A Novel Strategy to Identify Placebo Responders: Prediction Index of Clinical and Biological Markers in
the EMBARC Trial

Madhukar H. Trivedi, M.D.¹, Charles South, Ph.D.¹, Manish Jha, M.B.B.S.¹, A. John Rush, M.D.², Jing Cao, Ph.D.³, Benji Kurian, M.D.¹, Mary Phillips, M.D.⁴, Diego A. Pizzagalli, Ph.D.⁵, Joseph M. Trombello, Ph.D.¹, Maria A. Oquendo, M.D., Ph.D.⁶, Crystal Cooper, Ph.D.¹, Daniel G. Dillon, Ph.D.⁵, Christian Webb, Ph.D.⁵, Bruce D Grannemann, M.A.¹, Gerard Bruder, Ph.D.⁷, Patrick J. McGrath, M.D.⁷, Ramin Parsey, M.D., Ph.D.⁸, Myrna Weissman, Ph.D.⁷, Maurizio Fava, M.D.⁹

1. Department of Psychiatry, University of Texas Southwestern Medical Center
2. Duke-National University of Singapore, Singapore; Duke Medical School, Durham, NC, Texas Tech University Health Sciences Center, Permian Basin, TX
3. Department of Statistical Science, Southern Methodist University, Dallas, TX, USA
4. Department of Psychiatry, University of Pittsburgh School of Medicine
Columbia University
5. Center for Depression, Anxiety and Stress Research, Mclean Hospital,
6. Department of Psychiatry, University of Pennsylvania Perelman School of Medicine
7. New York State Psychiatric Institute & Department of Psychiatry, College of Physicians and Surgeons
8. Stony Brook University School of Medicine
9. Department of Psychiatry, Massachusetts General Hospital

*** Corresponding author**

Madhukar H. Trivedi, M.D.

Professor of Psychiatry, Betty Jo Hay Distinguished Chair in Mental Health,
Director, Center for Depression Research and Clinical Care,
University of Texas Southwestern Medical Center

5323 Harry Hines Blvd., Dallas, TX 75390-9119

Phone: 214-648-0188

Fax: 214-648-0167

Email: madhukar.trivedi@utsouthwestern.edu

Abstract

Background: One-in-three clinical trial patients with Major Depressive Disorder report symptomatic improvement with placebo. Strategies to mitigate the effect of placebo response have focused on modifying study design with variable success. Identifying and excluding or controlling for individuals with high likelihood of responding to placebo may improve clinical trial efficiency and avoid unnecessary medication trials.

Methods: Participants included those assigned to the placebo arm (n=141) of the Establishing Moderators and Biosignatures for Antidepressant Response in Clinical Care (EMBARC) trial. The elastic net was used to evaluate 283 baseline clinical, behavioral, imaging, and electrophysiological variables to identify the most robust yet parsimonious features that predicted depression severity at the end of the double-blind 8-week trial. Variables retained in at least 50% of the 100 imputed datasets were used in a Bayesian multiple linear regression model to simultaneously predict the probabilities of response and remission.

Results: Lower baseline depression severity, younger age, absence of melancholic features or history of physical abuse, less anxious arousal, less anhedonia, less neuroticism, and higher average theta current density in the rostral anterior cingulate predicted higher likelihood of improvement with placebo. The Bayesian model predicted remission and response with a actionable degree of accuracy (both AUC> 0.73). An interactive calculator was developed predicting likelihood of placebo response at the individual level.

Conclusion: Easy-to-measure clinical, behavioral and electrophysiological assessments can be used to identify placebo responders with a high degree of accuracy. Development of this calculator based on these findings can be used to identify potential placebo responders.

Introduction

Major depressive disorder (MDD) affects one in six adults during their lifetime and is estimated to cost the United States over \$200 billion per year [1, 2]. Yet, only one-third of MDD patients achieve remission on any given antidepressant treatment [3], with over one-third of the patients not responding to two or more antidepressants [4, 5]. Despite substantial financial investments [6, 7], efforts to improve MDD treatment outcomes by developing non-monoaminergic antidepressants have failed. High placebo response rates have been a common factor in the failure of several novel antidepressant medications in phase 2 and 3 clinical trials [8, 9]. With limited/variable success, attempts to mitigate the increase in placebo response rate over the last three decades [10] have focused mainly on study design related issues such as: 1) increasing the sample size to account for smaller drug-placebo difference; 2) incorporating a placebo lead-in period; 3) controlling measurement factors by using central raters or standardized interviews; or 4) implementing innovative study designs such as the sequential parallel comparison design (SPCD) [11-14]. A patient-centered approach, which identifies the individual characteristics that define placebo-responders, may offer an alternative way to reduce placebo response rate in clinical trials. These characteristics may be useful, additionally, in clinical practice, where efforts to maximize the placebo response can improve treatment outcomes [15].

The efforts to identify predictors of placebo response have been limited by the focus on subjective disease severity assessment and demographic features. This has frequently resulted in conflicting findings that often did not consider a wide variety of potential predictors from behavioral and biological domains. Among individual factors, low pre-treatment symptom severity has been associated with higher likelihood of placebo response [16, 17]. However, other factors such as gender and age, while significant in venlafaxine vs. placebo studies [18], did not replicate in a meta-analysis by Holmes et al. [19]. The neurobiological basis of placebo response is characterized by an increase in the metabolic activity of the frontal and striatal cortical regions [20] and increased endogenous opioid release in the

subgenual anterior cingulate cortex, nucleus accumbens, midline thalamus and amygdala [21, 22].

Placebo response has also been linked to increased baseline resting state functional connectivity of the rostral anterior cingulate cortex (rACC) within the salience network [23] and to increased pre-treatment rACC activity in two EEG studies [24, 25]. Previous studies of neuroimaging biomarkers of placebo response have been limited by small sample sizes and a lack of comparison with other clinical and biobehavioral markers [21, 23]. Despite extensive research to characterize placebo responders, a set of clinical and objective predictors and tools to filter out this subgroup from clinical trials has yet to be agreed upon and implemented. There is also a potential real-world clinical implication. A portion of treatment-seeking depressed individuals might not need a long-term antidepressant prescription if they are placebo responders. Hence, by identifying such placebo responders in advance, briefer low cost, low side-effect interventions may be recommended for these particular patients.

The goal of this report is to identify a parsimonious set of markers among assessments across units of analyses (clinical, demographic, neuroimaging, electrophysiological, behavioral and cognitive assessments) that most strongly predict the likelihood of placebo response, and can be implemented in research settings and clinical practice. Traditional approaches to handling a statistical problem like this – such a stepwise regression or factor analysis – are not well-equipped to handle a rich database with 1) a large number of predictors relative to the number of subjects or 2) missing data. To attain this goal of identifying the parsimonious set of markers, we utilized data from the Establishing Moderators and Biosignatures for Antidepressant Response in Clinical Care (EMBARC) trial to systematically explore 283 variables using an advanced variable selection method to identify the variables that most strongly predict the likelihood of improvement with placebo. Next, we used a novel Bayesian method to simultaneously predict the degree of symptom change, as well as the probability of remission and response, using these variables. Such a method allows flexibility in the definition of “placebo response” and can be implemented using a simple web-based tool.

Methods

Participants

Participants for this report were recruited as part of the EMBARC trial, which is a two-stage, multi-site, double-blind randomized controlled study designed to evaluate possible moderators and mediators of antidepressant treatment response in patients with MDD. The study design has been described in detail by Trivedi et al. [26]. Data for this report were obtained from participants (n = 141) who were assigned to the placebo arm during stage 1 of the EMBARC trial and received at least one placebo dose. Stage-1 of the EMBARC trial included an 8-week double-blind placebo-controlled trial of sertraline that enrolled 309 participants. Participants were 18-65 years old with MDD diagnosed by the Structured Clinical Interview for DSM-IV Axis I Disorders (SCID). A complete list of inclusion/exclusion criteria and their justification has been previously published [26]. Briefly, participants had to have a Quick Inventory of Depressive Symptomatology-Self Report (QIDS-SR) score ≥ 14 at both screening and randomization visits. Only patients whose first Major Depressive Episode began before age 30, with either a chronic current episode (duration ≥ 2 years) or recurrent MDD (at least 2 lifetime episodes) were eligible. Participants were recruited from four clinical sites (UT Southwestern, Massachusetts General Hospital, Columbia University, and University of Michigan) after approval from the institutional review boards at each site and after obtaining written informed consent from each participant. In the EMBARC trial, 141 participants were assigned to placebo and received at least one dose and constitute the sample for this report.

Outcome

The Hamilton Rating Scale for Depression 17-item (HAMD₁₇): the 17-items of this clinician-rated scale to assess depression severity have 3-5 choices which are scored either from 0-2 or 0-4 [27]. The individual items are summed to measure depression severity [none (<6), mild (6-13), moderate (14-18),

severe (19-23) and very severe (>23)] [27]. In the EMBARC trial, the HAMD₁₇ was administered at each study visit (baseline and weeks 1, 2, 3, 4, 6, and 8 of stage-1). For the purpose of this report, the primary outcome was the last-observed HAMD₁₇ score, which also contributed to the definitions of outcome of remission (defined as last observed HAMD₁₇ ≤ 7) and response (a decrease in HAMD₁₇ of ≥ 50% at the last observation as compared to the baseline). In total, 126 subjects completed 8 weeks of treatment, 5 had their last observed value at week 6, 4 each at weeks 3 and 4, and 1 each at weeks 1 and 2.

Baseline Assessments

Clinical & demographic: 42 clinical and demographic variables were identified *a priori*. In addition to standard demographic characteristics (e.g., sex, age), clinical factors previously reported to have predictive power for treatment outcome were assessed for moderating effects [28]. Clinical variables were derived from the HAMD₁₇, Altman Self-Rating Mania Scale (ASRM), Anger Attacks Questionnaire, Childhood Trauma Questionnaire (CTQ), Concise Associated Symptoms Tracking (CAST), Concise Health Risk Tracking (CHRT), Edinburgh Handedness Inventory (EHI), Family History Screen (FHS), Mood and Anxiety Symptom Questionnaire (MASQ-30), NEO-Five Factor Inventory (NEO), Quick Inventory of Depressive Symptomatology- Self Report (QIDS-SR), Self-administered Comorbidity Questionnaire (SCQ), Snaith-Hamilton Pleasure Scale (SHAPS), Social Adjustment Scale Short Form (SAS SF), Structured Clinical Interview for DSM-IV (SCID) and Wechsler Abbreviated Scale of Intelligence (WASI). Sixteen of the variables had at least some missing data, with the percentage ranging from 0.7% to 11.3%.

Behavioral and cognitive performance: 14 behavioral and cognitive performance measures were assessed for moderating effects. Reaction time, psychomotor slowing, cognitive control, working memory, and reward learning were derived from the following tasks: Choice Reaction Time; Word Fluency Test; “A not B” Working Memory task, Flanker (FT), and Probabilistic Reward Task (PRT) [29]. All variables had at least some missing data, ranging from 0.7% to 17.7%.

Electrophysiological: 15 EEG measures, derived from a pre-treatment EEG recording, were pre-selected for moderator analysis. These included resting (task-free) measures of power in the alpha (recorded during eyes-closed and eyes-open) and theta (eyes-closed) bands, loudness-dependence of auditory evoked potentials (LDEAP), and theta current density extracted from the rACC [25]. For details regarding paradigms and methods, see Tenke et al. [30]. All variables had at least some missing data, ranging from 5.0% to 20.6%.

Neuroimaging: 212 structural or functional imaging variables were analyzed for moderating effects [26]. Structural imaging variables included volumetric magnetic resonance imaging (MRI) assessments using FreeSurfer (FS) and diffusion tensor imaging (DTI) [31, 32]. Functional imaging variables were collected from MRI during resting state (two separate blocks collected on the same day; pre- and post-task), the Reward Processing Task (RPT), and the Emotional Recognition Task (ERT). All variables had at least some missing data, ranging from 28.4% to 34.8%.

Statistical analysis plan

Variable selection method

With the goal to identify a parsimonious set of variables that predict response to placebo, the elastic net [33] – a penalized regression technique that is becoming increasingly popular in the statistical literature [34] with respect to variable selection – was used. The elastic net was selected over other methods such as random forests, support vector machines, gradient boosting machines, and multivariate adaptive regression splines, to ensure the analysis would result in an easily interpretable linear model. The aforementioned alternatives – while popular as tools for prediction – often require large amounts of data [35] and can be difficult to interpret due to the inclusion of complex, non-linear interactions [36]. The elastic net was implemented via the *glmnet* package [37] in R 3.3.3 [38].

Approach for missing data

Due to missing values of predictor baseline variables, multivariate imputation by chained equations was used [39, 40] under the assumption that the data were missing at random [41] to generate a total of 100 imputed datasets. No imputation was used for the outcome variable, which as described above in the outcome sub-section utilized the last-observed HAMDD₁₇ score (last observation carried forward). As there is no universally accepted approach to apply variable selection methods in the context of multiply imputed data [42-44], the elastic net was run independently on each of the 100 datasets. For each run, 10-fold cross-validation was used to tune the models. The variables retained in at least 50% of imputed datasets were used for the Bayesian linear predictor model of placebo outcome.

Model for prediction of placebo outcomes

The variables selected from elastic net were used in a Bayesian linear regression model to predict the HAMDD₁₇ score at acute phase exit using the *bayesreg* package in R [45]. The Bayesian framework was chosen to allay concerns of “double dipping” and applying ordinary least squares (OLS) regression after choosing variables via the elastic net [46, 47]. The Bayesian framework can combat some of these problems by assigning ridge regression prior distributions to the regression parameters, which causes shrinkage on the selected subset of variables that mimics that which occurred when applying the elastic net to all variables. For specific details on the implementation of ridge regression priors, see Makalic and Schmidt [45]. Additionally, the samples from the posterior distributions of the regression parameters can be utilized to estimate the predicted HAMDD₁₇ score at the end of stage-1 and construct a distribution of these predicted scores. The Bayesian framework also handles multiply-imputed data [48] by running separate analyses and then mixing posterior draws from each run [49, 50]. Additionally, the distribution of predicted HAMDD₁₇ scores at the end of stage-1 can be used to estimate the probability of remission (by calculating the proportion of posterior distribution scores ≤ 7 in the distribution) and response (by calculating the proportion of posterior distribution scores $\leq 50\%$ of the baseline HAMDD₁₇ value) for each subject. To correct for some of the biases that exist when training and

testing a model on the same data, each imputed data set was bootstrapped 20 times (resulting in 2000 imputed/bootstrapped datasets: 100 datasets x 20 bootstraps) and the Bayesian model was subsequently fit on each bootstrapped data set to better estimate the variability of the selected predictors; the posterior draws were then mixed as described earlier. While 20 bootstrap repetitions is smaller than one might typically see, given that there were 100 imputed data sets and 2,000 draws from the posterior distributions for each imputed/bootstrapped pair, we ended with a total of 4 million data points – making more than 20 bootstrap replications nearly computationally intractable.

To understand the individual contributions, predictions were made with each variable at a time (univariate AUC) and the decrement to overall model-fit was estimated after removal of variables one at a time.

Sensitivity analysis

In addition to the *a priori* cut-off of retention in at least 50% of the datasets in the elastic net model for selection of variables in the previously-described Bayesian model, the percentage of time each variable was retained was plotted to visually identify drop-offs in clustering of candidate variables. Visual inspection of this ordered index suggested two additional thresholds: variables retained in 100% of elastic net run and those retained in 30% of the runs, see Supplementary Figure 1 for the ordered index. AUC curves are presented along with the comparison of these three thresholds (100%, 50%, and 30%) using the DeLong test [51], implemented in the *pROC* package in R [52].

Visualization of study findings

To demonstrate the practical application of the Bayesian linear regression model in clinical practice, an interactive calculator was developed using Shiny [53].

Results

Of the 141 participants randomized to the placebo arm of the EMBARC trial, the mean (standard deviation, SD) HAMD₁₇ at baseline was 18.6 (4.3) and at study exit was 12.0 (7.5). At study exit, 47/141 (33.3%) attained remission and 55/141 (39.0%) attained response. Baseline clinical and sociodemographic features are presented in Table 1.

Prediction of placebo outcomes

After applying the elastic net to 100 imputed datasets, 8-out-of-283 variables were retained in at least 50% of the runs. These variables, listed in Table 2, included baseline HAMD₁₇, age, anhedonia and anxious arousal as measured by MASQ, neuroticism, the presence of melancholic features, history of physical abuse, and average theta current density in the rACC. These variables were then entered in a Bayesian linear regression model to predict final HAMD₁₇ scores at the end of stage-1 of EMBARC trial. Table 2 gives the posterior median values for the chosen set of predictors. Higher posterior median values predicted higher HAMD₁₇ scores at the end of stage-1 of EMBARC trial, thus signifying worse outcomes. For example, a one (1) unit increase in baseline HAMD₁₇ was associated with a final HAMD₁₇ score that was (on average) 0.52 points higher, holding all other predictors in the model constant. Thus, while higher baseline depression severity, age, neuroticism, anxiety and anhedonia severity, as well as the presence of melancholic features and history of physical abuse predicted worse outcomes with placebo, larger values of pre-treatment theta current density localized to the rACC predicted better outcomes with placebo.

Taking advantage of the posterior distribution of the predicted HAMD₁₇ scores in the Bayesian framework, the probability of remission (HAMD₁₇ ≤ 7) and response (reduction of HAMD₁₇ by ≥ 50%) at the end of stage-1 of EMBARC trial were estimated simultaneously and compared with the observed values at individual participant level; the receiver operating curve (ROC) for both remission and

response are shown in Figure 1. The area under the curve (AUC) values (0.758 for remission and 0.728 for response) indicate a moderate fit, supporting the validity of using a Bayesian linear regression model to derive these values.

In univariate analyses to predict remission, baseline depression severity had the highest AUC and the greatest decrement in AUC to the model, as shown in Table 2. Two other variables with significant univariate AUC (i.e., their bootstrapped confidence interval did not include 0.50) were age and history of physical abuse. Notably, the highest univariate AUC (0.635 for baseline depression severity) was substantially lower than the AUC including all 8 variables (0.758 for remission), thus highlighting the predictive utility of our multivariate model (see Tables 2 and 3).

Further, the probability threshold for classifying subjects as remitters or responders can be optimized based on the desire of individual clinician or researcher. The four quantities often of interest in classification problems are: sensitivity (in this case, the percentage of correctly identified remitters from the entire population of remitters), specificity (the percentage of correctly identified non-remitters from the entire population of non-remitters), positive predictive value (the percentage of predicted remitters who are truly remitters), and negative predictive value (the percentage of predicted non-remitters who are truly non-remitter). Assuming a high certainty (minimum of 70%) to accurately identify participants likely to be a remitter – that is, when you predict someone to be a remitter you will be correct 70% of the time – the optimal probability threshold as identified by the *OptimalCutpoints* package in R will make correct predictions 73.3% of the time and with 23.4% sensitivity [54]. Such a threshold might be desirable for a clinician, who would hope to be quite confident in the likelihood of placebo response before making a treatment decision for his/her patients. Clinical researchers, on the other hand, might be more interested in capturing a larger percentage of placebo responders (i.e. increased sensitivity) – at the cost of more false positives – in order to screen them out of clinical trials to get a purer estimate of a treatment effect.

Sensitivity analyses

In contrast to the 8 variables retained with the *a priori* threshold of 50%, only 3 variables (baseline HAMD₁₇, age, and presence of melancholic features) were retained in all of the 100 elastic net runs and 15 variables were retained in at least 30% of the runs. The additional 7 variables retained in 30% of the runs that indicated improvement in depression severity included lower baseline self-reported depression severity, higher openness to experience, shorter duration of major depressive episode, higher resting-state functional connectivity between right ventral striatum and dorsal anterior cingulate, higher resting-state functional connectivity between left and right ventral striatum as well as left and right insula, and higher Flanker Accuracy effect (calculated as: $[Accuracy_{\text{Congruent trials}} - Accuracy_{\text{Incongruent trials}}]$).

The bootstrapped Bayesian linear regression models with ridge priors were conducted and the estimated root mean squared error (RMSE) for the HAMD₁₇ at the end of stage-1 of EMBARC trial, as well as AUC values for the probability of remission and response, were compared between the *a priori* threshold of 50% retention as well as 100% retention and 30% retention (Table 3). The comparison of AUC of the *a priori* threshold with 100% retention to 50% retention ($p = 0.076$) and 50% retention to 30% retention ($p = 0.055$) showed modest evidence of statistical significance while that of 100% and 30% was stronger ($p = 0.012$). With respect to RMSE, the 50% retention group had an 8% improvement in accuracy over the 100% retention group, with the 30% retention group showing a 13% improvement in accuracy. Taking all this information into account (including the costs of measuring the predictors in each group), the *a priori* defined threshold of retention in 50% of the models provided the best combination of practical and statistical utility.

Visualization of study findings

The *Shiny* package in R [53] was used to adapt the Bayesian linear regression model using the 8 variables identified by the *a priori* threshold of 50% retention in elastic net runs. This interactive

calculator allows the user to adjust the predictor values (based on the range observed in EMBARC trial) and obtain a distribution of posterior predicted HAMD₁₇ scores at the end of 8-week of placebo administration. The posterior distribution of HAMD₁₇ simultaneously allows estimation of the probability of remission and response. A screenshot of this calculator is presented in Figure 2 and will be made available to the broad scientific community using a web-based interface.

Discussion

In this large sample of depressed outpatients, a systematic exploration of a broad range of clinical and biological markers identified baseline depression severity, age, neuroticism, anhedonia, anxious arousal, presence of melancholic features, history of physical abuse, and theta current density in the rACC as predictors of placebo response in a double-blind randomized clinical trial. These markers can be implemented within a Bayesian framework to simultaneously predict the likelihood of response as well as remission with an easy-to-use calculator. Clinicians or clinical researchers could use this calculator to make more informed decisions about whether to prescribe a treatment or enroll subjects in clinical trials with a placebo arm.

The current findings are consistent with previous reports that have found that higher baseline depression severity is associated with lower likelihood of improvement in the placebo arm [10, 17]. The most powerful single variable predicting poorer outcome in response to placebo treatment was greater depression severity. This is a very important finding, as enrichment in terms of severity of illness can be easily implemented in the context of a clinical trial. The failure to reduce placebo response by implementing a greater illness severity threshold for inclusion in studies has been attributed to the “grade inflation” when the severity of illness is determined only by clinicians at the site, where the bias towards enrollment can be significant, thereby driving up severity measures [12]. Independent verifications of subject severity at entry by themselves can reduce the placebo response by allowing a greater depression severity enrichment [55].

The finding that higher theta current density in the rACC predicts greater improvement with placebo differs from previous work by Leuchter et al. which did not find any difference in rACC theta density between placebo responders and non-responders [56]. This difference may be related to the smaller sample size of the Leuchter et al. study. Moreover, in a later study, the same group reported that increased rACC theta current density predicted a greater placebo response. Notably, in the International Study to Predict Optimized Treatment in Depression (iSPOT-D) trial, higher rACC theta was associated with worse outcomes in depressed outpatients treated with either escitalopram, sertraline or placebo [57]. Theta current density and other objective biomarkers, such as resting-state connectivity, may offer greater predictive value at more cost but may prove promising over costly drug treatments that are potentially ineffective or harmful to the outcome.

There are several strengths of this report. The *a priori* defined threshold of 50% retention in elastic net runs have identified clinical and demographic assessments as well as electrophysiological measures, which are relatively inexpensive and easily implemented. The more expensive neuroimaging variables were retained when the threshold for inclusion was more liberal, i.e., included in $\geq 30\%$ of elastic net runs. Arguably, the most meaningful benefit of fitting the model in the Bayesian framework was to take advantage of the posterior distribution of the predicted HAMD₁₇ scores to simultaneously predict the likelihood of response and remission. In a more traditional analysis, three separate models would have to be run to accomplish this: a multiple linear regression to predict HAMD₁₇ at exit and two logistic regression models, one with remission as the outcome and one with a response as the outcome. Additionally, the interactive calculator informed by our predictive model provides a web-based tool, which could be easily implemented in clinical practice or research settings to predict the likelihood of placebo response at the individual level.

These findings should be interpreted in the context of certain limitations. While extensive, the variable selection procedure may have missed some additional clinical or biological features that could

increase the predictive ability of placebo response. However, when identified, such features may be integrated into future iterations of the calculator. The choice of elastic net was deliberate to maximize the practical application of these findings; however, this method may have missed more complicated (i.e., non-linear) interactions between features that might be uncovered with other variable selection methods such as random forest. Further, due to the unique combination of variables collected for the EMBARC trial, we cannot validate the model on an external data set. While the cross-validation and bootstrap procedures employed add strength to the analysis, but there is no substitute for external validation.

To conclude, a set of 8 clinical and biological markers can predict treatment outcomes with placebo with a fair degree of accuracy. It is also possible to further increase the accuracy of prediction if all 15 variables are included. By integrating these markers in an easy-to-use interactive calculator, the findings of this report can be implemented in research and clinical care.

Acknowledgements

Dr. Pizzagalli: funding from NIMH and the Dana Foundation; consulting fees from BlackThorn Therapeutics, Boehringer Ingelheim, Pfizer Inc. and Posit Science. **Dr. Weissman:** funding from NIMH, the National Alliance for Research on Schizophrenia and Depression (NARSAD), the Sackler Foundation, and the Templeton Foundation; royalties from the Oxford University Press, Perseus Press, the American Psychiatric Association Press, and MultiHealth Systems. **Dr. Trivedi** has received consulting fees from or has served on the advisory boards of Alkeremes Inc., Akili Interactive, Allergan Pharmaceuticals, Arcadia Pharmaceuticals, Avanir Pharmaceuticals, Brintellix Global, Bristol Myers Squibb, Caudex, Cerecor, Forest Pharmaceuticals, Global Medical Education Inc, Health Research Associates, Insys, Johnson & Johnson Pharmaceutical Research & Development, Lilly Research Laboratories, Lundbeck Research USA, Medscape, Merck & Co. Inc, Mitsubishi Pharma, MSI Methylation Sciences – PamLab Inc., Navitor, Otsuka America Pharmaceutical Inc., One Carbon Therapeutics, Otsuka America Pharmaceutical Inc., Pfizer Inc, and Takeda Global Research; royalties from Janssen Research and Development LLC; honoraria from the American Psychiatric Association; grants from the Agency for Healthcare Research and Quality (AHRQ), Cancer Prevention and Research Institute of Texas (CPRIT), National Institute of Mental Health (NIMH), National Institute of Drug Abuse (NIDA), National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK), National Center for Advancing Translational Sciences (NCATS), Johnson & Johnson, PCORI; and holds author agreements with Janssen Asia Pacific and Oxford University Press. **Drs. Trombello** currently owns stock in Merck and Gilead Sciences and within the past 36 months previously owned stock in Johnson & Johnson. **Dr. Fava** has received research support from Abbott Laboratories, Alkermes, Inc., American Cyanamid, Aspect Medical Systems, AstraZeneca, Avanir Pharmaceuticals, BioResearch, BrainCells Inc., Bristol-Myers Squibb, CeNeRx BioPharma, Cephalon, Cerecor, Clintara, LLC, 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., NARSAD, National Center for Complementary and Alternative Medicine, Neuralstem, Inc., NIDA, NIMH, 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; has served on the advisory board or consulted for 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., Pharmaceutical Product Development, LLC, 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 Therapeutics; 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, Lippincott Williams & Wilkins, MGH Psychiatry Academy/Primedia, MGH Psychiatry Academy/Reed Elsevier, Novartis AG, Organon Pharmaceuticals, Pfizer Inc., PharmaStar, United BioSource Corp., Wolters Kluwer, World Scientific Publishing Co. Pte. Ltd., and Wyeth-Ayerst Laboratories; has equity holdings in Compellis and PsyBrain, Inc.; holds patents for Sequential Parallel Comparison Design, licensed by MGH to Pharmaceutical Product Development, LLC, and patent application for a combination of Ketamine plus Scopolamine in Major Depressive Disorder, licensed by MGH to Biohaven; and holds copyrights for the MGH Cognitive & Physical Functioning Questionnaire, Sexual Functioning Inventory, Antidepressant Treatment Response Questionnaire, Discontinuation-Emergent Signs & Symptoms, Symptoms of Depression Questionnaire, and SAFER. MW received funding from NIMH, NIDA, NARSAD, Sackler Foundation, and Templeton Foundation and receives royalties from Oxford University Press, Perseus Press, American Psychiatric Association Press, and MultiHealth Systems. MHT is or has been an advisor/consultant and received fees from Abbott Laboratories, Inc., Abdi Ibrahim, Akzo (Organon Pharmaceuticals Inc.), Alkermes,

AstraZeneca, Axon Advisors, Bristol-Myers Squibb, Cephalon, Inc., Cerecor, CME Institute of Physicians, Concert Pharmaceuticals, Inc., Eli Lilly and Company, Evotec, Fabre Kramer Pharmaceuticals, Inc., Forest Pharmaceuticals, Inc., GlaxoSmithKline, Janssen Global Services, LLC, Janssen Pharmaceutica Products, LP, Johnson & Johnson Pharmaceutical Research & Development, LLC, 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 Inc., Sierra, SK Life and Science, Sunovion, Takeda, Tal Medical/Puretech Venture, Targacept, Transcept, VantagePoint, Vivus, and Wyeth-Ayerst Laboratories and has received grants and/or research support from Agency for Healthcare Research and Quality, Cyberonics, Inc., NARSAD, NIMH, and NIDA. **Dr. Kurian** has received research grant support from Targacept, Inc., Pfizer Inc., Johnson & Johnson, Evotec, Rexahn, Naurex, Forest Pharmaceuticals, Inc., and NIMH. **Dr. McGrath** has received research grant support from Forest Research Laboratories, Sunovion Pharmaceuticals, and Naurex. **Dr. Oquendo** receives royalties for use of the Columbia Suicide Severity Rating Scale. Her family owns stock in Bristol Myers Squibb. **Dr. Rush** has received consulting fees from Akili, Brain Resource Inc., Compass Inc., Curbstone Consultant LLC, Eli Lilly, Emmes Corp, Liva-Nova, Mind Linc., Sunovion, Takeda USA, Taj Medical; speaking fees from Liva-Nova and Sing-Health; and royalties from Guilford Press and the University of Texas Southwestern Medical Center, Dallas, TX. (for the Inventory of Depressive Symptoms and its derivatives). He is also named co-inventor on two patents: U.S. Patent No. 7,795,033: Methods to Predict the Outcome of Treatment with Antidepressant Medication, Inventors: McMahon FJ, Laje G, Manji H, Rush AJ, Paddock S, Wilson AS and U.S. Patent No. 7,906,283: Methods to Identify Patients at Risk of Developing Adverse Events During Treatment with Antidepressant Medication, Inventors: McMahon FJ, Laje G, Manji H, Rush AJ, Paddock S. **Drs. South, Jha, Cao, Phillips, Cooper, Dillon, Webb, Bruder, and Parsey, and Mr. Grannemann** have no conflicts to report.

References

- 1 Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, Rush AJ, Walters EE, Wang PS: The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). *Jama* 2003;289:3095-3105.
- 2 Greenberg PE, Fournier AA, Sisitsky T, Pike CT, Kessler RC: The economic burden of adults with major depressive disorder in the United States (2005 and 2010). *The Journal of clinical psychiatry* 2015;76:155-162.
- 3 Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, Norquist G, Howland RH, Lebowitz B, McGrath PJ, Shores-Wilson K, Biggs MM, Balasubramani GK, Fava M: Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. *The American journal of psychiatry* 2006;163:28-40.
- 4 Rush AJ, Trivedi MH, Wisniewski SR, Nierenberg AA, Stewart JW, Warden D, Niederehe G, Thase ME, Lavori PW, Lebowitz BD, McGrath PJ, Rosenbaum JF, Sackeim HA, Kupfer DJ, Luther J, Fava M: Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. *The American journal of psychiatry* 2006;163:1905-1917.
- 5 Souery D, Papakostas GI, Trivedi MH: Treatment-resistant depression. *The Journal of clinical psychiatry* 2006;67 Suppl 6:16-22.
- 6 Murrough JW, Charney DS: Is there anything really novel on the antidepressant horizon? *Current psychiatry reports* 2012;14:643-649.
- 7 Nutt D, Goodwin G: ECNP Summit on the future of CNS drug research in Europe 2011: Report prepared for ECNP by David Nutt and Guy Goodwin. *European Neuropsychopharmacology* 2011;21:495-499.
- 8 DiMasi JA, Feldman L, Seckler A, Wilson A: Trends in risks associated with new drug development: success rates for investigational drugs. *Clinical pharmacology and therapeutics* 2010;87:272-277.
- 9 Schatzberg AF: Issues encountered in recent attempts to develop novel antidepressant agents. *Annals of the New York Academy of Sciences* 2015;1345:67-73.
- 10 Walsh B, Seidman SN, Sysko R, Gould M: Placebo response in studies of major depression: Variable, substantial, and growing. *Jama* 2002;287:1840-1847.
- 11 Desseilles M, Witte J, Chang TE, Iovieno N, Dording C, Ashih H, Nyer M, Freeman MP, Fava M, Mischoulon D: Massachusetts General Hospital SAFER Criteria for Clinical Trials and Research. *Harvard Review of Psychiatry* 2013;21:269-274.
- 12 Fava M, Evins AE, Dorner DJ, Schoenfeld DA: The problem of the placebo response in clinical trials for psychiatric disorders: culprits, possible remedies, and a novel study design approach. *Psychotherapy and psychosomatics* 2003;72:115-127.
- 13 Trivedi MH, Rush J: Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? *Neuropsychopharmacology* 1994;11:33-43.
- 14 Bret R. Rutherford, Steven P. Roose: A Model of Placebo Response in Antidepressant Clinical Trials. *American Journal of Psychiatry* 2013;170:723-733.
- 15 Enck P, Bingel U, Schedlowski M, Rief W: The placebo response in medicine: minimize, maximize or personalize? *Nature reviews Drug discovery* 2013;12:191-204.
- 16 Fournier JC, DeRubeis RJ, Hollon SD, et al.: Antidepressant drug effects and depression severity: A patient-level meta-analysis. *Jama* 2010;303:47-53.
- 17 Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT: Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. *PLoS medicine* 2008;5:e45.

- 18 Entsua R, Vinall P: Potential predictors of placebo response: Lessons from a large database. *Drug information journal* 2007;41:315-330.
- 19 Holmes RD, Tiwari AK, Kennedy JL: Mechanisms of the placebo effect in pain and psychiatric disorders. *The pharmacogenomics journal* 2016;16:491-500.
- 20 Mayberg HS, Silva JA, Brannan SK, Tekell JL, Mahurin RK, McGinnis S, Jerabek PA: The functional neuroanatomy of the placebo effect. *The American journal of psychiatry* 2002;159:728-737.
- 21 Pecina M, Bohnert AS, Sikora M, Avery ET, Langenecker SA, Mickey BJ, Zubieta JK: Association Between Placebo-Activated Neural Systems and Antidepressant Responses: Neurochemistry of Placebo Effects in Major Depression. *JAMA psychiatry* 2015;72:1087-1094.
- 22 Fava M: Implications of a Biosignature Study of the Placebo Response in Major Depressive Disorder. *JAMA psychiatry* 2015;72:1073-1074.
- 23 Sikora M, Heffernan J, Avery ET, Mickey BJ, Zubieta JK, Pecina M: Salience Network Functional Connectivity Predicts Placebo Effects in Major Depression. *Biological psychiatry Cognitive neuroscience and neuroimaging* 2016;1:68-76.
- 24 Korb AS, Hunter AM, Cook IA, Leuchter AF: Rostral anterior cingulate cortex activity and early symptom improvement during treatment for major depressive disorder. *Psychiatry research* 2011;192:188-194.
- 25 Pizzagalli DA, Webb CA, Dillon DG, Tenke CE, Kayser J, Goer F, Fava M, McGrath P, Weissman M, Parsey R, Adams P, Trombello J, Cooper C, Deldin P, Oquendo M, McInnis M, Carmody T, Bruder G, Trivedi MH: The incremental predictive validity of rostral anterior cingulate cortex activity in relation to treatment response in depression: A randomized clinical trial. *JAMA psychiatry* 2018;In press
- 26 Trivedi MH, McGrath PJ, Fava M, Parsey RV, Kurian BT, Phillips ML, Oquendo MA, Bruder G, Pizzagalli D, Toups M, Cooper C, Adams P, Weyandt S, Morris DW, Grannemann BD, Ogden RT, Buckner R, McInnis M, Kraemer HC, Petkova E, Carmody TJ, Weissman MM: Establishing moderators and biosignatures of antidepressant response in clinical care (EMBARC): Rationale and design. *Journal of psychiatric research* 2016;78:11-23.
- 27 Hamilton M: A rating scale for depression. *Journal of neurology, neurosurgery, and psychiatry* 1960;23:56-62.
- 28 Rush AJ, Wisniewski SR, Warden D, Luther JF, Davis LL, Fava M, Nierenberg AA, Trivedi MH: Selecting among second-step antidepressant medication monotherapies: predictive value of clinical, demographic, or first-step treatment features. *Archives of general psychiatry* 2008;65:870-880.
- 29 Webb CA, Dillon DG, Pechtel P, Goer FK, Murray L, Huys QJ, Fava M, McGrath PJ, Weissman M, Parsey R, Kurian BT, Adams P, Weyandt S, Trombello JM, Grannemann B, Cooper CM, Deldin P, Tenke C, Trivedi M, Bruder G, Pizzagalli DA: Neural Correlates of Three Promising Endophenotypes of Depression: Evidence from the EMBARC Study. *Neuropsychopharmacology* 2016;41:454-463.
- 30 Tenke CE, Kayser J, Pechtel P, Webb CA, Dillon DG, Goer F, Murray L, Deldin P, Kurian BT, McGrath PJ, Parsey R, Trivedi M, Fava M, Weissman MM, McInnis M, Abraham K, J EA, Alschuler DM, Cooper C, Pizzagalli DA, Bruder GE: Demonstrating test-retest reliability of electrophysiological measures for healthy adults in a multisite study of biomarkers of antidepressant treatment response. *Psychophysiology* 2017;54:34-50.
- 31 Fischl B, Dale AM: Measuring the thickness of the human cerebral cortex from magnetic resonance images. *Proceedings of the National Academy of Sciences of the United States of America* 2000;97:11050-11055.
- 32 Fischl B, van der Kouwe A, Destrieux C, Halgren E, Segonne F, Salat DH, Busa E, Seidman LJ, Goldstein J, Kennedy D, Caviness V, Makris N, Rosen B, Dale AM: Automatically parcellating the human cerebral cortex. *Cerebral cortex (New York, NY : 1991)* 2004;14:11-22.
- 33 Zou H, Hastie T: Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 2005;67:301-320.

- 34 Iniesta R, Stahl D, McGuffin P: Machine learning, statistical learning and the future of biological research in psychiatry. *Psychological medicine* 2016;46:2455-2465.
- 35 van der Ploeg T, Austin PC, Steyerberg EW: Modern modelling techniques are data hungry: a simulation study for predicting dichotomous endpoints. *BMC medical research methodology* 2014;14:137.
- 36 Obermeyer Z, Emanuel EJ: Predicting the future—big data, machine learning, and clinical medicine. *The New England journal of medicine* 2016;375:1216.
- 37 Friedman J, Hastie T, Tibshirani R: glmnet: Lasso and elastic-net regularized generalized linear models. R package version 2009;1
- 38 Team RC: R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2014, 2014,
- 39 Raghunathan TE, Lepkowski JM, Van Hoewyk J, Solenberger P: A multivariate technique for multiply imputing missing values using a sequence of regression models. *Survey methodology* 2001;27:85-96.
- 40 Van Buuren S: Multiple imputation of discrete and continuous data by fully conditional specification. *Statistical methods in medical research* 2007;16:219-242.
- 41 Schafer JL, Graham JW: Missing data: our view of the state of the art. *Psychological methods* 2002;7:147.
- 42 Chen Q, Wang S: Variable selection for multiply-imputed data with application to dioxin exposure study. *Statistics in medicine* 2013;32:3646-3659.
- 43 Wan Y, Datta S, Conklin D, Kong M: Variable selection models based on multiple imputation with an application for predicting median effective dose and maximum effect. *Journal of statistical computation and simulation* 2015;85:1902-1916.
- 44 Liu Y, Wang Y, Feng Y, Wall MM: Variable selection and prediction with incomplete high-dimensional data. *The annals of applied statistics* 2016;10:418.
- 45 Makalic E, Schmidt DF: High-Dimensional Bayesian Regularised Regression with the BayesReg Package. *arXiv preprint arXiv:161106649* 2016
- 46 Chatfield C: Model Uncertainty, data mining and statistical inference (with discussion). *Journal of the Royal Statistical Society, series B* 1995;158:419-466.
- 47 Belloni A, Chernozhukov V: Least squares after model selection in high-dimensional sparse models. *Bernoulli* 2013;19:521-547.
- 48 White IR, Royston P, Wood AM: Multiple imputation using chained equations: Issues and guidance for practice. *Stat Med* 2011;30:377-399.
- 49 Gelman A, Carlin J, Stern H, Rubin D: Bayesian data analysis, vol. vol. 2. FL, USA: Chapman & Hall/CRC Boca Raton, 2014,
- 50 Zhou X, Reiter JP: A note on Bayesian inference after multiple imputation. *The American Statistician* 2010;64:159-163.
- 51 DeLong ER, DeLong DM, Clarke-Pearson DL: Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* 1988;44:837-845.
- 52 Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez JC, Muller M: pROC: an open-source package for R and S+ to analyze and compare ROC curves. *BMC bioinformatics* 2011;12:77.
- 53 Chang W, Cheng J, Allaire J, Xie Y, McPherson J: shiny: Web Application Framework for R. R package version 1.0. 3, 2017,
- 54 López-Ratón M, Rodríguez-Álvarez MX, Cadarso-Suárez C, Gude-Sampedro F: OptimalCutpoints: an R package for selecting optimal cutpoints in diagnostic tests. *Journal of statistical software* 2014;61:1-36.

- 55 Freeman MP, Pooley J, Flynn MJ, Baer L, Mischoulon D, Mou D, Fava M: Guarding the Gate: Remote Structured Assessments to Enhance Enrollment Precision in Depression Trials. *Journal of clinical psychopharmacology* 2017;37:176-181.
- 56 Korb AS, Hunter AM, Cook IA, Leuchter AF: Rostral anterior cingulate cortex theta current density and response to antidepressants and placebo in major depression. *Clinical Neurophysiology* 2009;120:1313-1319.
- 57 Arns M, Etkin A, Hegerl U, Williams LM, DeBattista C, Palmer DM, Fitzgerald PB, Harris A, deBeuss R, Gordon E: Frontal and rostral anterior cingulate (rACC) theta EEG in depression: Implications for treatment outcome? *European Neuropsychopharmacology* 2015;25:1190-1200.

Table 1. Baseline sociodemographic and clinical features of EMBARC trial participants randomized to placebo who received at least one dose (n=141)

Categorical variables	n	%
Gender		
Male	52	36.9
Female	89	63.1
Race		
White	98	69.5
Black	23	16.3
Other	20	14.2
Employment status		
Employed	78	55.3
Unemployed	58	41.1
<u>Continuous variables</u>	Mean	SD
Age	37.4	12.9
Age of onset	16.4	5.6
Years of education	15.2	2.7
Number of MDE*	4	9
Duration of current episode (in months)	41.5	75.8
QIDS-SR	17.8	2.7
HAMD ₁₇	18.6	4.3

*Median and Interquartile Range reported due to outlying values, EMBARC is Establishing Moderators and Biosignatures for Antidepressant Response in Clinical Care, n is number, SD is standard deviation,

MDE - major depressive episode; QIDS-SR - Quick Inventory of Depressive Symptomatology Self-Report version; HAMD₁₇ - Hamilton Rating Scale for Depression 17-item.

Table 2. Posterior median values of variables selected by elastic net that were included in the Bayesian linear regression model to predict treatment outcomes with placebo

Variable	Posterior Median	95% CI	Univariate AUC for remission	Decrement in AUC for remission upon removal
Baseline HAMD ₁₇ score	0.52	(0.186, 0.877)	0.635 (0.541-0.729)	0.092 (-0.003-0.187)
Age	0.12	(0.016, 0.222)	0.611 (0.517-0.705)	0.058 (-0.032-0.148)
Melancholic depression indicator ¹	2.06	(-0.764, 4.959)	0.388 (0.309-0.468)	0.003 (-0.077-0.083)
Anhedonia ²	0.13	(-0.123, 0.367)	0.537 (0.454-0.619)	0.017 (-0.067-0.100)
Anxious arousal ³	0.22	(-0.085, 0.537)	0.545 (0.444-0.646)	0.014 (-0.072-0.100)
Neuroticism ⁴	0.12	(-0.092, 0.336)	0.505 (0.411-0.600)	0.004 (-0.077-0.084)
Physical Abuse ⁵	0.18	(-0.185, 0.543)	0.625 (0.533-0.718)	-0.010 (-0.090-0.070)
rACC theta current density	-5.93	(-13.417, 1.166)	0.510 (0.412-0.608)	0.019 (-0.067-0.105)

HAMD₁₇ is Hamilton Rating Scale for Depression 17-item; rACC is rostral anterior cingulate

¹ Based on specifier questions on the SCID

² Anhedonic depression scale from the Mood and Anxiety Symptom Questionnaire

³ Anxiety specific scale (Anxious Arousal) from the Mood and Anxiety Symptom Questionnaire

⁴ Based on 12 neuroticism items from the NEO-FFI

⁵ Scale from the CTQ

Table 3. Comparison of model fit statistics for different thresholds for inclusion of variables in Bayesian linear regression model for prediction of outcomes with placebo

Metric	100% Retention	≥ 50% Retention	≥ 30% Retention
<i>Root Mean Squared Error</i>	6.94	6.38	6.01
<i>AUC (Remission)</i>	0.686	0.758	0.793
<i>AUC (Response)</i>	0.649	0.728	0.772

AUC – “Area Under the Curve.”

Figure 1. ROC curves for Bayesian model with a priori threshold of 50% variable retention

ROC is Receiver operator curve, AUC – “Area Under the Curve.”

Figure 2. An interactive web-based calculator to predict likelihood of placebo response

Predictors of Placebo Response

Baseline HAMD17 Score:

9

16

30

Melancholic Depression:

0

1

Age:

18

30

65

Mood and Anxiety Symptom Questionnaire - Anhedonia:

22

32

50

Mood and Anxiety Symptom Questionnaire - Anxious Arousal:

10

28

35

NEO Five Factor Inventory - Neuroticism:

10

35

48

Childhood Trauma Questionnaire - Physical Abuse:

5

10

22

Average Theta Current Density (rAcc):

-4.032

-3.3

-2.898

Update

Estimate	Value
Posterior Median HAMD17 at Acute Phase Exit:	11.59
Probability of Remission (shaded in dark red above):	0.05
Probability of Response:	0.10

Supplementary Figure 1 – Ordered Index of Percent Retention in Elastic Net Models

Weissman, Myrna (NYSPI)

From: notifications@alphaecommerce.gr
Sent: Monday, February 26, 2018 11:39 AM
To: Weissman, Myrna (NYSPI)
Subject: Electronic Transaction Confirmation for your Order (62269) at ERASMUS CONFERENCES TOURS

ATTENTION: This email came from an external source. Do not open attachments or click on links from unknown senders or unexpected emails.

Dear **Myrna Weissman**,

This e-mail verifies that your online transaction at **ERASMUS CONFERENCES TOURS** has been processed by Alpha e-Commerce Service. The details of your transaction are:

Merchant Name: ERASMUS CONFERENCES TOURS
Merchant URL: <http://www.erasmus.gr/>
TX ID: 354652031
Order Amount: 525.00 EUR
Date Started: 26/02/2018 18:39
Payment Method: amex
Wallet(s) Involved:
Wallet Status:
Order ID: 62269
Payment Reference / MyBank tx id: 289565
Number of Installments:
Number of Recurring Frequency:
Number of Recurring Sequence:

Redemption Information

Offer redeemed (Offer Code): N/A
Redemption Points: N/A
Redemption Amount: N/A

Please note that this confirmation e-mail only indicates that your transaction has been processed successfully by Alpha e-Commerce Service. It does not indicate that your order has been accepted. It is the responsibility of **ERASMUS CONFERENCES TOURS** to confirm that your order has been accepted and to deliver any goods or services you have purchased.

Please save this confirmation e-mail for future reference.

Sincerely,

Alpha e-Commerce Service
Alpha Bank