during completion of the emotional conflict task as an additional quality control metric (total accuracy  $\geq 70\%$  of trials correct). Functional runs displaying motion higher than our cutoff or accuracy below the minimum cutoff were excluded from further analyses.

For individual-level analyses for each participant, regressors modeling trials of interest were convolved with the hemodynamic response function. First-level general linear models were estimated in SPM 8 (ref. 55). Regressors corresponded to zero-duration markers set at the onset of stimuli, which were explicitly categorized by congruency (incongruent or congruent) and previous trial type (post-incongruent or post-congruent) to model conflict response and regulation effects. This resulted in four different trial types in total, in addition to nuisance regressors for error trials and post-error trials (when applicable) and six motion parameters.

**Single-pulse TMS and EEG acquisition.** Following an anatomical MRI (T1-weighted, 3T) to determine MRI-guided spTMS targets, subjects received spTMS using a Cool-B B6 5 butterfly coil and a MagPro X100 TMS stimulator (MagVenture). Stimulations were delivered to V1, M1, pDLPFC and aDLPFC in a randomized order for each subject. Among these stimulation sites, M1 was defined as the hand knob in the standard MNI space (MNI coordinate: (–38, –18, 64) for left M1 and (40, –18, 64) for right M1). V1 was defined by its MNI anatomical target (MNI coordinate: (0, –100, 2)). For pDLPFC and aDLPFC, the stimulation sites were targeted on the basis of the location of the frontoparietal (executive) control network and ventral attention (salience) network in separate resting-state fMRI data (MNI coordinate: (–32, 42, 34) for left aDLPFC, (30, 50, 26) for right aDLPFC, (–38, 22, 38) for left pDLPFC and (46, 26, 38) for right pDLPFC). Following our previous work, these targets were established using a group ICA on a separate cohort of 38 participants, with the pDLPFC and aDLPFC targets representing peak voxels within the middle frontal clusters of these two networks<sup>22,32</sup>. Coordinates for the pDLPFC and aDLPFC stimulation targets were then transformed to individual subject native space using nonlinear spatial normalization with FSL (<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki>) and used for TMS targeting. The resting motor threshold was determined as the minimum stimulation intensity that produced visible finger movement of the right hand at least 50% of the times when the subject's left M1 was stimulated. TMS coil placement was guided by Visor2 LT 3D neuronavigation system (ANT Neuro) based on co-registration of the functionally defined target to each participant's structural MRI (T1-weighted, slice distance 1 mm, slice thickness 1 mm, sagittal orientation, acquisition matrix 256  $\times$  256) acquired with a 3T GE DISCOVERY MR750 scanner. The TMS coil was placed in a posterior to anterior direction, with an angle of 45° to the nasion–inion axis (studying the optimal coil angles is beyond the scope of this paper). Each target site was stimulated with 60 pulses (biphasic TMS pulses, 280- $\mu$ s pulse width, 120% resting motor threshold, 1,500-ms recharge delay), interleaved at a random interval of 3 s  $\pm$  300 ms. A thin foam pad was attached to the surface of the TMS coil to decrease electrode movement. The subjects were instructed to relax and to fixate at a cross located on the opposing wall, while stimulations were administered by a research assistant.

The same TMS-compatible 64-channel BrainAmp DC amplifier as for rsEEG recordings was used to record spTMS and EEG data. Electrode impedances were kept below 5 k $\Omega$ . An electrode attached to the tip of the nose was used as the reference. DC correction was manually triggered at the end of the stimulations at each site to prevent the saturation of the amplifier due to DC drift.

**Single-pulse TMS and EEG preprocessing.** The recorded spTMS and EEG data were cleaned offline with ARTIST, which is a fully automated artifact rejection algorithm for spTMS and EEG<sup>56</sup> as follows: (1) the initial 10-ms data segment following TMS pulses was discarded to remove the large stimulation-induced electric artifact; (2) the EEG data were downsampled to 1 kHz; (3) big-decay artifacts were automatically removed using ICA on the basis of thresholding; (4) the 60-Hz AC line noise artifact was removed by a notch filter; (5) nonphysiological slow drifts in the EEG recordings were removed using a 0.01-Hz high-pass filter and high-frequency noise was removed by using a 100-Hz low-pass filter; (6) the spectrally filtered EEG data were then re-referenced to the common average and epoched with respect to the TMS pulse (–500 to 1,500 ms); (7) bad trials were rejected by thresholding the magnitude of each trial. Bad channels were rejected on the basis of the spatial correlations among channels. The rejected bad channels were then interpolated from the EEG of adjacent channels; and (8) remaining artifacts were automatically removed using ICA. ICs related to the scalp muscle artifact, ocular artifact and ECG artifact, were rejected using a pattern classifier trained on expert-labeled ICs from other TMS and EEG datasets.

#### Fourth depression study cohort (prediction of outcome with rTMS treatment).

**Participants.** This study was a naturalistic open-label clinical study and has been previously reported in greater detail elsewhere<sup>33</sup>. Briefly, patients were drawn from three outpatient mental healthcare clinics in the Netherlands (neuroCare Clinic Nijmegen, neuroCare Clinic The Hague and Psychologenpraktijk Timmers Oosterhout) between May 2007 and November 2016. Inclusion criteria included (1) a primary diagnosis of nonpsychotic major depressive disorder or dysthymia; (2) BDI, second edition, Dutch version (BDI-II-NL)  $\geq 14$  at baseline; and (3) treatment with at least ten sessions of rTMS over the DLPFC or response within the ten sessions. All participants signed an informed consent under an approved IRB-approved protocol. Additional exclusion criteria included previous ECT treatment, epilepsy, traumatic brain injury, current psychotic disorder, wearing a cardiac pacemaker or metal parts in the head and current pregnancy.

**Treatment.** All patients were treated with either a high-frequency (10 Hz) protocol over the left DLPFC or a low-frequency (1 Hz) protocol over the right DLPFC or both sequentially. The rTMS data included a long time span and the rTMS protocol applied was never based on clinical symptomatology. In the beginning (2006–2012) the standard protocol applied was 10-Hz left DLPFC rTMS and only in some cases 1-Hz right DLPFC rTMS was applied (when there were concerns for safety, such as, paroxysmal activity or seizure risk) due to 1-Hz rTMS being considered a safer protocol. On first inspection of those data<sup>28</sup>, it was found that the clinical benefits for 10 Hz and 1 Hz were indistinguishable, after which time period the standard protocol became 1 Hz right DLPFC<sup>33</sup>. The analyses reported here focus on patients who received only 10-Hz or 1-Hz rTMS, as too few datasets were available on patients who received both treatments or switched treatments mid-way. There were 73 patients in the 10-Hz arm, of which 64 had high quality EEG data, whereas in the 1-Hz arm there were 104 patients, of which 88 had high quality EEG data. Selection of the treatment protocol was not performed in a randomized manner, but rather in the context of clinical care and thus each arm is analyzed separately. rTMS was performed using a Magstim Rapid2 (Magstim Company) or a Deymed DuoMag XT-100 stimulator with a figure-of-eight coil, 70 mm in diameter. For the 10-Hz protocol, rTMS was administered at 10 Hz over the left DLPFC, 110–120% of the resting motor threshold, 30 trains of 5-s duration, inter-train interval of 30 s and 1,500 pulses per session. The 1-Hz protocol consisted of rTMS at 1 Hz over the right DLPFC, 110–120% motor threshold, 120 trains of 10 s duration, inter-train interval of 1 s and 1,200 pulses per session. In the case of both protocols, the low-frequency protocol was administered first with a shorter duration of 1,000 pulses per session and subsequently the high-frequency protocol was administered at full length. The DLPFC was localized using either the 5-cm rule or the Beam F3/F4 method. Furthermore, rTMS treatment was complemented with cognitive behavioral psychotherapy by a trained psychologist. Psychotherapy was performed concurrently with the rTMS treatment in 45-min sessions (rTMS lasting 20 min). Sessions took place with a minimum frequency of two to three times per week and a maximum frequency of two per day, as per the patient's availability.

As these data were drawn from naturalistic clinical care, the total number of sessions depended on clinical decisions and thus varied across patients. Decisions to continue treatment were based on the response to treatment, clinical evaluation of symptom severity and the patient's own request. Decisions followed several rules: if a BDI decrease was observed of at least 20% from baseline ten sessions, the treatment was continued and re-evaluated every five sessions. If no response occurred by session 20–25, treatment was recommended to be terminated unless the patient requested to extend it. If BDI scores reached 12 or below for five sessions, which indicated remission, the patient was given the option of ending or tapering treatment, with an option to extend into maintenance sessions (one session every 6–8 weeks). However, if the threshold of BDI = 12 was reached, but symptom improvement continued, treatment was continued until BDI scores ceased improving.

**Clinical outcome measures.** Clinical outcome was assessed on the BDI (which was the primary outcome measure for the decision rules above) as well as the DASS<sup>57</sup>. The DASS is a self-report questionnaire and consists of three subscales: depression (DASS<sub>D</sub>), anxiety (DASS<sub>A</sub>) and stress (DASS<sub>S</sub>). Each scale consists of 14 items with a four-point severity score, with a maximum total score of 42 on each scale. The patient is asked to fill in the items based on experiences in the previous week.

**Resting EEG acquisition.** EEG data were acquired from 26 channels according to the 10–20 electrode international system (Quickcap; NuAmps). Data were referenced to averaged mastoids with a ground at Fpz. The sampling rate of all channels was 500 Hz. A low-pass filter with attenuation of 40 dB per decade above 100 Hz was employed before digitization. Subjects were asked to rest quietly with their eyes open and eyes closed for 3 min each.

**Preprocessing of resting-state EEG and clinical outcome metrics.** The recorded rsEEG data were cleaned offline with the identical fully automated artifact rejection pipeline as used in EMBARC. Missing data in the clinical metrics were imputed in the same manner as in the EMBARC data, separately by treatment arm.

**Machine-learning analysis. Sparse EEG latent-space regression.** We developed an end-to-end machine-learning algorithm for predicting the treatment outcome from the baseline resting EEG. This algorithm, referred to as SELSER, optimizes a latent-space model that maps the resting EEG data to the treatment outcome by minimizing the prediction error, subject to a constraint on the dimensionality of the latent signals. The band powers in each of the four canonical EEG frequency bands ( $\theta$ , 4–7 Hz;  $\alpha$ , 8–12 Hz;  $\beta$ , 13–30 Hz;  $\gamma$ , 31–50 Hz; filtered using zero-phase FIR filters) are employed as the features. Due to the volume conduction, these band power features are best captured in a latent space rather than in the sensor space. For this purpose, SELSER optimizes a set of spatial filters to linearly transform the multichannel EEG signals in the sensor space to low-dimensional latent signals. A linear regression model is then built to relate the band powers of the latent signals to the treatment outcome.

More formally, SELSER models the treatment outcome  $y_i$  for the  $i$ th subject as follows ( $i = 1, \dots, M$ ):

$$\hat{y}_i = f(\mathbf{X}_i; \{\mathbf{w}_k\}_{k=1}^L, \{\beta_k\}_{k=1}^L, b) = \sum_{k=1}^L \beta_k \mathbf{w}_k^T \mathbf{X}_i \mathbf{X}_i^T \mathbf{w}_k / N + b \quad (1)$$

where  $\mathbf{X}_i \in \mathbb{R}^{C \times N}$  denotes the filtered EEG data for the  $i$ th subject,  $C$  is the number of channels and  $N$  is the number of sampled time points.  $\hat{y}_i$  denotes the predicted treatment outcome for the  $i$ th subject.  $\mathbf{w}_k \in \mathbb{R}^C$  is the  $k$ th spatial filter ( $k = 1, \dots, L$ ),  $\beta_k$  is the  $k$ th weight coefficient of a linear regression model and  $b$  is the intercept of the linear regression model. As can be seen in Supplementary Fig. 10a, prediction is carried out in three phases: (1) the multichannel EEG signals are transformed to  $L$  latent signals  $\{s_k\}_{k=1}^L$  via  $L$  spatial filters  $\{\mathbf{w}_k\}_{k=1}^L$ :  $s_k = \mathbf{X}_i^T \mathbf{w}_k$ ; (2) the band powers of the  $L$  latent signals  $\{z_k\}_{k=1}^L$  are calculated:  $z_k = s_k^T s_k / N$ ; and (3) a linear regression model  $\{\beta_k, b\}_{k=1}^L$  is used to combine the band powers of the latent

signals to predict the treatment outcome:  $\hat{y}_i = \sum_{k=1}^L \beta_k z_k + b$ . It is expected that each latent signal captures a certain portion of information predictive of the treatment outcome, quantified by the band power of the rhythmic EEG activity.

Unlike conventional approaches, where the unknown parameters in the spatial filters and regression model, namely  $\{\mathbf{w}_k, \beta_k, b\}_{k=1}^L$ , were optimized separately under distinct objective functions which may or may not be directly associated with the treatment outcome, we proposed a computationally efficient algorithm (see next section) for optimizing all the model parameters by directly minimizing the mean-squared prediction error  $\sum_{i=1}^N (\hat{y}_i - y_i)^2$  while preventing  $L$  from getting too large to guard against overfitting.

**Parameter optimization in SELSER.** Let  $\mathbf{C}_i = \mathbf{X}_i \mathbf{X}_i^T / N$  denote the EEG spatial covariance matrix. The predicted treatment outcome can be represented alternatively as the following<sup>58</sup>:

$$\hat{y}_i = \sum_{k=1}^L \beta_k \mathbf{w}_k^T [\mathbf{X}_i \mathbf{X}_i^T / N] \mathbf{w}_k + b = \sum_{k=1}^L \beta_k \mathbf{w}_k^T \mathbf{C}_i \mathbf{w}_k + b = \text{Tr}(\mathbf{W}^T \mathbf{C}_i) + b \quad (2)$$

where  $\mathbf{W} = \sum_{k=1}^L \beta_k \mathbf{w}_k \mathbf{w}_k^T \in \mathbb{R}^{C \times C}$  is a symmetric matrix.  $\text{Tr}(\cdot)$  stands for the trace operator, which takes the sum of the diagonal elements of a matrix. Suppose the spatial filters  $\{\mathbf{w}_k\}_{k=1}^L$  are orthogonal to each other, then  $\{\mathbf{w}_k\}_{k=1}^L$  and  $\{\beta_k\}_{k=1}^L$  are the eigenvectors and eigenvalues of  $\mathbf{W}$ , respectively. As a result, optimizing  $\{\mathbf{w}_k\}_{k=1}^L$  and  $\{\beta_k\}_{k=1}^L$  amounts to optimizing  $\mathbf{W}$ , after which  $\{\mathbf{w}_k\}_{k=1}^L$  and  $\{\beta_k\}_{k=1}^L$  can be obtained by performing eigendecomposition of  $\mathbf{W}$ .

However, the number of unknown parameters in  $\mathbf{W}$ ,  $C(C+1)/2$ , is typically much larger than the number of training samples, hence simply minimizing the prediction error is prone to model overfitting. To address this issue, in addition to minimizing the prediction error, we added the rank of  $\mathbf{W}$ , which is equal to  $L$ , as a penalty term into the objective to limit the dimensionality of the latent signals, yielding the following optimization problem:

$$\min_{\mathbf{W}, b} \sum_{i=1}^M (\hat{y}_i - y_i)^2 + \lambda \|\mathbf{W}\|_0, \quad (3)$$

where  $\|\mathbf{W}\|_0$  denotes the rank of  $\mathbf{W}$ . However, equation (3) is NP-hard (requiring non-deterministic polynomial time) and the rank penalty is nonsmooth. Alternatively, the following nuclear norm has been widely used as a convex surrogate of the rank of matrices in a wide range of applications in signal processing and machine learning<sup>58–60</sup>:

$$\|\mathbf{W}\|_* = \sum_{k=1}^L \sigma_k, \quad (4)$$

where  $\{\sigma_k\}_{k=1}^L$  are the singular values of  $\mathbf{W}$ . Consequently, replacing  $\mathbf{W}_0$  with  $\mathbf{W}_*$ , yields

$$\min_{\mathbf{W}, b} \sum_{i=1}^M (\hat{y}_i - y_i)^2 + \lambda \|\mathbf{W}\|_*. \quad (5)$$

Equation (5) is a convex optimization problem, as the objective function is a convex function of  $\mathbf{W}$  and  $b$ . The global minimum solution can be obtained with the accelerated proximal gradient method<sup>61</sup>.

To remove the intersubject variability due to the overall power variation, for each subject  $\mathbf{C}_i$  is normalized by dividing its trace before SELSER analysis:  $\mathbf{C}_i = \mathbf{C}_i / \text{Tr}(\mathbf{C}_i)$ .

**RVM.** To compare with SELSER, RVM<sup>24</sup> with the linear kernel was used to build sparse linear regression models for treatment prediction from non-SELSER-optimized features. By leveraging a sparse prior to penalize overly complex models under the sparse Bayesian learning framework, RVM is able

to automatically select relevant features for prediction via marginal likelihood maximization. Hence, RVM obviates the need of additional validation data to determine the hyperparameters.

**Treatment prediction using SELSER.** We applied SELSER to each canonical EEG frequency band to predict the treatment outcome. To increase the sample size, each of the two blocks of EEG in each resting-state condition (eyes open or eyes closed) was treated as a separate sample for training the model. However, during the leave-subject-out cross-validation, the two blocks from each subject were always grouped together, such that if one block was included in the training set, the other was also included. The predicted outcome of each subject was the average of the predicted outcomes of the two EEG blocks from the subject.

**Visualizing spatial maps of latent signals.** Each latent signal has a spatial map that could be visualized on both the scalp and cortical surface to facilitate neurophysiological interpretation. The scalp spatial maps were calculated as follows<sup>18</sup>:

$$\mathbf{A}_s = \bar{\mathbf{C}} \cdot \mathbf{V} \cdot \bar{\mathbf{C}}_z \quad (6)$$

where  $\mathbf{A}_s \in \mathbb{R}^{C \times L}$  contains the scalp spatial maps of the latent signals as columns,  $\bar{\mathbf{C}} = \sum_{i=1}^N \mathbf{C}_i \in \mathbb{R}^{C \times C}$  is the mean spatial covariance matrix of the EEG signals across subjects,  $\mathbf{V} = [\mathbf{w}_1, \dots, \mathbf{w}_L] \in \mathbb{R}^{C \times L}$  and  $\bar{\mathbf{C}}_z = \mathbf{V}^T \mathbf{C} \mathbf{V} \in \mathbb{R}^{L \times L}$  is the mean covariance matrix of the latent signals across subjects.

To obtain the cortical spatial maps of the latent signals, a three-layer (scalp, skull and cortical surface) boundary element head model was computed with OpenMEEG<sup>62</sup> based on the FreeSurfer average brain template<sup>63</sup>. A total of  $M$  ( $M = 15,002$ ) fixed-orientation dipoles whose orientations were normal to the cortical surface were generated. The lead-field matrix  $\mathbf{B} \in \mathbb{R}^{M \times C}$  relating the dipole activities to the EEG was obtained as a result of the boundary element modeling. A linear inverse matrix  $\mathbf{H} \in \mathbb{R}^{C \times M}$  that maps the EEG signals to the cortical sources was then computed via the sLORETA algorithm<sup>64</sup>. The cortical spatial maps could subsequently be calculated as follows:

$$\mathbf{A}_c = \mathbf{H} \cdot \mathbf{A}_s \quad (7)$$

where  $\mathbf{A}_c \in \mathbb{R}^{M \times L}$  contains the cortical spatial maps of the latent signals as columns.

The spatial maps used for visualization were obtained by training SELSER on the entire EMBARC sample.

**Treatment prediction using the channel-level  $\alpha$ -band power or  $\vartheta$  cordance features.**

To contrast with prediction approaches based on channel-level measures, we trained the RVM using the channel-level  $\alpha$ -band power (Supplementary Fig. 10c) and  $\vartheta$  cordance features, respectively. Cordance is a quantitative EEG measure that has been implicated as a predictive biomarker for antidepressant treatment<sup>13</sup>. In particular, it was reported that lower prefrontal  $\vartheta$  cordance during the placebo lead-in phase predicted better antidepressant (fluoxetine, which is an selective serotonin reuptake inhibitor and venlafaxine, which is a serotonin-norepinephrine reuptake inhibitor) treatment outcome<sup>65</sup>. The  $\vartheta$  cordance was calculated for each participant via the following steps. (1) EEG power re-attribution: absolute re-attributed  $\vartheta$  power of each electrode was calculated as the average of the  $\vartheta$  band (4–7 Hz) power of all bipolar neighboring electrode pairs that share that electrode and absolute re-attributed total power of each electrode was calculated as the average of the total (1–50 Hz) power of all bipolar neighboring electrode pairs that share that electrode. Relative re-attributed power of each electrode was calculated as the absolute re-attributed  $\vartheta$  power divided by the average of the absolute re-attributed total power. (2) Spatial normalization: the absolute and relative re-attributed  $\vartheta$  power values were each normalized by their average across channels. (3) Combination of absolute and relative power:  $\vartheta$  cordance was calculated as the sum of the spatially normalized absolute and relative re-attributed  $\vartheta$  power.

**Treatment prediction using band power features of the latent signals extracted with ICA or PCA.** SELSER was also benchmarked against prediction approaches using band power features of the latent signals estimated by ICA and PCA, respectively. ICA and PCA are widely used unsupervised approaches for estimating latent signals from EEG signals<sup>66,67</sup>, based on different statistical criteria (statistical independence is maximized among latent signals in ICA, whereas variances of the latent signals are maximized in PCA). Over the years, a multitude of ICA algorithms have been developed, as statistical independence could be quantified in a variety of ways<sup>68</sup>. In this work, we used the information maximization (Infomax) algorithm<sup>20</sup> for performing ICA.

To align with the SELSER analyses, the prediction framework based on ICA and PCA followed the same workflow as used in SELSER (Supplementary Fig. 10b). Each subject's EEG signal was first normalized by dividing the square root of its total power across channels. The latent signals were then estimated by applying ICA or PCA to temporally concatenated EEG signals across subjects. After that, the band powers of the full set of latent signals were computed as the

features, followed by an RVM with the linear kernel that related the band power features to the HAMD<sub>17</sub> score changes.

**Performance evaluation using cross-validation.** Stratified tenfold leave-subject-out cross-validation<sup>69</sup> (Supplementary Fig. 2) was employed to assess the predictive performance of each prediction approach. More specifically, the data were randomly partitioned into ten subsets, such that each subset containing an approximately equal number of subjects from each of the four study sites. A subset was left out as the test data, and the remaining nine subsets were used as the training data. For SELSER, the regularization parameter  $\lambda$  was determined using an inner tenfold cross-validation on the training data. The process was then repeated ten times, where each of the ten subsets was used exactly once as test data. As a result, each subject had a predicted HAMD<sub>17</sub> score change. To enhance the stability of the prediction, the data were randomized ten times and the stratified tenfold cross-validation was run on each randomized data. The median of the resulting ten predicted HAMD<sub>17</sub> score changes of each subject was used as the final prediction. The prediction performance was then quantified by the Pearson's correlation coefficient and r.m.s.e. between the cross-validated prediction of the HAMD<sub>17</sub> score change and the true HAMD<sub>17</sub> score change. The  $P$  value for the one-tailed alternative hypothesis that the Pearson's correlation coefficient was greater than 0 was also reported.

Specificity of the prediction was tested by applying, at each fold of stratified  $10 \times 10$  cross-validation, the prediction model to the data from the other treatment arm, which was summarized for each participant by taking the median of the 100 folds of cross-validation.

**Statistical test.** A nonparametric permutation test was used to assess the statistical significance of the treatment prediction results. The observed HAMD<sub>17</sub> score changes were randomly shuffled across subjects 1,000 times. Each time the cross-validated prediction procedure was repeated, resulting in a distribution of the Pearson's correlation coefficient. The  $P$  value was then defined as the proportion of cross-validated correlation coefficients that were greater than the cross-validated correlation coefficient without permutation.

**Application of machine-learning models to independent major depressive disorder data.** *Calculating rsEEG predictions in the second major depressive disorder study.* We applied the result of the  $\alpha$  SELSER model trained on the sertraline arm of the EMBARC sample to data from the second major depressive disorder study in which rsEEG data were collected from 72 patients with depression who were assessed cross-sectionally at the baseline visit. Because 37 patients' EEG data were recorded with a new amplifier (EGI Net Amps 300) distinct from those used in EMBARC, the mean-removal site correction procedure was performed on both the EMBARC data to train the model and the 37 patients' EEG data in the second study, as in the leave-study-site-out analysis. The rsEEG were fed into equation (1) to yield predictions with the SELSER model trained at each fold of cross-validation in EMBARC. The prediction was summarized for each participant by taking the median of the predictions from the 100 folds of cross-validation. This yielded a measure of the EMBARC SELSER model expression strength for each individual in the second major depressive disorder study, expressed as a predicted HAMD<sub>17</sub> change for each patient. Given the focus in the present study on testing the generalizability of the EMBARC sertraline-predictive signature in an independent dataset, we compared the SELSER model-predicted HAMD<sub>17</sub> change scores between patients with a treatment-resistant profile on the ATRQ (failed two or more medications in the current episode) versus those who showed a partial response to antidepressant treatment within-episode.

*Calculating rsEEG predictions in the third major depressive disorder study.* We applied the result of the  $\alpha$  SELSER model trained on the sertraline arm of the EMBARC sample to data from the third major depressive disorder study, in which rsEEG data were collected from 24 patients with depression who were assessed in a cross-sectional manner (without treatment). Since the EEG data were recorded with yet another amplifier distinct from those used in EMBARC, the mean-removal site correction procedure was performed, as for the leave-study-site-out analysis. Due to the difference of electrode montages between this study and EMBARC, the SELSER model trained on EMBARC cannot be directly applied to this study. To address this issue, we source localized the rsEEG from the study based on the linear inverse matrix obtained similarly as in EMBARC and then mapped the source activity to the EMBARC electrodes via the lead-field matrix from EMBARC. Next, the mapped rsEEG at EMBARC electrodes were fed into equation (1) to yield predictions with the SELSER model trained at each fold of cross-validation in EMBARC. The prediction was summarized for each participant by taking the median of the predictions from the 100 folds of cross-validation.

*Calculating task-fMRI predictions in the third major depressive disorder study.* We also applied a previously described RVM model trained on emotional conflict task-fMRI data from EMBARC to data from the third major depressive disorder study<sup>22</sup>. Development of the model on EMBARC data is described in brief below. Extractions were conducted on cortical regions of interest (ROIs), defined based on a recently published cortical parcellation derived from applying a combination

of local gradient analysis and global signal similarity on an independent resting-state fMRI cohort<sup>70</sup>. Since functional parcellations typically rely on resting-state connectivity patterns, which may or may not adequately describe activity patterns in the emotional conflict task, we pooled ROIs from the 200, 400 and 600 region parcellations to limit parcellation-related specificity. ROIs were mapped to seven previously identified functional networks based on the spatial overlap between each ROI and each network<sup>70</sup>. In addition to these cortical ROIs, subcortical ROIs included striatal and cerebellar parcellations based on the same seven functional networks, amygdala ROIs, anterior and posterior hippocampal ROIs and the thalamus. We then regressed imaging sites out of these data using multiple linear regression within the training set at each run of the RVM model and the residualized brain signals were then used for predicting the HAMD<sub>17</sub> score change with the RVM model trained at each fold of cross-validation in EMBARC. The prediction was summarized for each participant by taking the median of the predictions from the 100 folds of cross-validation. The fMRI data from our third major depressive disorder study were preprocessed in the same manner, and the EMBARC-derived weight vector was applied to the extracted ROI data to determine each participant's strength of expression of the EMBARC fMRI RVM model.

**Correlating spTMS and EEG with rsEEG predictions in the third major depressive disorder study.** To quantify in the third major depressive disorder study the correlation between the spTMS and EEG responses and the EMBARC-defined rsEEG phenotype, we employed SELSER to learn predictive models from the spTMS and EEG data to the rsEEG predictions and calculated the leave-one-out cross-validated Pearson's correlation coefficients between the predicted rsEEG predictions and true rsEEG predictions. The SELSER analysis was performed separately for the seven stimulation sites (bilateral pDLPFC, bilateral aDLPFC, bilateral M1 and V1), with the same set of frequency bands as used in the rsEEG prediction analysis ( $\theta$ ,  $\alpha$ ,  $\beta$  and  $\gamma$ ) and for three time windows relative to the TMS pulse (0–200 ms, 200–400 ms and 400–600 ms). For each SELSER analysis, the spTMS and EEG data were concatenated across trials. Significance was evaluated after correcting for the FDR ( $P < 0.05$ ) across all SELSER models (encompassing stimulation sites  $\times$  frequency bins  $\times$  time windows).

**Testing the relationship between rsEEG predictions and treatment outcome in the fourth major depressive disorder study.** We computed each patient's expression of the EMBARC-trained SELSER rsEEG model (expressed as predicted HAMD<sub>17</sub> change) using the same mean site removal procedure as above. We then conducted linear mixed models (SPSS v.25, IBM Corporation) between the SELSER rsEEG-generated predicted HAMD<sub>17</sub> change and outcome on the BDI as well as each of the DASS subscales, separately by rTMS protocol. Terms were time, predicted HAMD<sub>17</sub> change and predicted HAMD<sub>17</sub> change  $\times$  time, using a random intercept and fixed slope. A Bonferroni correction for eight comparisons (two stimulation frequencies and four outcome measures) was then conducted on the predicted HAMD<sub>17</sub> change  $\times$  time results.

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

## Data availability

The EMBARC data are publicly available through the National Institute of Mental Health (NIMH) Data Archive ([https://nda.nih.gov/edit\\_collection.html?id=2199](https://nda.nih.gov/edit_collection.html?id=2199)).

## Code availability

Code for SELSER is available for noncommercial use only at [altoneuroscience.com](http://altoneuroscience.com). For commercial use, please contact Alto Neuroscience at [info@altoneuroscience.com](mailto:info@altoneuroscience.com).

## References

- van Buuren, S. & Groothuis-Oudshoorn, K. Mice: multivariate imputation by chained equations in R. *J. Stat. Softw.* **45**, 1–67 (2011).
- Etkin, A., Egner, T., Peraza, D. M., Kandel, E. R. & Hirsch, J. Resolving emotional conflict: a role for the rostral anterior cingulate cortex in modulating activity in the amygdala. *Neuron* **51**, 871–882 (2006).
- Etkin, A., Buchel, C. & Gross, J. J. The neural bases of emotion regulation. *Nat. Rev. Neurosci.* **16**, 693–700 (2015).
- Egner, T. & Hirsch, J. Cognitive control mechanisms resolve conflict through cortical amplification of task-relevant information. *Nat. Neurosci.* **8**, 1784–1790 (2005).
- Jenkinson, M., Beckmann, C. F., Behrens, T. E., Woolrich, M. W. & Smith, S. M. FSL. *Neuroimage* **62**, 782–790 (2012).
- Friston, K. J. et al. Statistical parametric maps in functional imaging: a general linear approach. *Hum. Brain Mapp.* **2**, 189–210 (1995).
- Wu, W. et al. ARTIST: a fully automated artifact rejection algorithm for single-pulse TMS-EEG data. *Hum. Brain Mapp.* **39**, 1607–1625 (2018).
- Lovibond, P. F. & Lovibond, S. H. The structure of negative emotional states: comparison of the depression anxiety stress scales (DASS) with the Beck depression and anxiety inventories. *Behav. Res. Ther.* **33**, 335–343 (1995).
- Tomioka, R. & Müller, K.-R. A regularized discriminative framework for EEG analysis with application to brain–computer interface. *Neuroimage* **49**, 415–432 (2010).
- Srebro, N. & Jaakkola, T. Weighted low-rank approximations. in *Proc. 20th International Conference on Machine Learning (ICML-03)* 720–727 (AAAI Press, 2003).
- Candès, E. J., Li, X., Ma, Y. & Wright, J. Robust principal-component analysis? *J. ACM* **58**, 11 (2011).
- Parikh, N. & Boyd, S. Proximal algorithms. *Found. Trends Mach. Learn.* **1**, 127–239 (2014).
- Gramfort, A., Papadopoulos, T., Olivi, E. & Clerc, M. OpenMEEG: opensource software for quasistatic bioelectromagnetics. *Biomed. Eng. Online* **9**, 45 (2010).
- Fischl, B. FreeSurfer. *Neuroimage* **62**, 774–781 (2012).
- Pascual-Marqui, R. D. Standardized low-resolution brain electromagnetic tomography (sLORETA): technical details. *Method. Find. Exp. Clin. Pharmacol.* **24**, 5–12 (2002).
- Hunter, A. M., Leuchter, A. F., Morgan, M. L. & Cook, I. A. Changes in brain function (quantitative EEG cordance) during placebo lead-in and treatment outcomes in clinical trials for major depression. *Am. J. Psychiatry* **163**, 1426–1432 (2006).
- Makeig, S., Bell, A. J., Jung, T.-P. & Sejnowski, T. J. Independent component analysis of electroencephalographic data. in *Advances in Neural Information Processing Systems* Vol. 8 (eds Touretzky, D. et al.) (MIT Press, 1996).
- Ghosh-Dastidar, S., Adeli, H. & Dadmehr, N. Principal component analysis-enhanced cosine radial basis function neural network for robust epilepsy and seizure detection. *IEEE Trans. Biomed. Eng.* **55**, 512–518 (2008).
- Cichocki, A. & Amari, S. *Adaptive Blind Signal and Image Processing: Learning Algorithms and Applications*, Vol. 1 (John Wiley & Sons, 2002).
- Witten, I. H., Frank, E., Hall, M. A. & Pal, C. J. *Data Mining: Practical Machine Learning Tools and Techniques* (Morgan Kaufmann, 2016).
- Schaefer, A. et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb. Cortex* **28**, 3095–3114 (2018).

## Acknowledgements

We thank C.J. Keller and S. Kim. The EMBARC study was supported by the National Institute of Mental Health of the National Institutes of Health (NIH) under award nos, U01MH092221 (M.H.T.) and U01MH092250 (P.J.M., R.V.P. and M.M.W.). Data from the second depressed cohort were acquired under R01MH103324 (A.E.) and Big Idea in Neuroscience research funds from the Stanford Neurosciences Institute (A.E.). This work was also funded in part by the Hersh Foundation (M.H.T.). A.E. and W.W. were additionally funded by NIH grant no. DP1 MH116506. W.W. was also funded by National Key Research and Development Plan of China (grant no. 2017YFB1002505) and the National Natural Science Foundation of China (grant nos. 61876063 and 61836003).

## Author contributions

W.W. contributed to the analysis and interpretation of the data and the drafting and revision of the manuscript. Y.Z. and J.J. contributed to the analysis and interpretation of the data and drafting of the manuscript. M.V.L. and G.A.F. contributed to the drafting and revision of the manuscript. C.E.R., C.C., C.C.F., N.K., C.A.C., R.W., R.T., H.M.T., K.M., T.L.C., K.S., M.K.J. and J.M.T. contributed to the conduct of the study, analysis and interpretation of the data, and revision of the manuscript. T.D., P.A., P.J.M., M.M.W. and M.F. contributed to the design and conduct of the study. D.A.P., M.A. and M.H.T. contributed to the design and conduct of the study, and the drafting and revision of the manuscript. A.E. contributed to the design and conduct of the study, the analysis and interpretation of the data and the drafting and revision of the manuscript.

## Competing interests

A.E. (lifetime disclosure) has been receiving salary and equity from Alto Neuroscience since 1 November 2019, to which the pending patent for SELSER has been licensed from Stanford. He holds equity in Mindstrong Health, Akili Interactive and Sizung for unrelated work, has received research funding from the National Institute of Mental Health, Department of Veterans Affairs, Cohen Veterans Bioscience, Brain and Behavior Research Foundation, Dana Foundation, Brain Resource Inc. and the Stanford Neurosciences Institute and consulted for Cervel, Takaeda, Posit Science, Acadia, Otsuka, Lundbeck and Janssen. Over the past 3 years, D.A.P. has received consulting fees from Alkermes, BlackThorn Therapeutics, Boehringer Ingelheim, Posit Science and Takeda Pharmaceuticals. He has received funding from NIMH, the Dana Foundation and Brain and Behavior Research Foundation. T.D.'s research has been funded by NIH, NIMH, National Alliance for Research on Schizophrenia & Depression, TSA, IOCDF, Tufts University, DBDAT and Otsuka Pharmaceuticals. He has received honoraria, consultation fees and/or royalties from the MGH Psychiatry Academy, BrainCells Inc., Clintara, LLC, Inc., Systems Research and Applications Corporation, Boston University, the Catalan Agency for Health Technology Assessment and Research, the National Association of Social Workers Massachusetts, the Massachusetts Medical Society, Tufts University, National Institute of Drug Abuse, NIMH, Oxford University Press, Guilford Press and Rutledge. He has also participated in research funded by DARPA, NIH, NIA, Agency for Healthcare Research and Quality, PCORI, Janssen Pharmaceuticals,

The Forest Research Institute, Shire Development Inc., Medtronic, Cyberonics, Northstar and Takeda. P.M. has received funding from the National Institute of Mental Health, New York State Department of Mental Hygiene, Research Foundation for Mental Hygiene (New York State), Forest Research Laboratories, Sunovion Pharmaceuticals and Naurex Pharmaceuticals (now Allergan). In the past 2 years, M.W. received funding from the NIMH, the National Institute on Drug Abuse, the National Alliance for Research on Schizophrenia and Depression, the Sackler Foundation, the Templeton Foundation; and receives royalties from the Oxford University Press, Perseus Press, the American Psychiatric Association Press and MultiHealth Systems. M.F. has received research support from Abbot Laboratories; Alkermes, Inc.; American Cyanamid; Aspect Medical Systems; AstraZeneca; Avanir Pharmaceuticals; BioResearch; BrainCells Inc.; Bristol-Myers Squibb; CeNeRx BioPharma; Cephalon; Clintara, LLC; Cerecor; Covance; Covidien; Eli Lilly and Company; EnVivo Pharmaceuticals, Inc.; Euthymics Bioscience, Inc.; Forest Pharmaceuticals, Inc.; Ganeden Biotech, Inc.; GlaxoSmithKline; Harvard Clinical Research Institute; Hoffman-LaRoche; Icon Clinical Research; i3 Innovus/Ingenix; Janssen R&D, LLC; Jed Foundation; Johnson & Johnson Pharmaceutical Research & Development; Lichtwer Pharma GmbH; Lorex Pharmaceuticals; Lundbeck Inc.; MedAvante; Methylation Sciences Inc.; National Alliance for Research on Schizophrenia & Depression; National Center for Complementary and Alternative Medicine; National Institute of Drug Abuse; NIMH; Neuralstem, Inc.; Novartis AG; Organon Pharmaceuticals; PamLab, LLC.; Pfizer Inc.; Pharmacia-Upjohn; Pharmaceutical Research Associates., Inc.; Pharmavite LLC; PharmoRx Therapeutics; Photothera; Reckitt Benckiser; Roche Pharmaceuticals; RCT Logic, LLC (formerly Clinical Trials Solutions, LLC); Sanofi-Aventis US LLC; Shire; Solvay Pharmaceuticals, Inc.; Stanley Medical Research Institute; Synthelabo; Tal Medical; and Wyeth-Ayerst Laboratories. He has served as advisor or consultant to Abbott Laboratories; Acadia; Affectis Pharmaceuticals AG; Alkermes, Inc.; Amarin Pharma Inc.; Aspect Medical Systems; AstraZeneca; Auspex Pharmaceuticals; Avanir Pharmaceuticals; AXSOME Therapeutics; Bayer AG; Best Practice Project Management, Inc.; Biogen; BioMarin Pharmaceuticals, Inc.; Biovail Corporation; BrainCells Inc; Bristol-Myers Squibb; CeNeRx BioPharma; Cephalon, Inc.; Cerecor; CNS Response, Inc.; Compellis Pharmaceuticals; Cypress Pharmaceutical, Inc.; DiagnoSearch Life Sciences (P) Ltd.; Dinippon Sumitomo Pharma Co. Inc.; Dov Pharmaceuticals, Inc.; Edgemont Pharmaceuticals, Inc.; Eisai Inc.; Eli Lilly and Company; EnVivo Pharmaceuticals, Inc.; ePharmaSolutions; EPIX Pharmaceuticals, Inc.; Euthymics Bioscience, Inc.; Fabre-Kramer Pharmaceuticals, Inc.; Forest Pharmaceuticals, Inc.; Forum Pharmaceuticals; GenOmind, LLC; GlaxoSmithKline; Grunenthal GmbH; i3 Innovus/Ingenix; Intracellular; Janssen Pharmaceutica; Jazz Pharmaceuticals, Inc.; Johnson & Johnson Pharmaceutical Research & Development, LLC; Knoll Pharmaceuticals Corp.; Labopharm Inc.; Lorex Pharmaceuticals; Lundbeck Inc.; MedAvante, Inc.; Merck & Co., Inc.; MSI Methylation Sciences, Inc.; Naurex, Inc.; Nestle Health Sciences; Neuralstem, Inc.; Neuronetics, Inc.; NextWave Pharmaceuticals; Novartis AG; Nutrition 21; Orexigen Therapeutics, Inc.; Organon Pharmaceuticals; Osmotica; Otsuka Pharmaceuticals; PamLab, LLC.; Pfizer Inc.; PharmaStar; Pharmavite LLC.; PharmoRx Therapeutics; Precision Human Biolaboratory; Prexa Pharmaceuticals, Inc.; Puretech Ventures; PsychoGenics; Psylin Neurosciences, Inc.; RCT Logic, LLC Formerly Clinical Trials Solutions, LLC; Rexahn Pharmaceuticals, Inc.; Ridge Diagnostics, Inc.; Roche; Sanofi-Aventis US LLC.; Sepracor Inc.; Servier Laboratories; Schering-Plough Corporation; Solvay Pharmaceuticals, Inc.; Somaxon Pharmaceuticals, Inc.; Somerset Pharmaceuticals, Inc.; Sunovion Pharmaceuticals; Supernus Pharmaceuticals, Inc.; Synthelabo; Taisho

Pharmaceutical; Takeda Pharmaceutical Company Limited; Tal Medical, Inc.; Tetragenex Pharmaceuticals, Inc.; TransForm Pharmaceuticals, Inc.; Transcept Pharmaceuticals, Inc.; Vanda Pharmaceuticals, Inc.; and VistaGen. He has received speaking or publishing fees from Adamed, Co; Advanced Meeting Partners; American Psychiatric Association; American Society of Clinical Psychopharmacology; AstraZeneca; Belvoir Media Group; Boehringer Ingelheim GmbH; Bristol-Myers Squibb; Cephalon, Inc.; CME Institute/Physicians Postgraduate Press, Inc.; Eli Lilly and Company; Forest Pharmaceuticals, Inc.; GlaxoSmithKline; Imedex, LLC; MGH Psychiatry Academy/Primedia; MGH Psychiatry Academy/Reed Elsevier; Novartis AG; Organon Pharmaceuticals; Pfizer Inc.; PharmaStar; United BioSource, Corp.; and Wyeth-Ayerst Laboratories. He has equity holdings in Compellis and PsyBrain, Inc.; he has a patent for Sequential Parallel Comparison Design, which is licensed by MGH to Pharmaceutical Product Development, LLC (PPD); and a patent application for a combination of Ketamine plus Scopolamine in Major Depressive Disorder, licensed by MGH to Biohaven; and he receives copyright royalties for the MGH Cognitive & Physical Functioning Questionnaire, Sexual Functioning Inventory, ATRQ, Discontinuation-Emergent Signs & Symptoms, Symptoms of Depression Questionnaire, and SAFER; Lippincott, Williams & Wilkins; Wolters Kluwer; and World Scientific Publishing Co. Pte. Ltd. M.H.T. is or has been an advisor/consultant and received fees from (lifetime disclosure): Abbott Laboratories, Inc., Abdi Ibrahim, Akzo (Organon Pharmaceuticals Inc.), Alkermes, AstraZeneca, Axon Advisors, Bristol-Myers Squibb Company, Cephalon, Inc., Cerecor, CME Institute of Physicians, Concert Pharmaceuticals, Inc., Eli Lilly and Company, Evotec, Fabre-Kramer Pharmaceuticals, Inc., Forest Pharmaceuticals, GlaxoSmithKline, Janssen Global Services, LLC, Janssen Pharmaceutica Products, LP, Johnson & Johnson PRD, Libby, Lundbeck, Meade Johnson, MedAvante, Medtronic, Merck, Mitsubishi Tanabe Pharma Development America, Inc., Naurex, Neuronetics, Otsuka Pharmaceuticals, PamLab, Parke-Davis Pharmaceuticals, Inc., Pfizer Inc., PgxHealth, Phoenix Marketing Solutions, Rexahn Pharmaceuticals, Ridge Diagnostics, Roche Products Ltd., Sepracor, SHIRE Development, Sierra, SK Life and Science, Sunovion, Takeda, Tal Medical/Puretech Venture, Targacept, Transcept, VantagePoint, Vivus and Wyeth-Ayerst Laboratories. In addition, he has received grants/research support from: Agency for Healthcare Research and Quality, Cyberonics, Inc., National Alliance for Research in Schizophrenia and Depression, National Institute of Mental Health and the National Institute on Drug Abuse. J.M.T. currently owns stock in Merck and Gilead Sciences and within the past 36 months has previously owned stock in Johnson & Johnson. M.A. holds options from Brain Resource, is Director and Owner of Research Institute Brainclinics, has equity in neuroCare Group and is co-inventor on four patent applications (A61B5/0402; US2007/0299323, A1; WO2010/139361 A1) related to EEG, neuromodulation and psychophysiology, but does not own these nor receives any proceeds related to these patents; he receives Research Institute Brainclinics funding from Brain Resource and neuroCare Group and equipment support from Deymed, neuroConn and Magventure. All other authors report no competing interests.

### Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41587-019-0397-3>.

**Correspondence and requests for materials** should be addressed to A.E.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

## Reporting Summary

Nature Research 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 Research policies, see [Authors & Referees](#) 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

EEG data were collected using the acquisition software provided with the EEG amplifiers. MRI data were collected using conventional imaging acquisition software on each of the MRI platforms noted. Clinical and self-report data were collected by web and paper forms into the StudyTrax database

Data analysis

EEG data were analyzed using the EEGLAB toolbox and custom Matlab code. Imaging data were analyzed using tools from FSL 5 and SPM8, as well as custom Matlab code. Other statistical analyses used SPSS v25.

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/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [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 list of figures that have associated raw data
- A description of any restrictions on data availability

EMBARC data are publicly available through the NIMH data archive

## 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](https://www.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	EMBARC study: 266 patients. Second depression study: 72 patients. Third depression study: 24 patients. Fourth depression study: 177 patients.
Data exclusions	Subjects whose total powers across all the channels were beyond three standard deviations of the mean total power or having too many bad EEG channels were discarded.
Replication	The machine learning models used a 10-fold cross-validation procedure. Leave-site-out analysis was performed for the EMBARC study. The EMBARC model was further validated with the second, third, and fourth depression data sets.
Randomization	EMBARC study: Patients were randomized 1:1 into the sertraline versus placebo treatment arms. Fourth depression study: Selection of the treatment protocol was not done in a randomized manner, but rather in the context of clinical care.
Blinding	For the EMBARC study, all MRI/behavioral data acquisition occurred prior to randomization, and was thus blind to treatment arm. Clinical assessors were also blind to treatment arm.

## 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
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input type="checkbox"/>	<input checked="" type="checkbox"/> Human research participants
<input type="checkbox"/>	<input checked="" type="checkbox"/> Clinical data

### Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input type="checkbox"/>	<input checked="" type="checkbox"/> MRI-based neuroimaging

## Human research participants

Policy information about [studies involving human research participants](#)

### Population characteristics

The participant characteristics are presented here in brief. For the EMBARC study, Written informed consent was obtained from each participant under institutional review board-approved protocols at each of the four clinical sites (University of Texas Southwestern Medical Center, Massachusetts General Hospital, Columbia University, and University of Michigan). Data reported here are based on EMBARC participants who were randomly assigned to sertraline or placebo during stage 1 of the trial (N=309 total). Key eligibility for the study included being 18-65 years old, having major depression as a primary diagnosis by the Structured Clinical Interview for DSM-IV Axis I Disorders, at least moderate depression severity with a score  $\geq 14$  on the Quick Inventory of Depressive Symptomatology-Self Report (QIDS-SR) at screening and randomization, a major depressive episode beginning before age 30, either a chronic recurrent episode (duration  $\geq 2$  years) or recurrent MDD (at least 2 lifetime episodes), and no antidepressant failure during the current episode. Exclusion criteria included: current pregnancy, breastfeeding, no use of contraception; lifetime history of psychosis or bipolar disorder; substance dependence in the past six months or substance abuse in the past two months; unstable psychiatric or general medical conditions requiring hospitalization; study medication contraindication; clinically significant laboratory abnormalities; history of epilepsy or condition requiring an anticonvulsant; electroconvulsive therapy (ECT), vagal nerve stimulation (VNS), transcranial magnetic stimulation (TMS) or other somatic treatments in the current episode; medications (including but not limited to antipsychotics and mood stabilizers); current psychotherapy; significant suicide risk; or failure to respond to any antidepressant at adequate dose and duration in the current episode.

For the second depression study, written informed consent was obtained under institutional review board-approved protocol at UTSW. Individuals were eligible for the study if they were aged 10 years or older, and had the ability to speak, read, and understand English. To be included, participants needed to have a lifetime or current diagnosis of a mood disorder (i.e., major

depressive disorder, persistent depressive disorder, bipolar I/II/NOS, bipolar/mood disorder with psychotic features, depressive disorder other specified (i.e., subthreshold) based upon a semi-structured diagnostic interview.

Exclusion criteria included the following: history of schizophrenia, schizoaffective disorders or chronic psychotic disorders based upon a semi-structured diagnostic interview; inability to provide a stable home address and contact information; having had any condition for which, in the opinion of the investigator or designee, study participation would not be in their best interest (including but not limited to significant cognitive impairment, unstable general medical condition, intoxication, active psychosis) or that could prevent, limit, or confound the protocol-specified assessments; or requirement for immediate hospitalization for psychiatric disorder or suicidal risk as assessed by a licensed study clinician.

For the third depression study, written informed consent was obtained from each participant under institutional review board-approved protocol at Stanford University. Participants in the study underwent several assessments during a clinical intake interview to determine eligibility and classification for the study. The diagnosis of depression (and comorbid conditions) was assessed by a clinician using the SCID, as for EMBARC. Key eligibility for the study included the following: age 18-50 years old; no current psychotherapy; free of metal or ferrous implant; good English comprehension and non-impaired intellectual abilities to ensure understanding of task instructions; no history of neurological disorders, brain surgery, electroconvulsive or radiation treatment, brain hemorrhage or tumor, stroke, epilepsy, hypo- or hyperthyroidism; no daily use of PRN benzodiazepines or opiates (max: 3x/week), or daily thyroid medications, and no antidepressant, anticonvulsant or antipsychotic medications for > 2 weeks (fluoxetine > 6 weeks). Exclusion criteria included the following: left-handed; did not graduate from an English-speaking high school and English was not their first language; psychiatric medications (including but not limited to antipsychotics and mood stabilizers) and hormonal and/or cancer medications; current psychotherapy; current repetitive transcranial magnetic stimulation (rTMS) treatment or electroconvulsive therapy treatment; loss of consciousness greater than 30 minutes and/or a loss of memory greater than 24 hours; lifetime evidence of psychosis, mania, hypomania, or bipolar disorders and/or manic episodes on the SCID; diagnosis of substance dependence within the past 3 months (but not abuse).

For the fourth depression study, patients were drawn from three outpatient mental health care clinics in the Netherlands (neuroCare Clinic Nijmegen, neuroCare Clinic The Hague, and Psychologenpraktijk Timmers Oosterhout) between May 2007 and November 2016. Inclusion criteria included: 1) a primary diagnosis of non-psychotic MDD or dysthymia, 2) Beck Depression Inventory, second edition, Dutch version (BDI-II-NL)<sup>14</sup> at baseline, 3) treatment with at least 10 sessions of rTMS over the DLPFC or response within these 10 sessions. All participants signed an informed consent under an approved IRB protocol. Additional exclusion criteria included: previous ECT treatment, epilepsy, traumatic brain injury, current psychotic disorder, wearing a cardiac pacemaker or metal parts in the head, and current pregnancy.

#### Recruitment

see above

#### Ethics oversight

The EMBARC study was approved by each of the four study sites: University of Texas Southwestern Medical Center, Massachusetts General Hospital, Columbia University, and University of Michigan. The second depression study was approved by University of Texas Southwestern Medical Center. The third depression study was approved by Stanford University. The fourth depression study was approved by neuroCare Clinic Nijmegen, neuroCare Clinic The Hague, and Psychologenpraktijk Timmers Oosterhout

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

Establishing Moderators and Biosignatures of Antidepressant Response for Clinical Care for Depression (EMBARC), NCT#0140709

#### Study protocol

EMBARC used a double-blind design, wherein participants were randomized to an 8-week course of sertraline or placebo. Randomization was stratified by site, depression severity, and chronicity using a block randomization procedure. Sertraline dosing began at 50mg using 50mg capsules and was increased as tolerated if the patient did not respond until a maximum of 200mg 2. A similar dosing approach was used for placebo capsules.

#### Data collection

Data reported here are based on EMBARC participants who were randomly assigned to sertraline or placebo during stage 1 of the trial (N=309 total). Resting-state EEG (rsEEG) were recorded from each of the four study sites: University of Texas Southwestern Medical Center (TX), Massachusetts General Hospital (MG), Columbia University (CU), and University of Michigan (UM). At all study sites, amplifier calibrations were performed. Experimenters were certified by the Columbia EEG team after demonstrating accurate EEG cap placement and delivery of task instructions via video conference, and then submitting satisfactory EEG data from a pilot subject.

rsEEG were recorded during four 2-minute blocks (2 blocks for eyes-closed, and 2 blocks for eyes open) in a counterbalanced order. Participants were instructed to remain still and minimize blinks or eye movements, and to fixate on a centrally presented cross during the eyes-open condition.

#### Outcomes

Our primary outcome was the Hamilton Depression Rating Scale (HAMD17). For participants lacking an endpoint HAMD17, multiple imputation by chained equations was conducted in R3 using the package mice 4. The following observed variables were utilized in order to impute endpoint HAMD17 values for missing data via Bayesian regression: baseline HAMD17, week 1 HAMD17, week 2 HAMD17, week 3 HAMD17, week 4 HAMD17, week 6 HAMD17, baseline Quick Inventory of Depressive Symptoms (QIDS) total score, baseline Mood and Symptom Questionnaire subscale scores for Anxious Arousal, Anhedonic Depression, and General Distress, Snaith-Hamilton Pleasure Scale (SHAPS) total score, age, years of education, gender, and Wechsler Abbreviated Scale of Intelligence (WASI) t-scores for Vocabulary and Matrix Reasoning.

## Magnetic resonance imaging

### Experimental design

Design type	Event-related task-based fMRI
Design specifications	One run of 148 trials, with duration 1s and jittered 3-5s inter-trial intervals
Behavioral performance measures	Button presses were recorded and used to calculate accuracy and reaction time

### Acquisition

Imaging type(s)	fMRI
Field strength	3T
Sequence & imaging parameters	gradient echo spiral-in/out for the second depression study. Acquisition parameters are in supplemental table 3
Area of acquisition	whole brain
Diffusion MRI	<input type="checkbox"/> Used <input checked="" type="checkbox"/> Not used

### Preprocessing

Preprocessing software	FSL tools were used to preprocess imaging data . Functional images were first realigned to structural images using an affine registration matrix and boundary-based registration based upon tissue segmentation as implemented in FSL's FLIRT, which was concatenated with a non-linear normalization of each participant's T1 image to the Montreal Neurological Institute (MNI) 152-person 1 mm3 T1 template using FNIRT from FSL 5.0 to result in a single transformation step from individual native functional space to structurally-aligned and spatially-normalized template space. Functional images were realigned to the middle volume of the run. Nuisance signals corresponding to segmented white matter and CSF were regressed out of motion-corrected functional images. A 6 mm full-width half max (FWHM) isotropic smoothing kernel was then applied to preprocessed time series images to account for individual anatomical variability.
Normalization	non-linear normalization with FNIRT
Normalization template	MNI152
Noise and artifact removal	Six motion parameters were used during first level models
Volume censoring	none

### Statistical modeling & inference

Model type and settings	n/a
Effect(s) tested	n/a...see multivariate modeling information below
Specify type of analysis:	<input checked="" type="checkbox"/> Whole brain <input type="checkbox"/> ROI-based <input type="checkbox"/> Both
Statistic type for inference (See <a href="#">Eklund et al. 2016</a> )	n/a
Correction	n/a

### Models & analysis

- n/a | Involved in the study
- Functional and/or effective connectivity
- Graph analysis
- Multivariate modeling or predictive analysis

Multivariate modeling and predictive analysis	We applied a previously-described Relevance Vector Machine (RVM) model trained on emotional conflict task fMRI data from EMBARC to data from the 2nd MDD study. Development of the model on EMBARC data is described in brief below. Extractions were conducted on cortical regions of interest (ROIs), defined based on a recently-published cortical parcellation derived from applying a combination of local gradient analysis and global signal similarity on an independent resting-state fMRI cohort. Since functional parcellations typically rely on resting-state connectivity patterns, which may or may not adequately describe activity patterns in the emotional conflict task, we pooled ROIs from the 200, 400, and 600 region parcellations in order to limit parcellation-related specificity. ROIs were mapped to seven previously
---	---

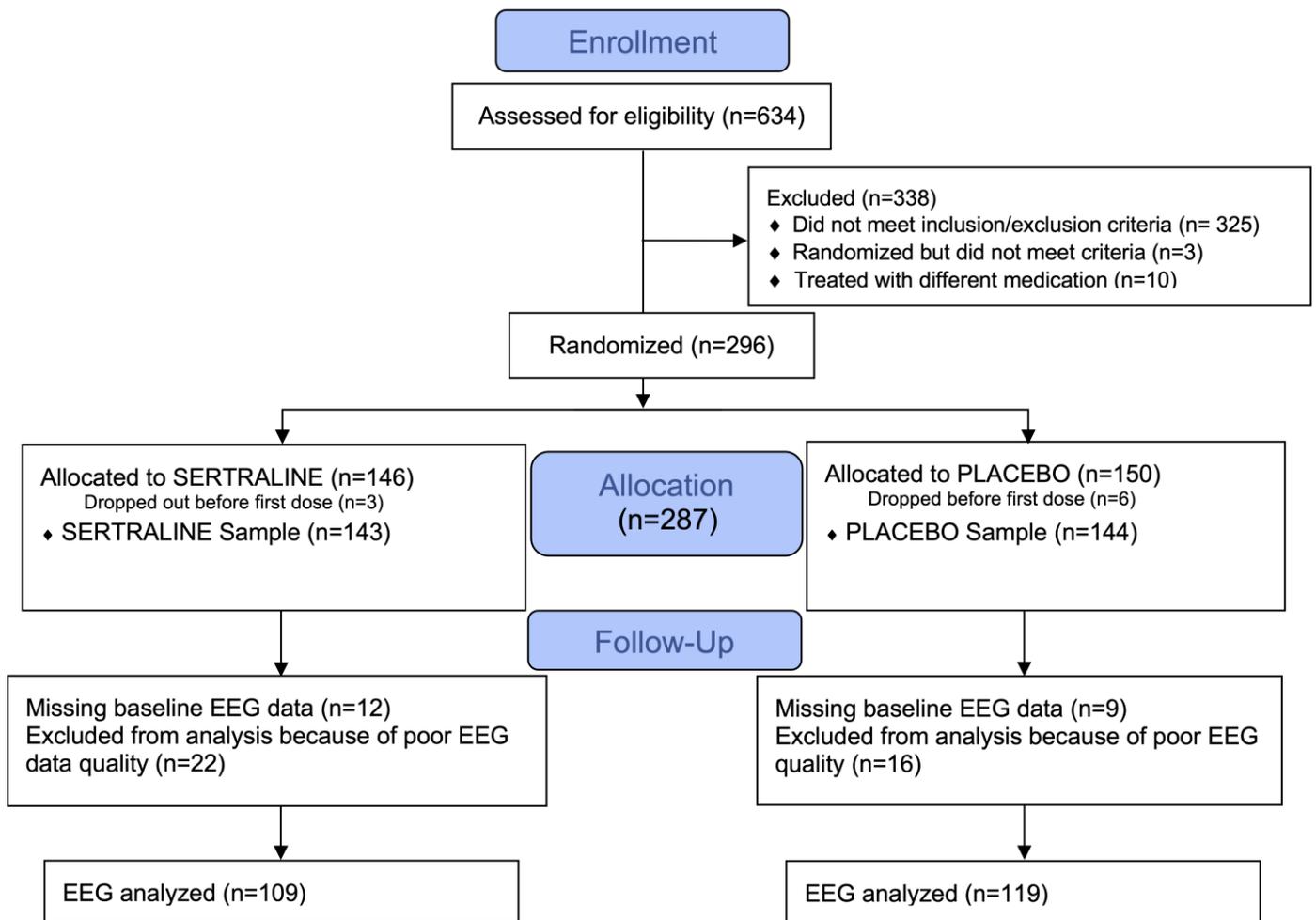
identified functional networks based on the spatial overlap between each ROI and each network. In addition to these cortical ROIs, subcortical ROIs included striatal and cerebellar parcellations based on the same seven functional networks, amygdala ROIs, anterior and posterior hippocampal ROIs and the thalamus<sup>59</sup>. We then regressed imaging site out of these data using multiple linear regression within the training set at each run of the RVM model, and the residualized brain signals were then used for calculating task fMRI predictions with the RVM model trained at each fold of cross-validation in EMBARC. The prediction was summarized for each participant by taking the median of the predictions from the 100 folds of cross-validation. fMRI data from our independent MDD study were preprocessed in the same manner, and the EMBARC-derived weight vector was applied to the extracted ROI data to determine each participant's strength of expression of the EMBARC fMRI RVM model.

In the format provided by the authors and unedited.

# An electroencephalographic signature predicts antidepressant response in major depression

Wei Wu<sup>1,2,3,4</sup>, Yu Zhang <sup>2,3,4</sup>, Jing Jiang<sup>2,3,4</sup>, Molly V. Lucas<sup>2,3,4</sup>, Gregory A. Fonzo<sup>2,3,4</sup>, Camarin E. Rolle<sup>2,3,4</sup>, Crystal Cooper<sup>5,6</sup>, Cherise Chin-Fatt<sup>5,6</sup>, Noralie Krepel<sup>7,8</sup>, Carena A. Cornelissen<sup>2,3,4</sup>, Rachael Wright <sup>2,3,4</sup>, Russell T. Toll<sup>2,3,4</sup>, Hersh M. Trivedi <sup>2,3,4</sup>, Karen Monuszko<sup>2,3,4</sup>, Trevor L. Caudle<sup>2,3,4</sup>, Kamron Sarhadi<sup>2,3,4</sup>, Manish K. Jha<sup>5</sup>, Joseph M. Trombello <sup>5,6</sup>, Thilo Deckersbach<sup>9</sup>, Phil Adams<sup>10</sup>, Patrick J. McGrath<sup>10</sup>, Myrna M. Weissman<sup>10</sup>, Maurizio Fava<sup>9</sup>, Diego A. Pizzagalli <sup>9</sup>, Martijn Arns <sup>7,11,12</sup>, Madhukar H. Trivedi <sup>5,6,13</sup> and Amit Etkin <sup>2,3,4,13</sup> ✉

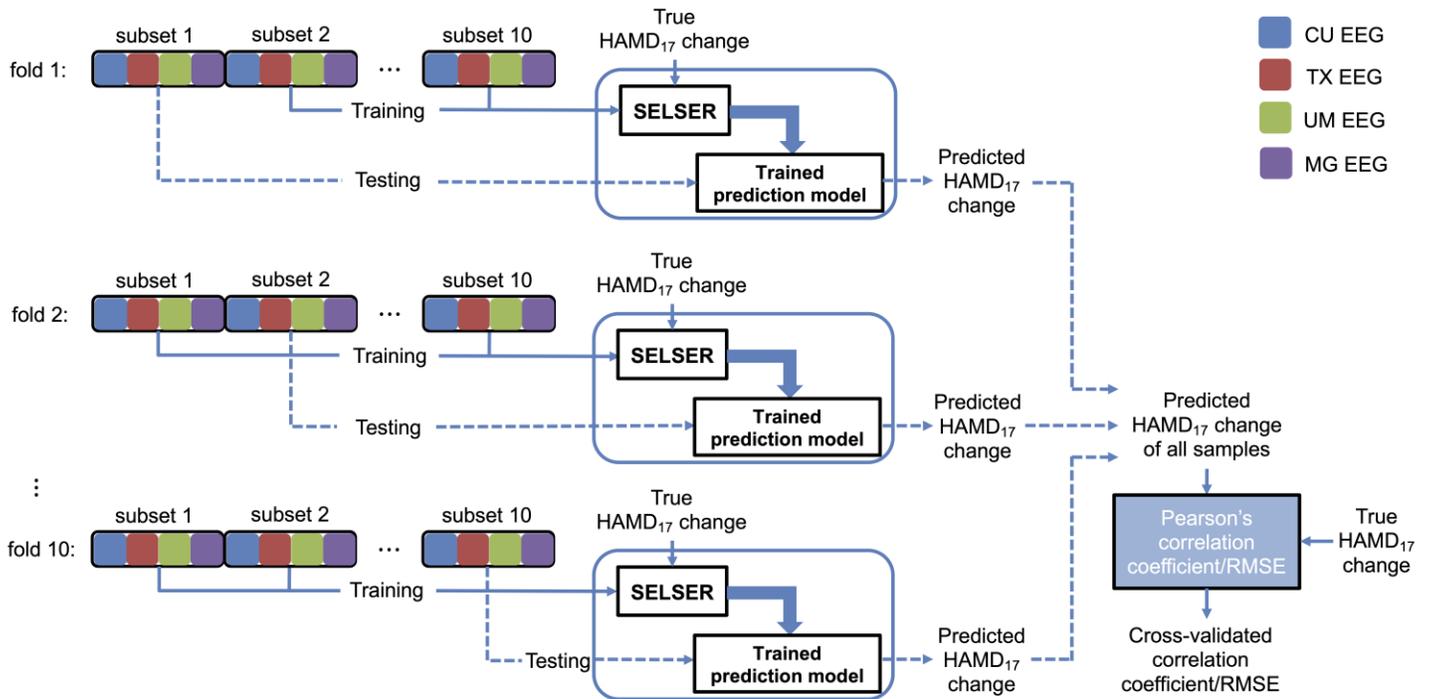
<sup>1</sup>School of Automation Science and Engineering, South China University of Technology, Guangzhou, Guangdong, China. <sup>2</sup>Department of Psychiatry and Behavioral Sciences, Stanford University, Stanford, CA, USA. <sup>3</sup>Wu Tsai Neuroscience Institute, Stanford University, Stanford, CA, USA. <sup>4</sup>Veterans Affairs Palo Alto Healthcare System and the Sierra Pacific Mental Illness, Research, Education and Clinical Center (MIRECC), Palo Alto, CA, USA. <sup>5</sup>Department of Psychiatry, University of Texas Southwestern Medical Center, Dallas, TX, USA. <sup>6</sup>O'Donnell Brain Institute, University of Texas Southwestern Medical Center, Dallas, TX, USA. <sup>7</sup>Research Institute Brainclinics, Brainclinics Foundation, Nijmegen, the Netherlands. <sup>8</sup>Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, the Netherlands. <sup>9</sup>Department of Psychiatry, Harvard Medical School and McLean Hospital, Belmont, MA, USA. <sup>10</sup>New York State Psychiatric Institute & Department of Psychiatry, College of Physicians and Surgeons of Columbia University, New York, NY, USA. <sup>11</sup>Department of Experimental Psychology, Utrecht University, Utrecht, the Netherlands. <sup>12</sup>neuroCare Group Netherlands, Nijmegen, the Netherlands. <sup>13</sup>These authors contributed equally: Madhukar H. Trivedi, Amit Etkin. ✉e-mail: [amitetkin@stanford.edu](mailto:amitetkin@stanford.edu)



**Supplementary Figure 1**

**EMBARC CONSORT Flow Diagram for the patients included in the treatment prediction analyses**

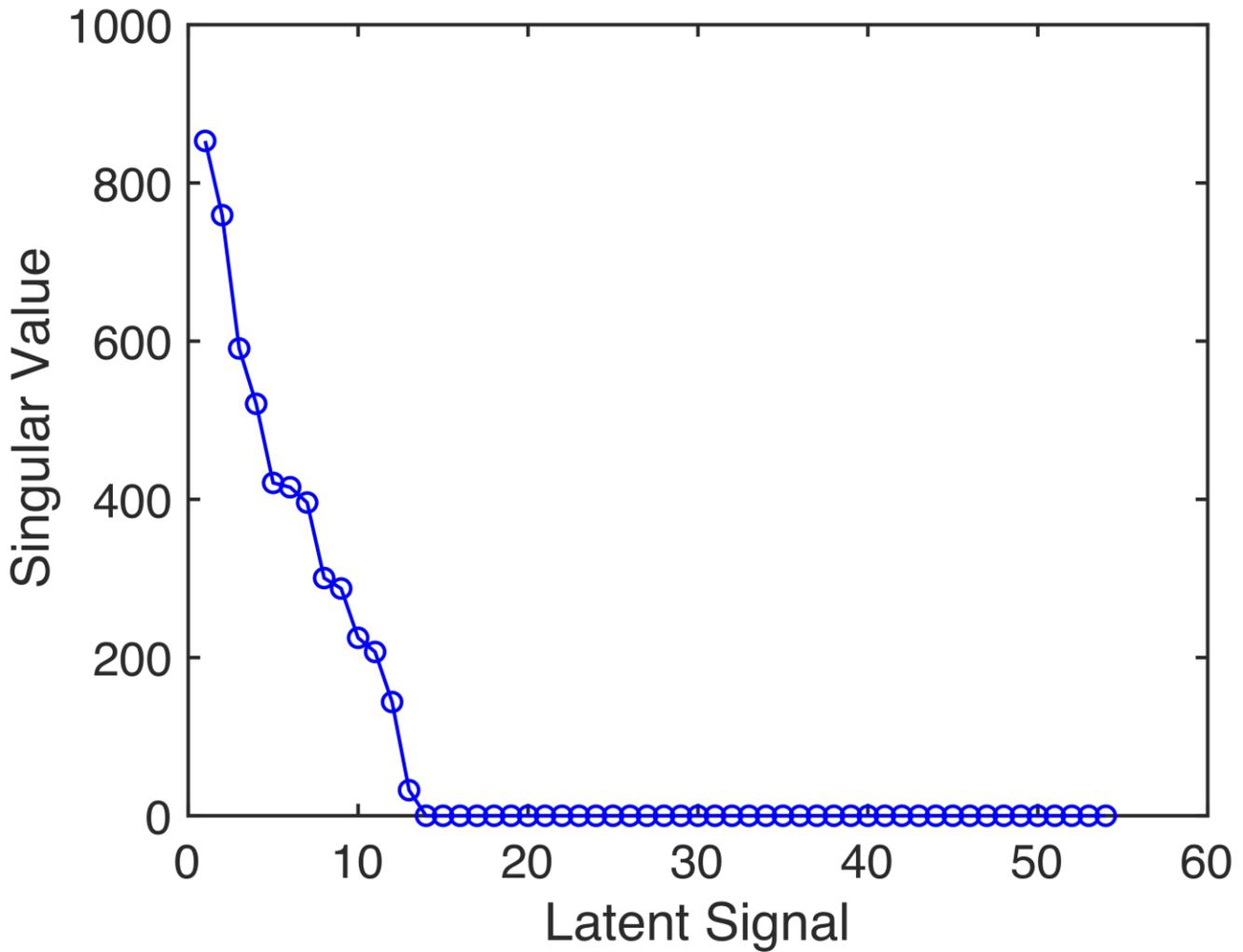
For this analysis, patients were included (1) regardless of their HAMD<sub>17</sub> score, and (2) if they had resting-state EEG data of sufficient quality.



**Supplementary Figure 2**

**Illustration of SERLSER training and evaluation using 10-fold stratified cross-validation.**

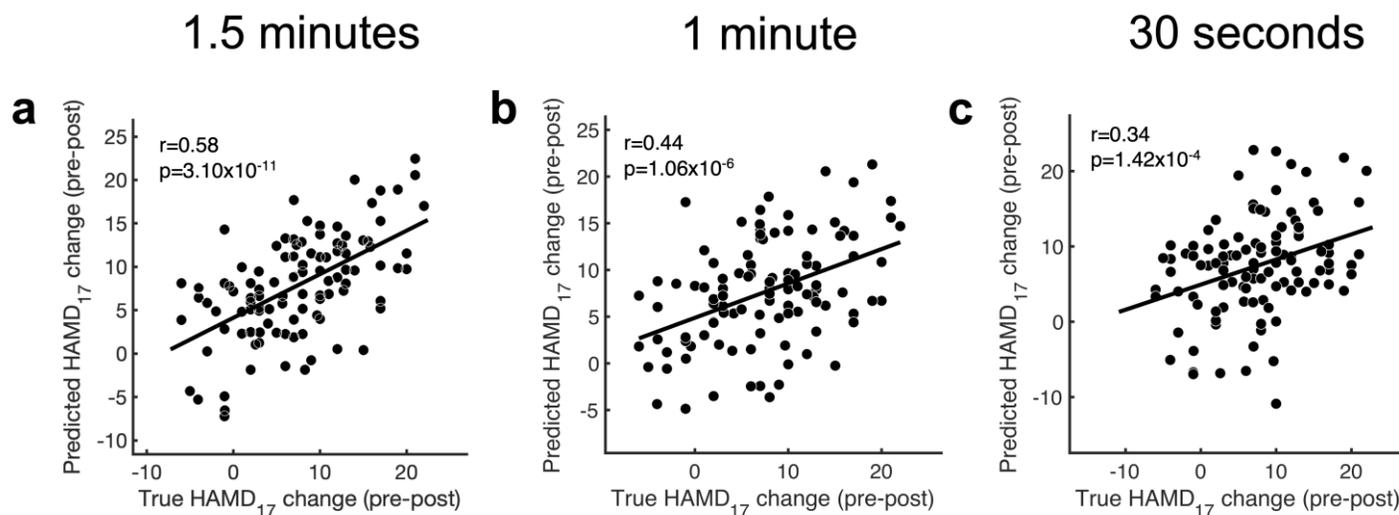
Study sites were Columbia University (CU), University of Texas Southwestern Medical Center (TX), University of Michigan (UM) and Massachusetts General Hospital (MG). Data were randomly partitioned into 10 subsets, such that each subset containing an approximately equal number of subjects from each of the four study sites. A subset was left out as the test data, and the remaining 9 subsets were used as the training data. The process was then repeated 10 times, where each of the 10 subsets was used exactly once as the test data. As a result, each subject had a predicted HAMD<sub>17</sub> score change. The prediction performance was then quantified by the Pearson's correlation coefficient and root mean square error (RMSE) between the cross-validated prediction of the HAMD<sub>17</sub> score change and the true HAMD<sub>17</sub> score change.



**Supplementary Figure 3**

**Singular values associated with alpha SELSER latent signals for the sertraline arm of EMBARC.**

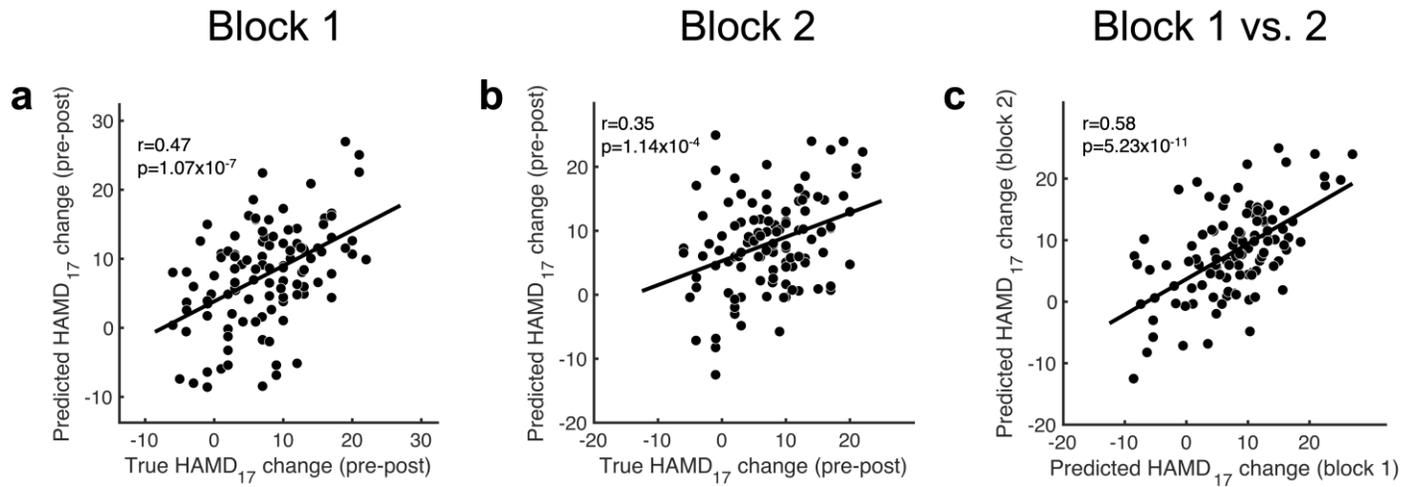
From left to right, the latent signals are sorted according to decreasing singular values (note that the singular values are the absolute values of the eigenvalues for a symmetric matrix). The alpha rsEEG data from all the participants in the sertraline arm were used to train the SELSER model.



**Supplementary Figure 4**

**Prediction of outcome specific to sertraline (n = 109) using SELSER trained on resting eyes open alpha-frequency range data of different lengths.**

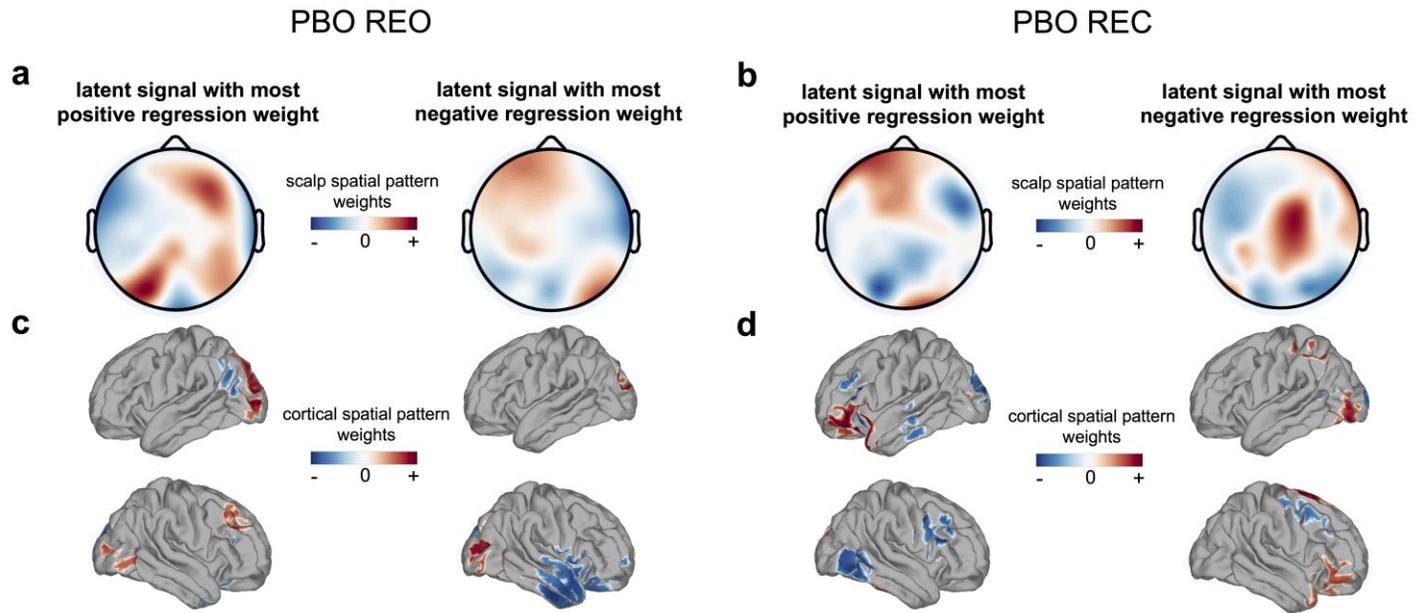
Prediction performance was assessed with 10x10 stratified cross-validation prediction. (a) 1.5 minutes/block. Pearson's  $r = 0.58$ ,  $p = 3.1 \times 10^{-11}$  based on the one-sided test against the alternative hypothesis that  $r > 0$ . (b) 1 minute/block. Pearson's  $r = 0.44$ ,  $p = 1.06 \times 10^{-6}$  based on the one-sided test against the alternative hypothesis that  $r > 0$ . (c) 30 seconds/block. Pearson's  $r = 0.34$ ,  $p = 1.42 \times 10^{-4}$  based on the one-sided test against the alternative hypothesis that  $r > 0$ .



### Supplementary Figure 5

#### Prediction of outcome specific to sertraline (n = 109) using SELSER trained on resting eyes open alpha-frequency range data of different blocks.

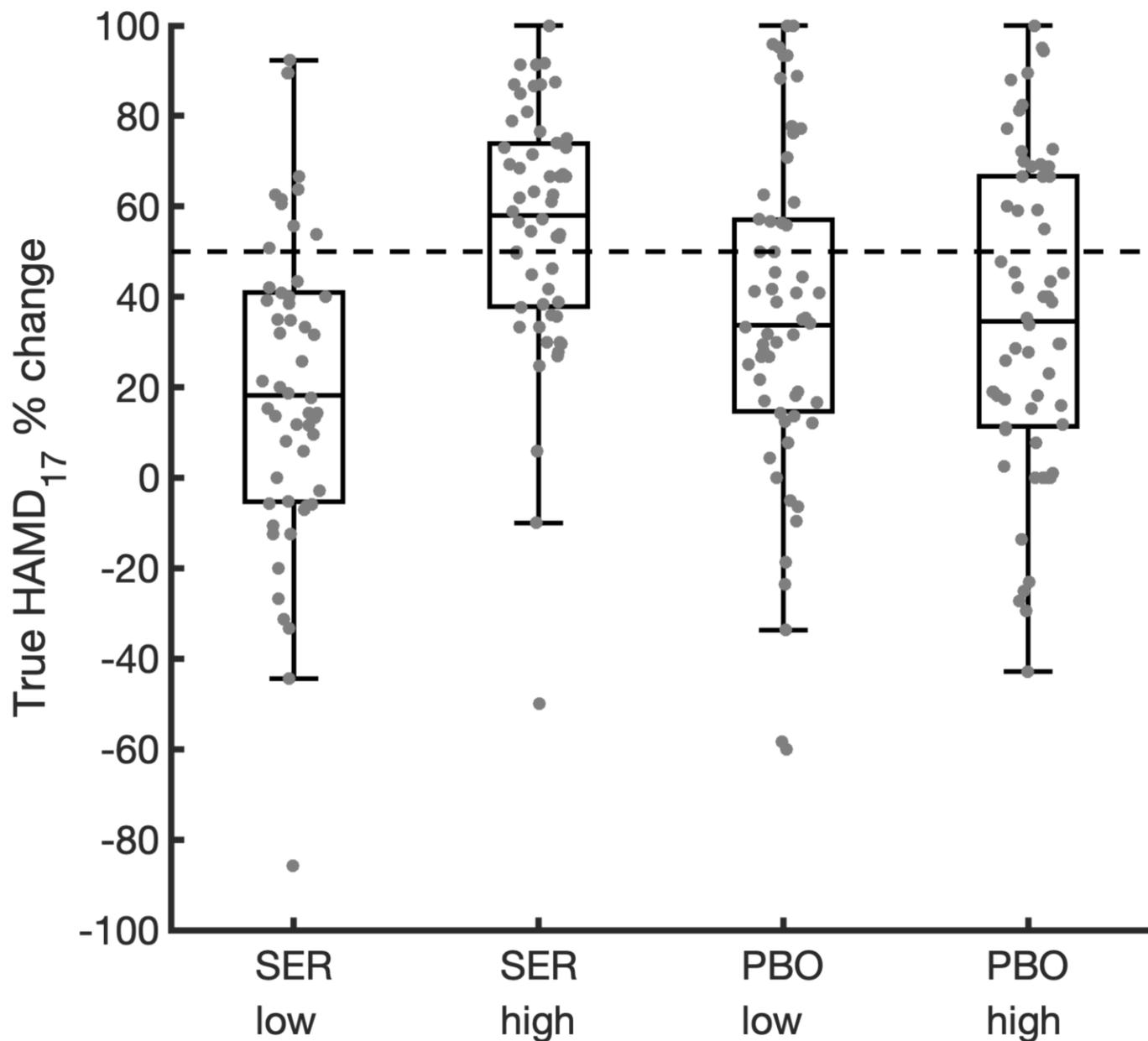
Prediction performance was assessed with 10x10 stratified cross-validation prediction. (a) Block 1. Pearson's  $r = 0.47$ ,  $p = 1.07 \times 10^{-7}$  based on the one-sided test against the alternative hypothesis that  $r > 0$ . (b) Block 2. Pearson's  $r = 0.35$ ,  $p = 1.14 \times 10^{-4}$  based on the one-sided test against the alternative hypothesis that  $r > 0$ . (c) Block 1 prediction vs. block 2 prediction. Pearson's  $r = 0.58$ ,  $p = 5.23 \times 10^{-11}$  based on the one-sided test against the alternative hypothesis that  $r > 0$ .



**Supplementary Figure 6**

**Scalp and cortical spatial patterns of the placebo (PBO) alpha SELSER latent signals (n = 119).**

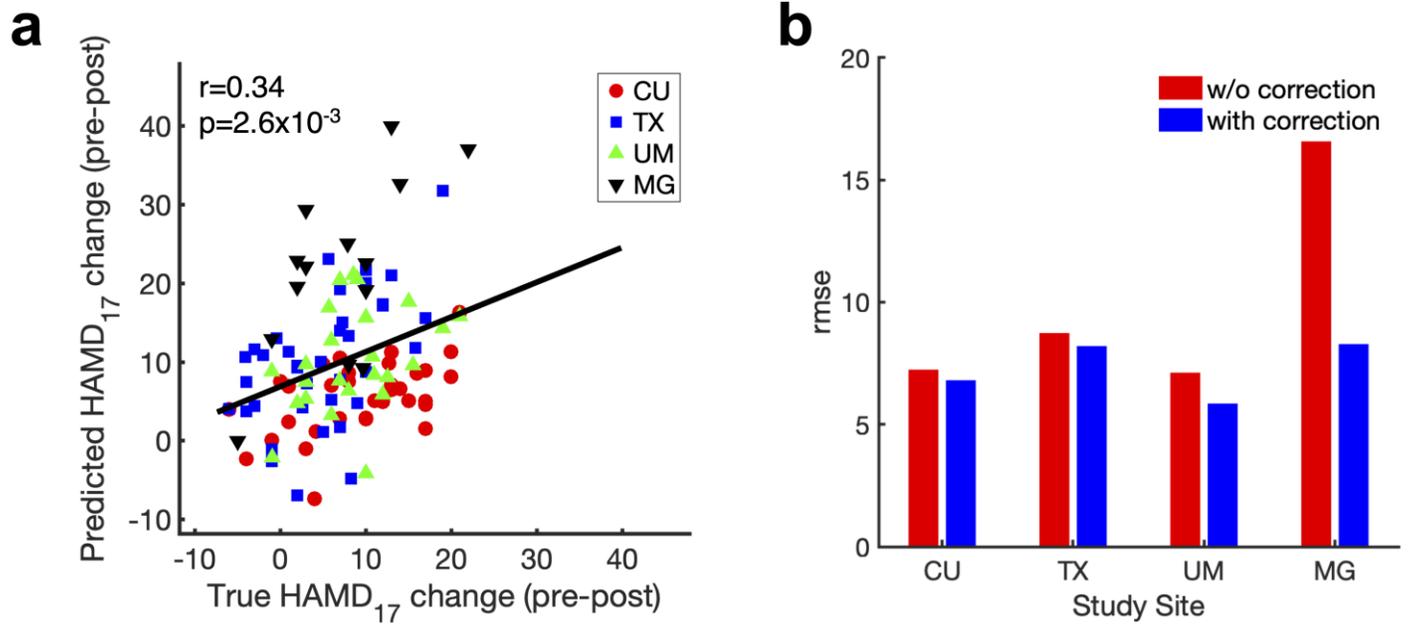
(a, c) Spatial patterns of the SELSER latent signals for the resting eyes open (REO) condition, with the most positive ( $\beta = 556.50.31$ ; left) and negative ( $\beta = -773.49$ ; right) regression weights, respectively. (b, d) Spatial patterns of the SELSER latent signals for the resting eyes closed (REC) condition, with the most positive ( $\beta = 840.85$ ; left) and negative ( $\beta = -801.23$ ; right) regression weights, respectively.



**Supplementary Figure 7**

**Treatment stratification using the rsEEG predictive signature.**

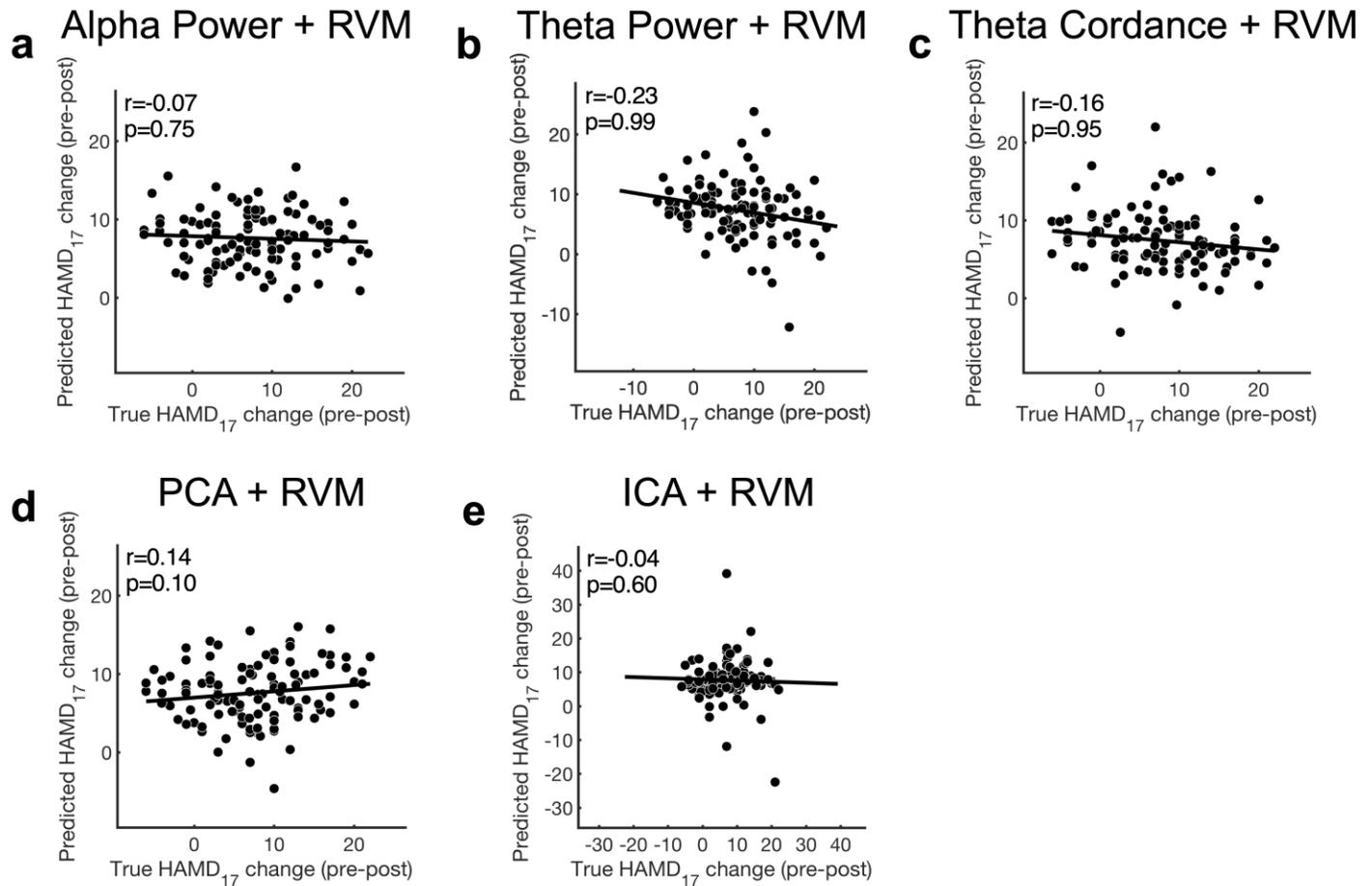
Patients in each arm were partitioned into the low and high groups by applying a median split on the cross-validated predicted HAMD17 score changes for sertraline response.  $n = 109$  for the sertraline arm, and  $119$  for the placebo arm. Each dot represents one patient. For each box, the central line depicts the median, the box extends vertically between the 25th and 75th percentiles, and the whiskers extend to the most extreme data that are not considered outliers. Dashed line indicates 50% change in the true HAMD17 score. SER = sertraline, PBO = placebo.



**Supplementary Figure 8**

**Influence of site correction on leave-study-site-out cross-validation performance (n = 109).**

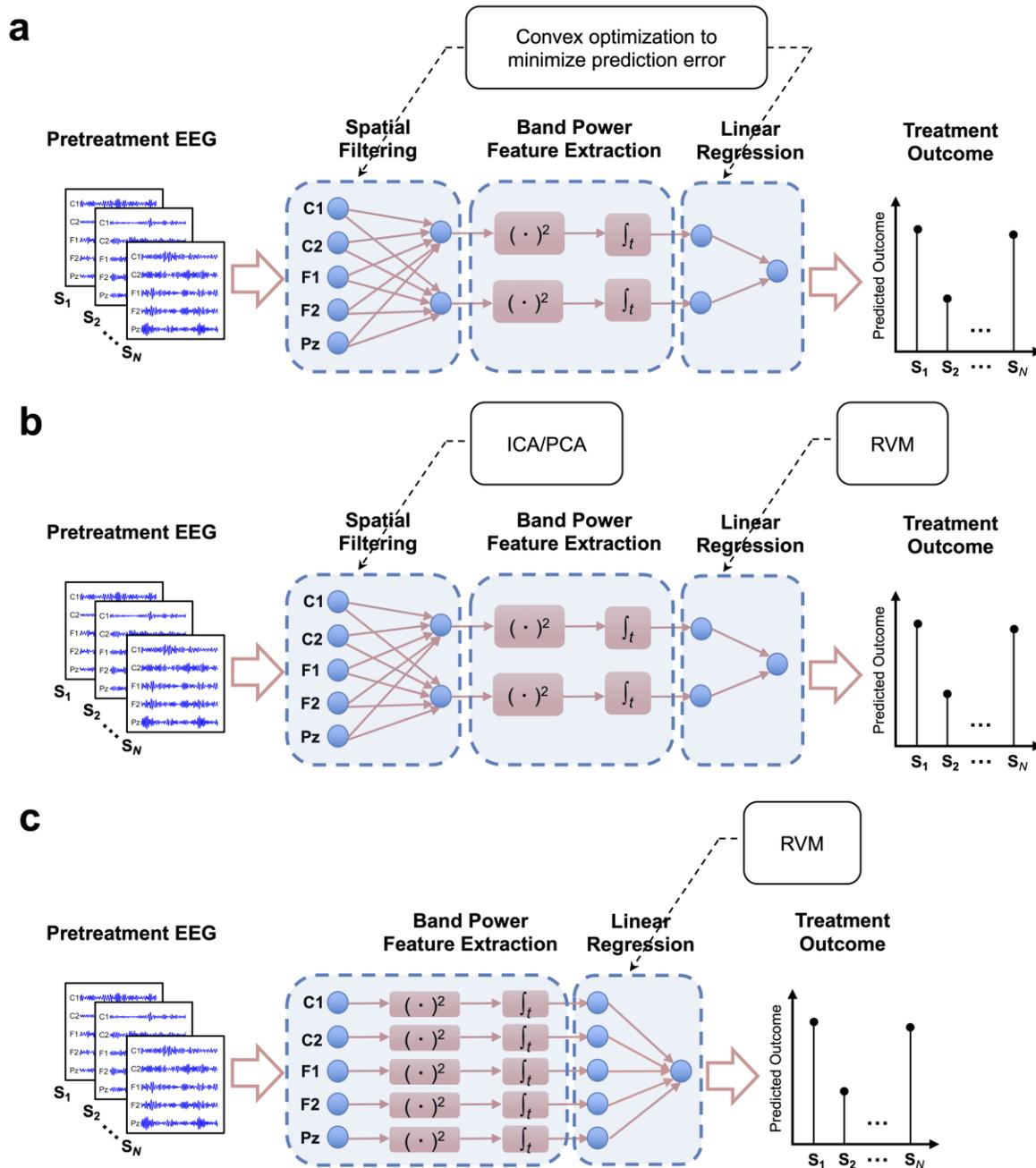
Study sites were Columbia University (CU), University of Texas Southwestern Medical Center (TX), University of Michigan (UM) and Massachusetts General Hospital (MG). Treatment prediction across study sites was assessed by a leave-study-site-out cross-validation on the alpha REO sertraline model. (a) Treatment prediction when site effect was not accounted for. Pearson's  $r = 0.34$ , Bonferroni-corrected  $p = 2.6 \times 10^{-3}$  based on the one-sided test against the alternative hypothesis that  $r > 0$ . (b) Comparison of root mean square error (RMSE) without and with site correction.



**Supplementary Figure 9**

**Machine learning prediction of treatment outcome using previously-suggested predictive metrics (alpha power, theta power, and theta cordance) and conventional latent space modeling approaches (PCA and ICA) on eyes open rsEEG data of the sertraline arm (n = 109).**

10x10 stratified cross-validation prediction using the relevance vector machine (RVM) on channel-level alpha power (a; Pearson's  $r = -0.07$ ,  $p = 0.75$  based on the one-sided test against the alternative hypothesis that  $r > 0$ ), theta power (b; Pearson's  $r = -0.23$ ,  $p = 0.99$  based on the one-sided test against the alternative hypothesis that  $r > 0$ ), theta cordance (c; Pearson's  $r = -0.16$ ,  $p = 0.95$  based on the one-sided test against the alternative hypothesis that  $r > 0$ ), alpha power of the PCA-extracted latent signals (d; Pearson's  $r = 0.14$ ,  $p = 0.1$  based on the one-sided test against the alternative hypothesis that  $r > 0$ ), or alpha power of the ICA-extracted latent signals (e; Pearson's  $r = -0.04$ ,  $p = 0.6$  based on the one-sided test against the alternative hypothesis that  $r > 0$ ) do not significantly predict outcome for sertraline.

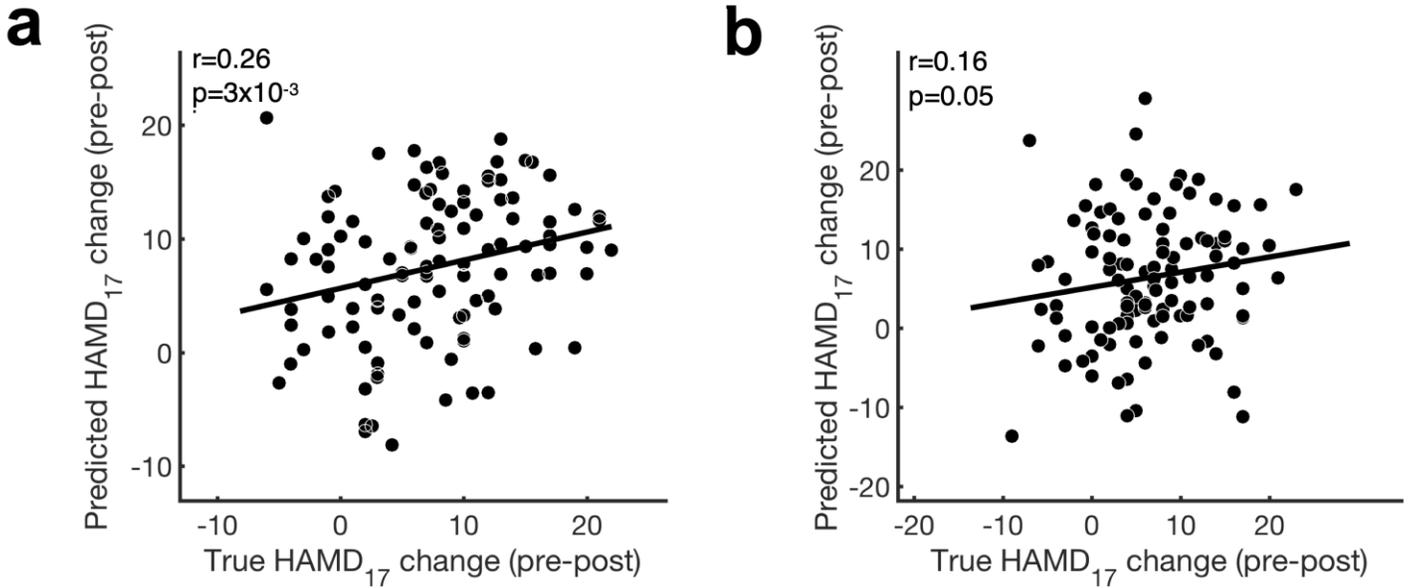


**Supplementary Figure 10**

**Comparison of different band-power based treatment prediction approaches.**

(a) End-to-end prediction with SELSER. All the unknown parameters (spatial filters and linear regression weight coefficients) are optimized in conjunction under a unified objective function via convex optimization. (b) Prediction with ICA/PCA. Spatial filters are optimized via ICA/PCA, and linear regression weight coefficients are optimized via RVM with a linear kernel. (c) Prediction with channel-level band power. EEG band power of each channel is fed directly into the linear regression model, which is optimized via RVM with a linear kernel.  $S_1$ ,  $S_2$ , and  $S_N$  refer to Subject 1, Subject 2, and the  $M$ th Subject, respectively. C1, C2, F1, F2 and Pz refer to electrode locations according to the 10/10 international system.  $(\cdot)^2$  denotes the square operator, and  $\int_t$  denotes the average of a time series over time.

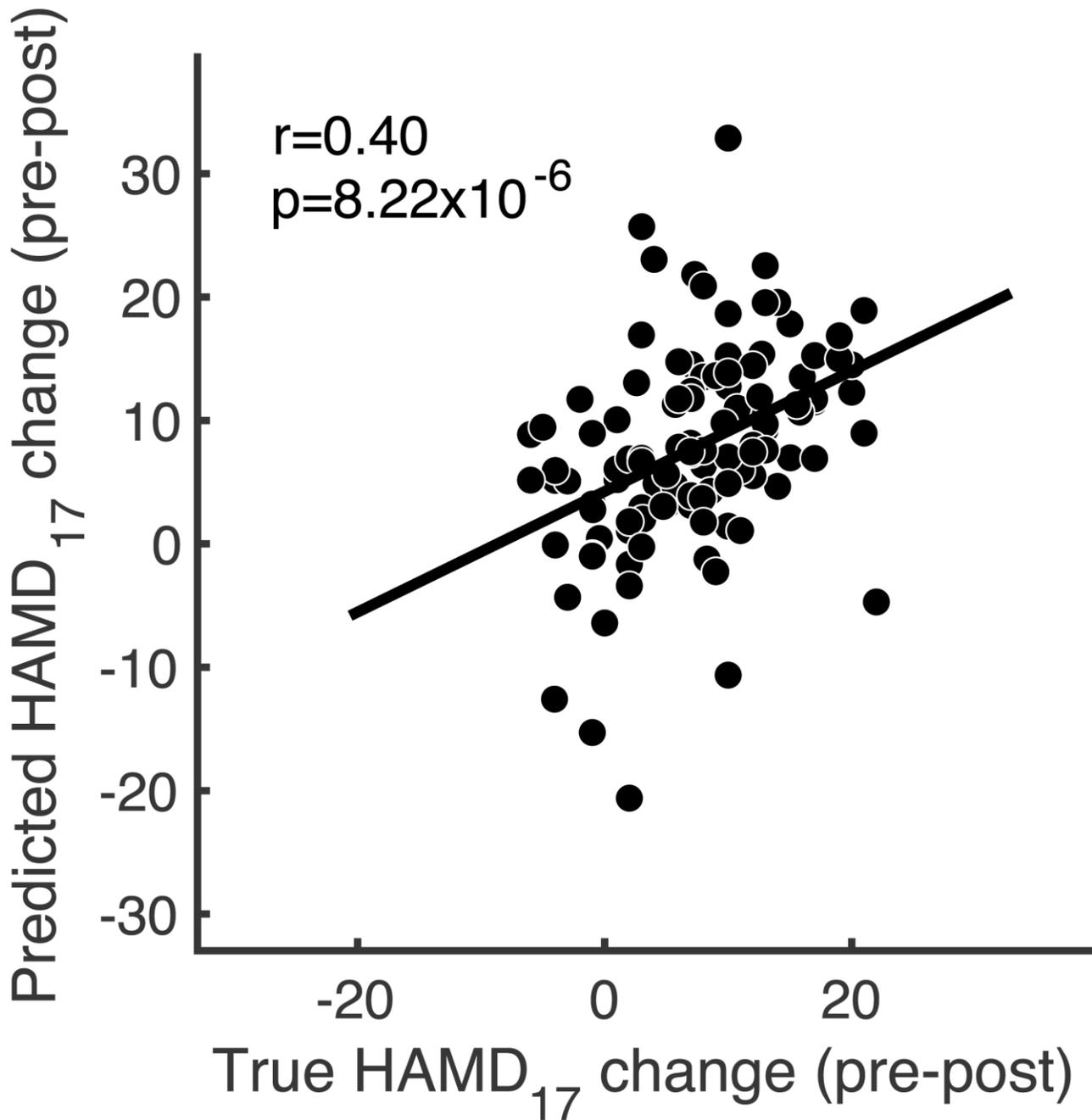
c



### Supplementary Figure 11

#### Machine learning prediction of treatment outcome from symptoms.

Prediction performance was assessed with 10x10 cross-validation prediction using the relevance vector machine (RVM). Included symptom measures were the Spielberger State-Trait Anxiety Inventory, the Quick Inventory of Depressive Symptoms, the Mood and Anxiety Questionnaire, the Childhood Trauma Questionnaire, age, and education. (a) Sertraline arm ( $n = 109$ ). Pearson's  $r = 0.26$ ,  $p = 3 \times 10^{-3}$  based on the one-sided test against the alternative hypothesis that  $r > 0$ . (b) Placebo arm ( $n = 119$ ). Pearson's  $r = 0.16$ ,  $p = 0.05$  based on the one-sided test against the alternative hypothesis that  $r > 0$ .



Supplementary Figure 12

**Prediction of outcome specific to sertraline using SELSER trained on resting eyes open alpha-frequency range data of posterior channels (n = 109).**

A total of 16 posterior electrodes were included: P1, P2, P3, P4, P5, P6, P7, P8, PO3, PO4, PO7, PO8, POz, O1, O2, and Oz. Prediction performance was assessed with 10x10 stratified cross-validation prediction. Pearson's  $r = 0.40$ ,  $p = 8.22 \times 10^{-6}$  based on the one-sided test against the alternative hypothesis that  $r > 0$ . The most positive regression weight is 759.31 and the most negative regression weight is -853.13.

**Supplemental Table 1. Baseline sociodemographic and clinical variables for the EMBARC study.** Statistics reflect comparisons of the Sertraline and Placebo arms

<u>Categorical variables</u>	<u>Sertraline</u>		<u>Placebo</u>		$\chi^2$	p value
	n	%	n	%		
Gender					4.15	0.04
Male	30	27.52	48	40.34		
Female	79	72.48	71	59.66		
Race					1.73	0.42
White	68	62.39	83	69.75		
African American	20	18.35	20	16.81		
Other	21	19.27	16	13.45		
Employment status					0.12	0.94
Employed	58	53.21	66	55.46		
Unemployed	46	42.20	48	40.34		
N/A or Missing	5	4.59	5	4.20		
<u>Continuous variables</u>	Mean	SD	Mean	SD	t value	p value
Age	37.06	13.94	38.41	12.63	-0.77	0.44
Age of onset	16.36	5.90	15.94	5.60	0.55	0.59
Years of education	15.06	2.57	15.36	2.54	-0.87	0.38
Number of MDE	30.78	121.29	45.43	160.15	-0.77	0.44
Duration of current episode	42.70	74.64	51.90	117.91	-0.70	0.49
HAMD <sub>17</sub>	18.16	4.70	18.72	4.41	-0.94	0.35
Medication dose	103.48	32.28	108.67	29.73	-1.18	0.24

Note.  $\chi^2$  = Pearson's chi-squared test. MDE = major depression episodes; \*4 MDD participants (1 placebo, and 3 with sertraline) with no employment status.

**Supplemental Table 2. EEG amplifier settings across study sites for the EMBARC study.**

	<u>CU</u>	<u>TX</u>	<u>UM</u>	<u>MG</u>
<u>Amplifier</u>	BioSemi	NeuroScan Synamp	NeuroScan Synamp	Geodesic Net
<u># channels</u>	72	62	60	129
<u>Sampling rate (Hz)</u>	256	250	250	250
<u>Online filter (Hz)</u>	0-251.3	0-100	0.5-100	0.01-100
<u>Reference electrode</u>	PPO1, PPO2	nose	nose	Cz

Note. CU = Columbia University; TX = University of Texas Southwestern Medical Center; UM = University of Michigan; MG = Massachusetts General Hospital. For the MG site, EEG data were collected at McLean Hospital.

**Supplemental Table 3. Structural MRI and fMRI acquisition scanning parameters for the third MDD study**

<i>Scanner</i>	General Electric 3T 750
<i>Structural</i>	Series = FSPGR BRAVO TR/TE = 8.57ms/3.38ms Flip Angle = 15° Thickness = 1mm Resolution = 0.9375×0.9375mm <sup>2</sup> Duration = 5min
<i>fMRI</i>	Spiral in/out acquisition TR/TE=2000/30msec Flip Angle = 80° Res.=3.4375x3.4375mm <sup>2</sup> Thickness = 4.0mm Matrix = 64×64