ARTICLE   OPEN

Check for updates

Multi-omics driven predictions of response to acute phase combination antidepressant therapy: a machine learning approach with cross-trial replication

Jeremiah B. Joyce [1,8], Caroline W. Grant [2,8], Duan Liu [2], Siamak MahmoudianDehkordi[3], Rima Kaddurah-Daouk[3], Michelle Skime[1], Joanna Biernacka [4], Mark A. Frye [1], Taryn Mayes [5], Thomas Carmody[6], Paul E. Croarkin [1], Liewei Wang [2], Richard Weinshilboum [2], William V. Bobo[7], Madhukar H. Trivedi [5] and Arjun P. Athreya [2]✉

Combination antidepressant pharmacotherapies are frequently used to treat major depressive disorder (MDD). However, there is no evidence that machine learning approaches combining multi-omics measures (e.g., genomics and plasma metabolomics) can achieve clinically meaningful predictions of outcomes to combination pharmacotherapy. This study examined data from 264 MDD outpatients treated with citalopram or escitalopram in the Mayo Clinic Pharmacogenomics Research Network Antidepressant Medication Pharmacogenomic Study (PGRN-AMPS) and 111 MDD outpatients treated with combination pharmacotherapies in the Combined Medications to Enhance Outcomes of Antidepressant Therapy (CO-MED) study to predict response to combination antidepressant therapies. To assess whether metabolomics with functionally validated single-nucleotide polymorphisms (SNPs) improves predictability over metabolomics alone, models were trained/tested with and without SNPs. Models trained with PGRN-AMPS' and CO-MED's escitalopram/citalopram patients predicted response in CO-MED's combination pharmacotherapy patients with accuracies of 76.6% ($p < 0.01$; AUC: 0.85) without and 77.5% ($p < 0.01$; AUC: 0.86) with SNPs. Then, models trained solely with PGRN-AMPS' escitalopram/citalopram patients predicted response in CO-MED's combination pharmacotherapy patients with accuracies of 75.3% ($p < 0.05$; AUC: 0.84) without and 77.5% ($p < 0.01$; AUC: 0.86) with SNPs, demonstrating cross-trial replication of predictions. Plasma hydroxylated sphingomyelins were prominent predictors of treatment outcomes. To explore the relationship between SNPs and hydroxylated sphingomyelins, we conducted multi-omics integration network analysis. Sphingomyelins clustered with SNPs and metabolites related to monoamine neurotransmission, suggesting a potential functional relationship. These results suggest that integrating specific metabolites and SNPs achieves accurate predictions of treatment response across classes of antidepressants. Finally, these results motivate functional investigation into how sphingomyelins might influence MDD pathophysiology, antidepressant response, or both.

Translational Psychiatry (2021)11:513 ; https://doi.org/10.1038/s41398-021-01632-z

INTRODUCTION

Major depressive disorder (MDD) is a significant public health challenge and an economic burden [1]. Selective serotonin reuptake inhibitors (SSRIs) are first-line pharmacotherapy for MDD, but more than 50% of patients fail to respond [2]. Multiple studies have investigated combination therapies as a subsequent strategy, with mixed results in terms of improved efficacy [3–6]. Given that the full effects of these medications are often not experienced for months, predicting whether a patient will respond prior to therapy or shortly thereafter would advance clinical practice and future clinical translational research.
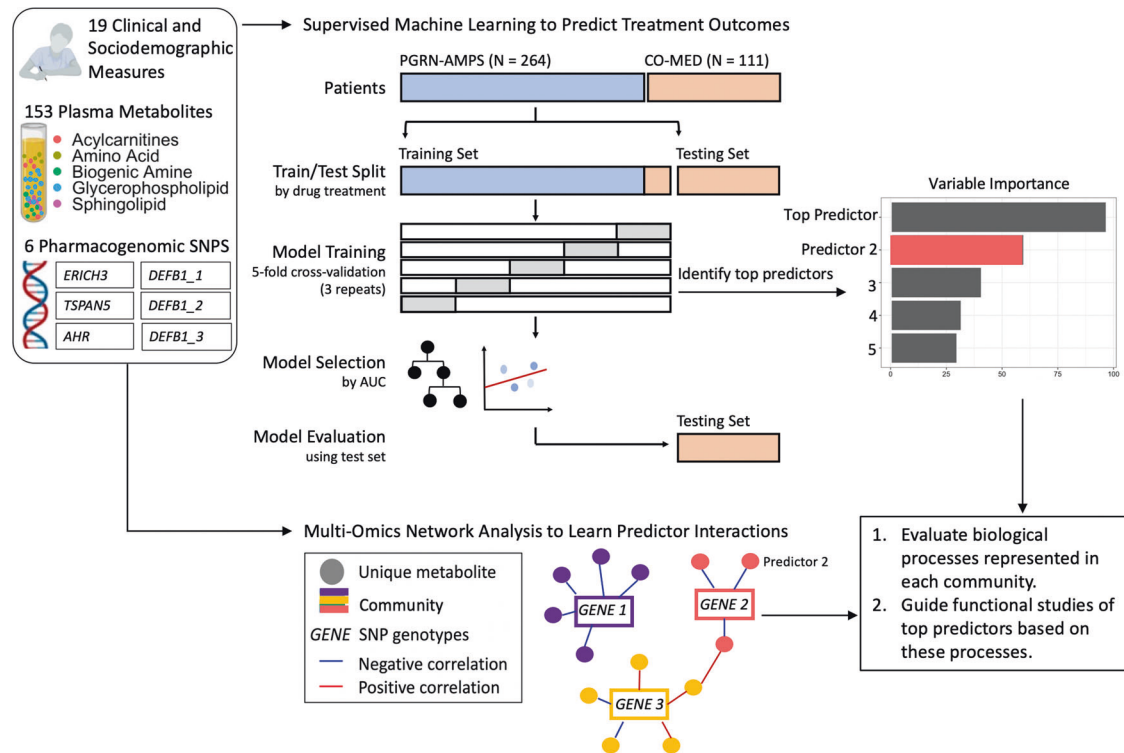
Several studies have established the predictability of antidepressant response by employing machine learning strategies [7–15]. Machine learning strategies using clinical and sociodemographic

factors predicted response to escitalopram/citalopram with accuracies of 59.6% but could not achieve statistically significant predictions across groups of patients receiving combination antidepressant therapies [12]. Two other independent strategies have demonstrated that augmenting clinical and sociodemographic factors with biological measures can improve the predictability of antidepressant treatment outcomes. First, including plasma p180 metabolomics improved predictability of changes in depression severity in a single cohort of the Combining Medications to Enhance Outcomes of Antidepressant Therapy (CO-MED [5]) subjects receiving mono or combination antidepressant therapies [13]. Second, including six functionally validated genomic biomarkers (i.e., single-nucleotide polymorphisms (SNPs) with mechanisms related to MDD severity, or citalopram or escitalopram

Fig. 1 Conceptual overview of model development and evaluation. PGRN-AMPS (Pharmacogenomic Research Network Antidepressant Medication Pharmacogenomic Study) and CO-MED (Combining Medications to Enhance Outcomes of Antidepressant Therapy) participants were partitioned into training/testing groups based upon treatment allocation.

response) in the Pharmacogenomic Research Network Antidepressant Medication Pharmacogenomic Study (PGRN-AMPS [16]) improved predictive accuracies of treatment response to >69% in patients treated with either citalopram or escitalopram [7]. These prior studies using either plasma metabolomics or genomics were limited, because they did not demonstrate cross-trial replication of predictions in patients receiving combination pharmacotherapy.

The present study examined data from the PGRN-AMPS and CO-MED studies with a machine learning and multi-omics strategy (Fig. 1) to address key questions. Can machine learning strategies combining plasma metabolomic and genomic measures from MDD patients receiving antidepressant monotherapy (citalopram or escitalopram) achieve statistically significant predictions of response to combination pharmacotherapy? If combining these multi-omics measures improves predictability of response to multiple classes of antidepressants, can multi-omics integration networks elucidate biologically meaningful relationships between metabolomic predictors of antidepressant response and functionally validated genomic biomarkers? This present study hypothesized that augmenting clinical measures (e.g., symptom severity scores) with multiple biological measures (e.g., metabolomics and genomics) might improve the predictability of response to combination antidepressant therapies.

MATERIALS AND METHODS
Data sources
MDD outpatients with complete clinical assessments, baseline metabolomics, and genomics from PGRN-AMPS and CO-MED were included in this work (Supplementary Fig. 1 and Supplementary Table 1). Patients were split into training and testing cohorts. The training cohort consisted of patients taking citalopram, escitalopram, or escitalopram plus placebo, whereas the testing cohort comprised patients receiving combination pharmacotherapies (Table 1). Briefly, the PGRN-AMPS 8-week SSRI clinical trial (NCT00613470) enrolled 529 MDD patients who scored ≥14 on the 17-item Hamilton Depression Rating Scale (HAMD-17). Patients received either

escitalopram (10 mg/day) or citalopram (20 mg/day). The CO-MED 7-month clinical trial (NCT00590863) enrolled 665 MDD patients who scored ≥16 on the HAMD-17 and who met criteria for either recurrent or chronic (current episode ≥2 years) depression. Patients were randomized to one of the following treatment regimens: (1) escitalopram (up to 20 mg/day) plus placebo, (2) bupropion (up to 400 mg/day) plus escitalopram (up to 20 mg/day), or (3) extended-release venlafaxine (up to 300 mg/day) plus mirtazapine (up to 45 mg/day). These studies were conducted in accordance with the approval of respective Institutional Review Boards to include informed consent. Both studies have been characterized in prior publications [5, 16].

Measures and outcomes
Race and ethnicity data were self-reported in both studies (Table 1 and Supplementary Table 1). History of prior suicide attempts was collected as a binary yes/no question; this variable represents the number of unique patients who have had at least one prior suicide attempt (Table 1 and Supplementary Table 1). Depression symptom severity and treatment outcomes were assessed using the Clinician-Rated Quick-Inventory of Depressive Symptomatology (QIDS-C). Response to therapy at 8 weeks was defined as ≥50% reduction in QIDS-C total score from baseline [17]. Remission (reported to characterize the samples in Table 1 and Supplementary Table 1) was defined as achieving a score of ≤5 on the QIDS-C [18].

Model definitions
We developed two sets of prediction models. The first model set ("Metabolomic Models") included clinical, sociodemographic, and metabolomic features common to both the PGRN-AMPS and CO-MED studies. All sociodemographic and metabolomic features common to both data sets were included (Table 1 and Supplementary Table 2). Clinical features included baseline and week 4 change of individual QIDS-C items belonging to a "core set" previously defined to extract homogeneous patterns of citalopram/escitalopram response from diverse response trajectories [19] (Supplementary Table 2). Independent studies also demonstrate the predictive utility of QIDS-C baseline totals and early change in total QIDS-C [20]. Therefore, QIDS-C baseline totals and QIDS-C week 4 percent change were also included as clinical predictors

Table 1. Clinical and sociodemographic features of training and testing cohorts.

Training and testing cohort				
	$N = 375$ Model Set 1 (Metabolomic)		$N = 348$ Model Set 2 (Multi-omics)	
	Training Set $N = 298$	Testing Set $N = 77$	Training Set $N = 277$	Testing Set $N = 71$
PGRN-AMPS patients (N)	264	0	245	0
CO-MED patients (N)	34	77	32	71
Sex [% female]	66.1%	71.4%	65.3%	70.4%
Age [mean (SD)]	40.6 (13.3)*	43.3 (11.5)*	41.0 (13.3)	43.4 (11.6)
Years of education [mean (SD)]	14.7 (2.5)*	13.9 (2.4)*	14.8 (2.5)*	13.9 (2.4)*
Race [% White]	93.6%*	77.9%*	96.0%*	77.5%*
Race [% Black or African American]	3.4%*	16.9%*	2.2%*	16.9%*
Race [% Other]	3.0%	5.2%	1.8%	5.6%
Ethnicity [% Hispanic]	2.0%†*	20.8%*	1.4%†*	21.1%*
Depression onset < age 18 years [%]	43.6%	39.0%	41.5%	39.4%
Prior suicide attempts [N (%)]	46 (15.4%)	7 (9.1%)	40 (16.3%)	6 (8.4%)
Antidepressants (N)	Citalopram (112) Escitalopram (152) Escitalopram + Placebo (34)	Venlafaxine + Mirtazapine (42) Escitalopram + Bupropion (35)	Citalopram (99) Escitalopram (146) Escitalopram + Placebo (32)	Venlafaxine + Mirtazapine (38) Escitalopram + Bupropion (33)
QIDS-C at baseline [mean (SD)]	15.1 (3.4)	15.5 (3.8)	15.1 (3.3)	15.3 (3.7)
QIDS-C response at week 4	47.0%	40.2%	47.7%	42.2%
QIDS-C remission at week 4	26.2%	26.0%	26.7%	28.2%
QIDS-C response at week 8	67.8%	62.3%	69.0%	63.3%
QIDS-C remission at week 8	47.0%	42.9%	48.4%	43.7%

*Significantly different ($p < 0.05$) between training and testing sets according to Mann-Whitney U or chi-square tests. †Ethnicity characterization in PGRN-AMPS is based off data from 205 out of 264 patients, per availability of data.

(Supplementary Table 2). The second model set ("Multi-Omics Models") augmented model set 1 with the inclusion of six previously functionally validated SNPs (Supplementary Table 2).

Metabolomics
Plasma metabolites in both the PGRN-AMPS and the CO-MED cohorts were measured by targeted metabolomics with the AbsoluteIDQ p180 assay platform (BIOCRATES Life Science AG, Innsbruck, Austria) [21] with quality-control and metabolomic profiling as previously published [13, 22, 23]. One hundred and fifty-three metabolites that met quality-control criteria in both the PGRN-AMPS and CO-MED data sets were included. The p180 kit includes all quality control (QC) samples, and calibration and internal standards; therefore, quantifications can be directly compared across studies. This assay detects metabolites from five analyte groups as follows: acylcarnitines, amino acids, biogenic amines, glycerophospholipids, and sphingolipids assayed by use of triple quadrupole tandem mass spectrometry operated in Multiple Reaction Monitoring mode.

Genomics
Six functionally validated pharmacogenomic SNP biomarkers in or near the *TSPAN5*, *ERICH3*, *DEFB1*, and *AHR* genes, and related to MDD pathophysiology or citalopram/escitalopram response [24–27], were included in the multi-omics models. These six SNPs were selected, because their mechanisms have been extensively characterized with multiple experimental models, including human induced pluripotent stem cells (iPSC)-derived astrocytes and neurons [25, 26]. These SNPs were initially pursued as top signals in metabolomics-informed-genomics studies in PGRN-AMPS depressed patients [24, 27]. The metabolomics-informed-genomics strategy moves beyond heterogeneous clinical disease and outcome phenotypes to help identify genomic and metabolomic variation, which might

contribute to individual differences in response to pharmacological agents [28]. In those studies, serotonin and kynurenine concentrations were most significantly associated with SSRI outcomes and baseline depression severity, respectively [24, 27]. Top SNPs that were functionally pursued from the serotonin genome-wide association study (GWAS) were in or near *TSPAN5* (rs10516436) and *ERICH3* (rs696692), and top SNPs from the kynurenine GWAS were in or near *DEFB1* (rs5743467, rs2741130, and rs2702877) and *AHR* (rs17137566). These SNPs were then demonstrated to be predictive of escitalopram/citalopram treatment outcomes in multiple large MDD cohorts [7].

These six SNPs are not exhaustive of the list of SNPs, which may be predictive of antidepressant outcomes. For example, the International Study to Predict Optimized Treatment in Depression found a significant main effect of the rs10245483 SNP, which alters P-glycoprotein expression in lymphoblast cells [29], on predicting remission [30]. To test whether the predictability of outcomes using PGRN-AMPS-derived metabolomic-informed-genomic SNPs may be augmented with rs10245483, we conducted additional experiments incorporating this SNP (Supplementary Table 3).

PGRN-AMPS genotyping was done using Illumina human 610-Quad BeadChips (Illumina, San Diego, CA, USA), with imputation and QC as previously reported [27]. CO-MED genotyping was done using Illumina Quad, Human Omni 2.5 bead chip, as previously published [31]. One of the six SNPs was genotyped in the CO-MED sample with a LooRsq > 99% and the remaining five SNPs were imputed using the Michigan Imputation Server with an imputation $R^2 > 97.5\%$ and a call rate > 99%.

Machine learning strategy
Data preprocessing. Features with ≥10% missingness and individuals missing ≥20% of features were excluded (Supplementary Fig. 1). Serotonin
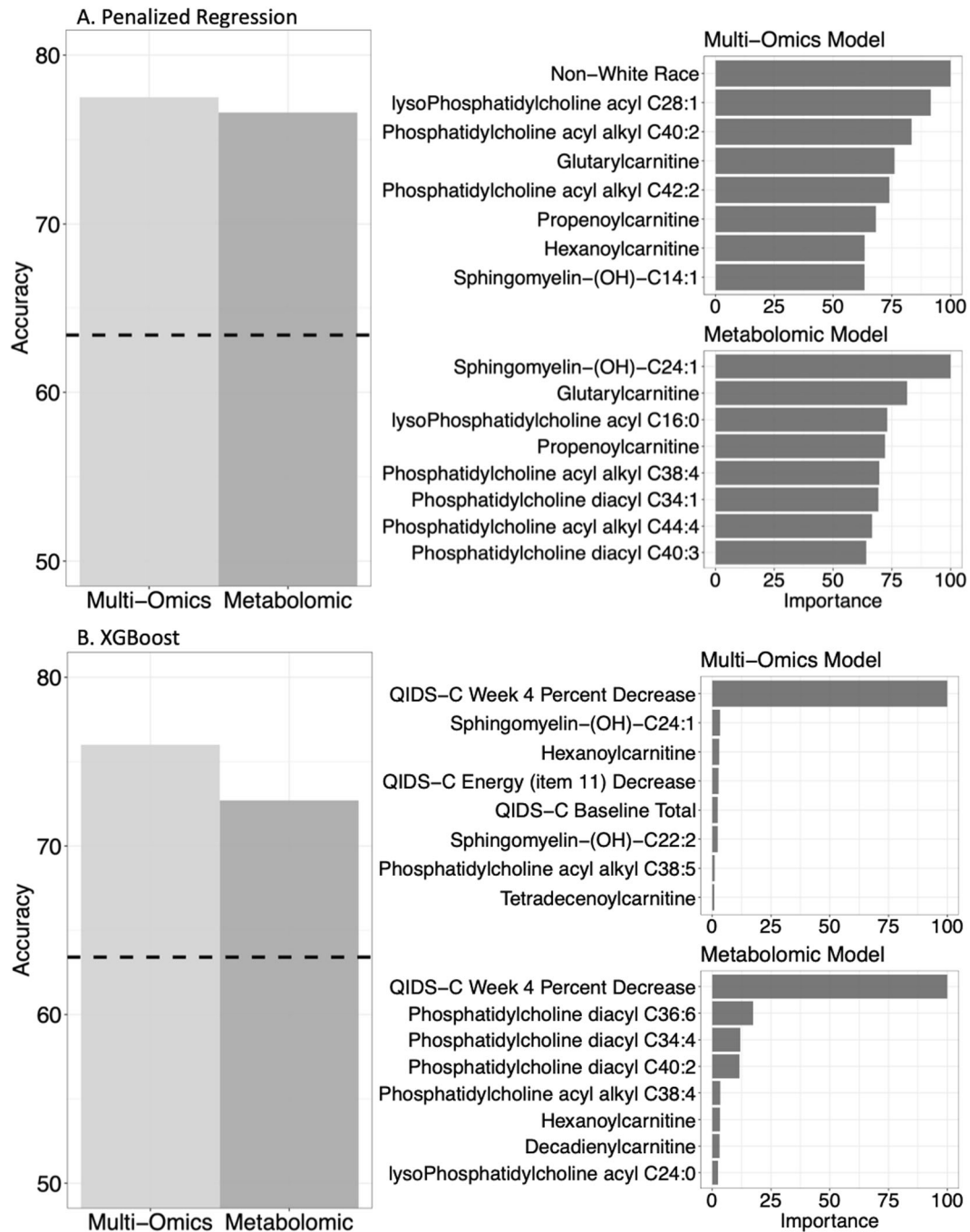
Fig. 2 Comparison of test-set accuracies for metabolomic and multi-omics models, with variable importance. Dashed line: null information rate (NIR). The NIR represents the response rate of 63.4% at 8 weeks of therapy. This serves as a benchmark to assess the significance of prediction accuracy.

and body mass index (BMI) were among the features that were excluded for having ≥10% missingness. This yielded a complete data set, except for one missing taurine value, which was imputed using K-nearest neighbor's imputation. Metabolites were transformed by the Yeo-Johnson transformation then centered at zero and scaled to unit variance. Nominal variables were converted to binary numerical values. Additive allele effects were assumed for SNPs.

Prediction model development. Models predicting antidepressant response at 8 weeks of therapy were trained on a treatment-homogenous set of patients taking the following SSRI monotherapies: citalopram, escitalopram, or escitalopram plus placebo (Fig. 1 and Table 1). In order to minimize the chance of overfit, estimate prediction

performance, and tune model hyperparameters, we used fivefold cross-validation with three repeats. For robustness, we also conducted experiments utilizing threefold and tenfold cross-validation with three repeats (Supplementary Table 4). Both linear and nonlinear algorithms were tested. As a linear penalized regression approach using clinical and metabolomic features successfully yielded predictive insights into change in QIDS Self-Report score in the CO-MED data set [13], we tested penalized regression performance in our combined cohorts. We then tested extreme gradient-boosted decision tree-based ensembles (XGBoost) as nonparametric models. Nonparametric models identify possible nonlinear relationships among predictors (e.g., metabolites and age [32]), while predicting treatment outcomes. These models both have the advantage of being tolerant to multi-collinear data [33]. For all models, upsampling was used

Table 2. Metrics of prediction performance.

Prediction model metrics				
2A: Models trained with PGRN-AMPS escitalopram, PGRN-AMPS citalopram, and CO-MED escitalopram + placebo patients.				
	Model Set 1 (Metabolomic)		**Model Set 2 (Multi-omics)**	
	XGBoost	**Penalized regression**	**XGBoost**	**Penalized regression**
Testing-set metrics				
AUC	0.76	0.85	0.83	0.86
Accuracy	0.727	0.766	0.761	0.775
NIR	0.62	0.62	0.63	0.63
p-Value	0.053	0.0045	0.017	0.0067
Sensitivity	0.75	0.69	0.69	0.71
Specificity	0.69	0.90	0.88	0.88
Training-set metrics (in cross-validation)				
AUC	0.69	0.69	0.68	0.72
Accuracy	0.64	0.66	0.67	0.65
NIR	0.68	0.68	0.69	0.69
2B: Models trained with PGRN-AMPS escitalopram and PGRN-AMPS citalopram patients.				
Testing-set metrics				
AUC	0.75	0.84	0.74	0.86
Accuracy	0.753	0.753	0.732	0.775
p-Value	0.026	0.026	0.085	0.0067
Sensitivity	0.65	0.73	0.80	0.71
Specificity	0.93	0.79	0.62	0.88
Training-set metrics (in cross-validation)				
AUC	0.68	0.68	0.72	0.72
Accuracy	0.64	0.65	0.67	0.68
NIR	0.69	0.69	0.70	0.70

For the penalized regression models, cartesian grid search was used to tune penalty and mixture hyperparameters, with 20 evenly spaced penalty values ranging from $1e-6$ to 10 and mixture values of 0, 0.05, 0.2, 0.4, 0.6, 0.8, and 1. For the XGBoost models, we tuned the number of trees, tree depth, minimum number of data points to split a node, learning rate, loss function reduction, and sample size using Bayesian optimization and a stopping criterion of no improvement over 30 iterations.

to correct for the class imbalance (67% response, 33% non-response). Upsampling injects minority class data points into the training set, equalizing the counts of both classes and preventing model inclination towards the majority class [34]. Variable importance plots were generated to show top predictors of antidepressant response (Fig. 2). We report the expanse of the grid search and tuned hyperparameters for both approaches in Supplementary Table 5.

The best-trained models based on the area under the receiver operating characteristic curve (AUC) during cross-validation were tested on an independent set of cases receiving combination pharmacotherapy (Table 2A). AUC is a widely used classification performance metric that ranges from 0.5 (random guessing) to 1.0 (perfect prediction). Model performance on the test set was evaluated according to accuracy, sensitivity, specificity, and AUC. Preprocessing and machine learning analyses were performed using the tidymodels package version 0.1.2 [35] implemented in R 4.0.3 [36] using RStudio version 1.3 [37].

Cross-trial replication experiment. We conducted an additional analysis to determine whether models trained with citalopram or escitalopram-treated patients (PGRN-AMPS monotherapy-treated patients) could predict outcomes in the CO-MED combination therapy-treated patients (Table 2B). This additional analysis enhances the validity of the cross-trial prediction

performance. These additional models were trained and tested according to the previously outlined approaches.

Multi-omics integration. The goal of the multi-omics integration analysis was to understand the interaction of predictive plasma metabolomic measures with functionally validated SNPs (Fig. 1). The multi-omics integration network analysis tool xMWAS [38] took all 153 metabolomic, 6 genomic, and clinical response features from all 348 participants as input (Supplementary Table 1). Network analysis was performed using sparse partial least squares (sPLS) regression analysis, a multivariate approach for data integration, with $|r| > 0.1$ and $p < 0.05$. sPLS simultaneously performs variable selection and feature integration, and was originally designed for scenarios with highly correlated variables [39]. As output, the community detection method in xMWAS generates a network of communities (sub-networks) comprising nodes (highly correlated SNPs and metabolites), which are tightly associated within the community but sparsely associated to nodes of other communities [40]. We hypothesized that we would find interactions between kynurenine and the *DEFB1* and *AHR* SNPs, as these SNPs were originally pursued as top signals from a GWAS for kynurenine [24]. Therefore, we performed a second multi-omics integration analysis excluding kynurenine, to assess changes in community membership in the absence of this known strong correlation. Multi-omics integration uncovers associations (via community membership) between top metabolite predictors and functionally characterized SNPs.

RESULTS
Prediction of combination therapy-treated patients with citalopram-, escitalopram-, and escitalopram + placebo-treated patients

Model training. In the citalopram/escitalopram/escitalopram + placebo monotherapy training set (which used repeated cross-validation to train prediction models), clinical, sociodemographic, and metabolomic features of the "metabolomics models" predicted response to treatment at 8 weeks with an AUC of 0.69 in both machine learning algorithms. When metabolomic and clinical predictor variables were further augmented with six pharmacogenomic SNPs ("multi-omics models"), the training-set prediction AUCs were 0.68 (XGBoost) and 0.72 (penalized regression)—representing an improvement of 0.03 for penalized regression (Table 2A).

Model testing. Using the "metabolomics models" feature set, the best-trained classifiers predicted response to combination anti-depressant therapies at 8 weeks with accuracies of 76.6% ($p < 0.005$; AUC: 0.85) and 72.7% ($p = 0.053$; AUC: 0.76) for penalized regression and XGBoost, respectively. Using the "multi-omics models" feature set, accuracies improved to 77.5% ($p < 0.01$; AUC: 0.86) and 76.1% ($p = 0.017$; AUC: 0.83) (Table 2A).

Prediction of combination therapy-treated patients with citalopram and escitalopram-treated patients

Model training. In the citalopram/escitalopram monotherapy training set (comprising only PGRN-AMPs patients and which used repeated cross-validation to train prediction models), clinical, sociodemographic, and metabolomic features of the "metabolomics models" predicted response to treatment at 8 weeks with an AUC of 0.68 in both machine learning algorithms. When metabolomic and clinical predictor variables were augmented with six pharmacogenomic SNPs ("multi-omics models"), the training-set prediction AUCs increased to 0.72 for both XGBoost and penalized regression—representing an improvement of 0.04 for both algorithms (Table 2B).

Model testing. Using the "metabolomics models" feature set, the best-trained classifiers from PGRN-AMPS subjects receiving citalopram or escitalopram monotherapy predicted response to combination antidepressant therapies at 8 weeks with accuracies of 75.3% ($p = 0.026$; AUC: 0.84) and 75.3% ($p = 0.026$; AUC: 0.75), and for penalized regression and XGBoost, respectively. Using the
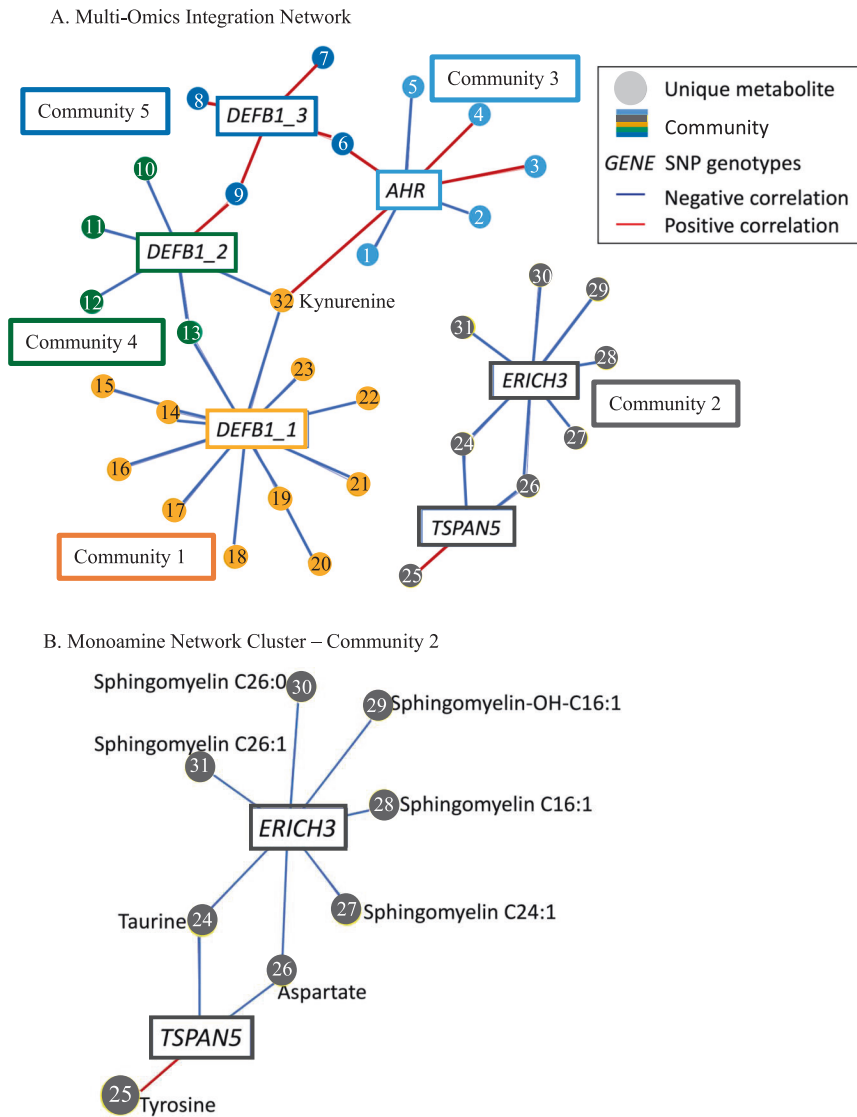
Fig. 3 Multi-omics integration network analysis. A Multi-omics integration network. Each metabolite is labeled with a number. Metabolite names corresponding to these numbers can be found in Supplementary Table 6. Correlation values between metabolites and SNPs can be found in Supplementary Table 7. **B** Labeled and enlarged community-2 containing sphingomyelins. Represented genotypes are: *TSPAN5* (rs10516436), *ERICH3* (rs696692), *AHR* (rs17137566), *DEFB1_1* (rs5743467), *DEFB1_2* (rs2741130), *DEFB1_3* (rs2702877).

"multi-omics models" feature set from these subjects, accuracies changed to 77.5% ($p = 0.0067$; AUC: 0.86) and 73.2% ($p = 0.085$; AUC: 0.74) (Table 2B)—representing an improvement of 2.2% for penalized regression.

Top predictor variables.   The top predictors were calculated using the models trained on escitalopram, citalopram, or escitalopram + placebo-treated patients, to maximize the number of patients in the variable importance calculations. The top predictors varied by algorithm and feature set, but hydroxylated sphingomyelins, glycerophospholipids, clinical/sociodemographic features, and acylcarnitines, and were all represented (Fig. 2). Although SNPs were not among top predictors in both approaches, augmenting the models with SNPs increased the predictive importance of clinical/sociodemographic features (ethnicity, baseline depression, and change in energy). In addition, in the XGBoost models, change in QIDS-C score was a top predictor of treatment outcomes and inclusion of SNPs increased the predictive importance of hydroxylated sphingomyelins.

Multi-omics network integration.   Integrative network analysis was used to establish relationships between top metabolite predictors (i.e., hydroxylated sphingomyelins) and SNPs whose MDD-related mechanisms have been functionally characterized. The community detection algorithm in this analysis identifies communities as groups of nodes (metabolites and SNPs), which are tightly associated with nodes inside but loosely associated with nodes outside the community [40].

The input to the network analysis included all 153 metabolites, 6 SNPs, and treatment response labels. The analysis identified five communities of statistically related ($|r| > 0.1$ and $p < 0.05$) entities, comprising 32 metabolites and 6 SNPs (Fig. 3). All communities contained both SNPs and metabolites. SNPs in or near the *DEFB1* and *AHR* genes clustered into communities 1, 3, 4, and 5, whereas SNPs in or near *TSPAN5* and *ERICH3* clustered with sphingomyelins and amino acids in community-2 (Fig. 3B). The *ERICH3* SNP in community-2 correlated negatively with various sphingomyelins and a hydroxylated sphingomyelin. Both the *TSPAN5* and *ERICH3* SNPs in community-2 correlated negatively with taurine and

aspartate, and the *TSPAN5* SNP correlated positively with tyrosine. Sphingomyelins were absent from communities 1, 3, 4, and 5, where glycerophospholipids, amino acids, and acylcarnitines were represented (Fig. 3 and Supplementary Table 6). Kynurenine clustered with the *DEFB1* and *AHR* SNPs, as expected, as these SNPs were originally pursued from a GWAS for kynurenine (Fig. 3A). When the analysis was repeated without kynurenine, community membership remained consistent (Supplementary Fig. 2). Pearson's correlation coefficients between SNPs and metabolites can be found in Supplementary Table 7.

DISCUSSION
The present findings integrated clinical and multiple biological measures to achieve 77.5% accuracy ($p = 0.0067$; AUC = 0.86) in predicting cross-trial outcomes to combination antidepressant therapy. The multi-omics driven augmented prediction performance marks an improvement over prior work in that (1) baseline clinical and sociodemographic measures alone previously achieved 51.4% accuracy in the CO-MED cohort [12] and (2) there is no prior evidence that omics-based models can achieve cross-trial replication in predicting response to combination antidepressant pharmacotherapies. Up to this point, it had been demonstrated that augmenting clinical and sociodemographic measures with either genomics [7] or metabolomics [13] alone improves predictability of treatment outcomes—this work demonstrates for the first time that incorporating SNPs and metabolites together can achieve cross-trial replication in predicting response to combination antidepressant therapies. Cross-trial replication was achieved despite characteristic differences in the PGRN-AMPS training and CO-MED testing cohorts, including race, ethnicity, and antidepressant therapy (Table 1). PGRN-AMPS patients represent a population based in the Upper Midwest of the United States, whereas CO-MED patients represent a more diverse population recruited from 15 sites throughout the United States. These findings encourage future antidepressant studies to assay a wider range of biological measures (e.g., proteomics, transcriptomics, and epigenomics), which might not only continue to improve predictability but might also advance our molecular understanding of MDD pathophysiology, antidepressant response, or both.

The six SNPs included in this work are not meant to be exhaustive or definitive of the entire set of SNPs, which are predictive of antidepressant therapy outcomes. Instead, they were selected, because their MDD-related mechanisms have been characterized extensively [24–27] and they previously increased the predictability of SSRI treatment outcomes compared with models utilizing clinical or sociodemographic factors alone [7]. Incorporating an additional P-glycoprotein SNP (rs10245483) yielded no further improvements in predictability (Supplementary Table 3). This may be related to drug-specific effects of rs10245483 [30], although future experiments would be needed to clarify this. Future experiments should aim to identify additional SNPs, which may synergize with the six metabolomics-informed-genomics SNPs included in this work to improve predictability in the future.

The multi-omics network analysis aimed to elucidate how the combination of these SNPs and metabolites (the "multi-omics models") continued to improve the predictability of treatment outcomes. The network analysis clustered *TSPAN5* and *ERICH3* in community-2, separate from communities that include *DEFB1* and *AHR* (Fig. 3). Human iPSC-derived astrocytes and neurons, and rodent models have demonstrated that *DEFB1* and *AHR* regulate central and peripheral inflammation via modulation of the kynurenine pathway, whereas *TSPAN5* and *ERICH3* impact monoamine neurotransmission via multiple mechanisms (reviewed in ref. [41]) [24–27]. These monoamine-related SNPs cluster with the monoamine-related metabolites taurine and tyrosine (Fig. 3). Tyrosine is a precursor to catecholamine monoamines, whereas

taurine is a natural analog to the alcohol-use-disorder treatment acamprosate, a drug which impacts monoamine-pathway enzyme concentrations [26]. Incorporating both SNPs and metabolites ("multi-omics" models) likely improved predictive accuracy over single-omics models through complimentary biological measurements of processes perturbed in MDD, including monoamine neurotransmission [38] and inflammation [42, 43].

This current analytical workflow ranks biomarkers by their contributions to improved predictions and elucidates opportunities for laboratory-based studies accordingly. Based on these results, laboratory-based experiments might explore individual variation in hydroxylated sphingomyelins, as they are both top predictors of treatment response and are also associated with the monoamine neurotransmission community-2 (Fig. 3B). The importance of hydroxylated sphingomyelins in predicting MDD treatment outcomes is also highlighted by the following evidence: hydroxylated sphingomyelins were represented in a signature of ketamine treatment response [44] and the baseline ratio of hydroxylated to non-hydroxylated sphingomyelins, as well as a larger change in this ratio by 12 weeks of therapy, predicted greater reduction in depressive symptoms with escitalopram or combined medication [13]. Our current analytical approach extends these findings to associate hydroxylated sphingomyelins with markers of monoamine neurotransmission, providing a foundation for follow-up mechanistic studies.

Sphingomyelins are dominant lipids in the outer leaflet of the plasma membrane of most cells and are especially prevalent in the brain [45, 46]. Their hydrogen-bonding capabilities confer unique functional properties distinct from glycerophospholipids and other membrane lipids [45]. Sphingomyelins help maintain the structural integrity of lipid rafts [47], which are highly ordered membrane domains involved in cell signaling [48, 49]. Hydroxylated sphingomyelins are a more polar species that promote fluidization of lipid rafts [47]. Sphingomyelin species, including hydroxylated and long-chain sphingomyelins, affect a multitude of processes such as the regulation of endocytosis [45], ligand binding to the serotonin$_{1A}$ receptor [50, 51], and receptor-mediated ligand uptake [45]. The sphingomyelin-ceramide system has also been demonstrated to play a role in antidepressant response through mechanisms related to autophagy [52].

Given the known roles of sphingomyelins at the plasma membrane and the correlations that we observed between sphingomyelins and the *TSPAN5* and *ERICH3* SNPs in our network analysis, we hypothesize a functional, as well as a statistical relationship between hydroxylated sphingomyelins (or their isomeric/isobaric relatives) and *TSPAN5/ERICH3* mechanisms of pathophysiology in MDD. Both TSPAN5 and ERICH3 physically interact with neurotransmitter vesicle-associated proteins; therefore, previous work suggests that neurotransmitter vesicle biogenesis and/or function is a mechanism by which both genes influence monoamine availability [25, 26]. Vesicles store monoamine neurotransmitters until exocytosis at the plasma membrane [53]. The plasma membrane itself plays a pivotal role in the neurotransmitter life cycle, influencing synaptic vesicle synthesis, storage, release, transport, and degradation [54]. As prevalent polar lipids in the outer leaflet of the plasma membrane [47], sphingomyelins (hydroxylated or not) may affect any of these processes. Gaining a clearer understanding of these processes in future studies may elucidate the biological processes that drive hydroxylated sphingomyelins to be top predictors of antidepressant treatment outcomes among the metabolites that were assayed.

Laboratory investigation of predictive metabolites (e.g., hydroxylated sphingomyelins) could begin with the repeatedly validated metabolomics-informed-genomics approach [28]. The success of this approach in discovering novel MDD or antidepressant response biology is exemplified by our identification and functional investigation of the *TSPAN5*, *ERICH3*, *DEFB1*, and *AHR*

genes. SNPs in or near these genes associated with concentrations of serotonin or kynurenine, which were in turn associated with citalopram/escitalopram treatment outcomes or baseline depression in a PGRN-AMPS cohort [24, 25, 27]. The metabolomics-informed-genomics studies, which identified those genes utilized a liquid chromatography electrochemical coulometric array (LC-ECA) metabolomics platform [27]. The LC-ECA platform quantified 31 known plasma metabolites in PGRN-AMPS patients, primarily in the tryptophan, tyrosine, purine, and tocopherol pathways [27]. In contrast, the 153 metabolites included in this work were quantified utilizing the p180 platform, which detects up to 40 acylcarnitines, 42 amino acids and biogenic amines, 90 glycerophospholipids, 15 sphingolipids, and sum of hexoses. The expanse of metabolites assayed by the p180 platform metabolites have previously improved predictability of change in QIDS self-reported score [13]. By combining SNPs of established predictability arising from LC-ECA detected metabolites with the expanse of metabolites available from the p180 platform, this work examines for the first time the joint predictability of these independently predictive biomarkers.

Compared with the electrochemical coulometric array platform, which quantified 31 plasma metabolites in previous work [27], the p180 platform detected 153 metabolites after quality control in this work. Therefore, we required a novel strategy to identify metabolites for follow-up mechanistic investigation. Linear associations between the 31 LC-ECA assayed metabolites and outcomes were previously successful in identifying serotonin and kynurenine for mechanistic investigation, linear associations between the 153 p180-assayed metabolites and outcomes yield many more candidate metabolites than can be immediately pursued in laboratory-based studies. The analytical strategy outlined here provides an alternative approach suited for expansive assays to identify metabolites for functional studies.

This study also highlights the clinical importance of developing better means to predict antidepressant treatment outcomes (i.e., incorporating biological markers for improving predictability and achieving cross-trial replications) in depressed patients [8, 55]. Monotherapy with antidepressants is the recommended first-line treatment approach. However, poor or incomplete responses to monotherapy occur frequently, even at optimized doses [56]. Antidepressant combinations are important strategies to consider in such circumstances [57]. Not surprisingly, antidepressant combinations, including the combinations studied in this work, are frequently encountered in clinical practice [58–62]. The field has begun to investigate clinically accessible parameters (such as BMI and markers of systemic inflammation, such as C-reactive protein), which have shown promise as predictive biomarkers for antidepressive effects with serotonergic and non-serotonergic antidepressants [63–65]. However, there are still no widely implemented evidence-based methods for predicting treatment outcomes with antidepressant combinations, either at the start of treatment or based on early (e.g., 4 weeks) response to therapy. Future studies that focus on integrating biological and symptom-based factors to predict clinical outcomes with other commonly used antidepressant combinations are needed. Supporting the validity of using a multi-omics machine learning process for predicting clinical outcomes is the ability of the models to generalize in diverse populations.

The top predictors varied by algorithm and feature set, as can be expected given the different mathematical underpinnings of penalized regression and boosting algorithms. In cross-validation, elastic net regression was demonstrated to be the optimal penalized regression approach for both the metabolomics and multi-omics data sets. Elastic net regression simultaneously performs continuous shrinkage and variable selection and is particularly useful for selecting predictive groups of correlated variables [66]. Elastic net regression determined sphingomyelins, phosphatidylcholines, and carnitines to be top metabolite

predictors (Fig. 2A). In contrast to elastic net regression, boosting algorithms are tree-based approaches. They first split the decision trees on the feature that best classifies the training samples, then they iteratively reduce erroneous classifications by exploring the joint predictive capabilities of other features [67]. Given that we know that early change in MDD severity associates with eventual treatment outcomes [68], it is likely that XGBoost begins classification based on early MDD severity change, thereby making this a top predictor in its models (Fig. 2B), before considering the contributions of other predictors in the final adjudication of whether a patient responds or not. Neither penalized regression nor boosting provides a definitively superior set of top predictors—jointly assessing the top predictors of both approaches demonstrates that hydroxylated sphingomyelins are consistently amongst top predictors. Therefore, investigating hydroxylated sphingomyelins in future work mechanistic work may lead to improved insights into the biology of antidepressant response.

In addition, non-White race was a top predictor in the multi-omics penalized regression model, but not in the metabolomics penalized regression model (Fig. 2A). This may have arisen from correlations between SNPs and race, i.e., differences in SNP minor allele frequencies among populations (Supplementary Table 8). These SNPs were initially discovered in the context of MDD in the PGRN-AMPS cohort, a primarily white population, which may explain their interaction with non-white race in the penalized regression models. Future work should aim to establish the predictability of these SNPs, along with any SNPs arising from mechanistic investigation of hydroxylated sphingomyelins, across populations.

Performance metrics were lower in the training data set compared to the testing data set (Table 2 and Supplementary Tables 3 and 4). This may have arisen from (a) higher inter-patient variability in measures in the training data set and/or (b) smaller sample sizes in testing data sets that may comprise subjects with lesser inter-patient variability in disease severity or drug response. Although the current results are encouraging in that multi-omics models developed with monotherapy patients established predictability of outcomes to combination antidepressant therapy, future work with additional biomarkers (e.g., neuroimaging, additional genomic SNPs, and proteomics) may achieve comparable performances in both larger and smaller sized studies.

This work has limitations. Lipids assayed by current mass spectrometry technology may actually reflect sum signals of all isomeric/isobaric compounds having the same parent and daughter ions, and that phenomenon may occur with the hydroxylated sphingomyelins [69]. Therefore, future approaches should validate the identified variables with other assays. Although eventual response may be predicted as early as 2 weeks post-treatment initiation [70], this was not feasible in this work due to lack of data at week 2. Minor allele frequencies for the included SNPs vary by population (Supplementary Table 8), so larger samples with individuals of diverse ancestry will be needed to validate prediction performance across populations. Although fasting status may not significantly affect laboratory variability for most metabolites [71], we cannot exclude the possibility of bias introduced by lack of fasting samples [72, 73]. This work's analysis includes only the clinical/sociodemographic features common to both cohorts, so relevant features including clinical comorbidities, socioeconomic status, and number of depressive episodes were not examined. In addition, high missingness (>35%) in BMI and serotonin concentration data, stemming from the absence of collection and quantifications below the lower limit of detection or lowest calibration standard, respectively, precluded their incorporation in these analyses. Both studies excluded patients with a history of psychotic symptoms, so these findings may not generalize to patients with psychotic depression. Finally, we did not have blood drug levels for all patients, but this limitation is

less concerning given that drug levels did not associate with clinical outcomes in previous studies [7, 10].

In summary, this is the first study that demonstrates improved predictability of antidepressant treatment outcomes in depressed adults receiving combination antidepressant therapy by augmenting clinical measures with multiple biological measures. The integration of multiple biological markers (e.g., SNPs and metabolites) suggested the prognostic importance of hydroxylated sphingomyelins in the context of monoamine neurotransmission. The current workflow ranks biomarkers according to their contributions to improved predictability of antidepressant response and provides a foundation for future laboratory-based studies, which may identify novel molecular mechanisms of MDD pathophysiology, drug response, or both.

CODE AVAILABILITY
All raw and analyzed data and related materials, including programming code, are available upon request to Mayo Clinic Ventures or University of Texas Southwestern. PGRN-AMPS data have been deposited on dbGaP, study accession phs000670.v1.p1

REFERENCES
1. Hasin DS, Sarvet AL, Meyers JL, Saha TD, Ruan WJ, Stohl M, et al. Epidemiology of adult DSM-5 major depressive disorder and its specifiers in the United States. JAMA Psychiatry. 2018;75:336–46. https://doi.org/10.1001/jamapsychiatry.2017.4602.
2. Trivedi MH, Rush AJ, Wisniewski SR, Nierenberg AA, Warden D, Ritz L, et al. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR*D: implications for clinical practice. Am J Psychiatry. 2006;163:28–40. https://doi.org/10.1176/appi.ajp.163.1.28.
3. Rush AJ, Fava M, Wisniewski SR, Lavori PW, Trivedi MH, Sackeim HA, et al. Sequenced treatment alternatives to relieve depression (STAR*D): rationale and design. Control Clin Trials. 2004;25:119–42. https://doi.org/10.1016/s0197-2456(03)00112-0.
4. Fava M, Rush AJ, Trivedi MH, Nierenberg AA, Thase ME, Sackeim HA, et al. Background and rationale for the sequenced treatment alternatives to relieve depression (STAR*D) study. Psychiatr Clin North Am. 2003;26:457–94. https://doi.org/10.1016/s0193-953x(02)00107-7.x.
5. Rush AJ, Trivedi MH, Stewart JW, Nierenberg AA, Fava M, Kurian BT, et al. Combining medications to enhance depression outcomes (CO-MED): acute and long-term outcomes of a single-blind randomized study. Am J Psychiatry. 2011;168:689–701. https://doi.org/10.1176/appi.ajp.2011.10111645.
6. Blier P, Ward HE, Tremblay P, Laberge L, Hébert C, Bergeron R. Combination of antidepressant medications from treatment initiation for major depressive disorder: a double-blind randomized study. Am J Psychiatry. 2010;167:281–8. https://doi.org/10.1176/appi.ajp.2009.09020186.
7. Athreya A, Iyer R, Neavin D, Wang L, Weinshilboum R, Kaddurah-Daouk R, et al. Augmentation of physician assessments with multi-omics enhances predictability of drug response: a case study of major depressive disorder. IEEE Comput Intell Mag. 2018;13:20–31. https://doi.org/10.1109/MCI.2018.2840660.
8. Athreya AP, Brückl T, Binder EB, John Rush A, Biernacka J, Frye MA, et al. Prediction of short-term antidepressant response using probabilistic graphical models with replication across multiple drugs and treatment settings. Neuropsychopharmacology. 2021;46:1272–1282. https://doi.org/10.1038/s41386-020-00943-x.
9. Athreya AP, Neavin D, Carrillo-Roa T, Skime M, Biernacka J, Frye MA, et al. Pharmacogenomics-driven prediction of antidepressant treatment outcomes: a machine-learning approach with multi-trial replication. Clin Pharm Ther. 2019;106:855–65. https://doi.org/10.1002/cpt.1482.
10. Athreya, AP, Banerjee, SS, Neavin, D, Kaddurah-Daouk, R, Rush, AJ, Frye, MA, et al. Data-driven longitudinal modeling and prediction of symptom dynamics in major depressive disorder: Integrating factor graphs and learning methods. In 2017 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology, CIBCB 2017 [8058559] Institute of Electrical and Electronics Engineers Inc. https://doi.org/10.1109/CIBCB.2017.8058559.
11. Chekroud AM, Gueorguieva R, Krumholz HM, Trivedi MH, Krystal JH, McCarthy G. Reevaluating the efficacy and predictability of antidepressant treatments: a symptom clustering approach. JAMA Psychiatry. 2017;74:370–8. https://doi.org/10.1001/jamapsychiatry.2017.0025.
12. Chekroud AM, Zotti RJ, Shehzad Z, Gueorguieva R, Johnson MK, Trivedi MH, et al. Cross-trial prediction of treatment outcome in depression: a machine learning approach. Lancet Psychiatry. 2016;3:243–50. https://doi.org/10.1016/S2215-0366(15)00471-X.
13. Czysz AH, South C, Gadad BS, Arning E, Soyombo A, Bottiglieri T, et al. Can targeted metabolomics predict depression recovery? Results from the CO-MED trial. Transl Psychiatry. 2019;9:11 https://doi.org/10.1038/s41398-018-0349-6.
14. Lin E, Kuo PH, Liu YL, Yu YW, Yang AC, Tsai SJ. A deep learning approach for predicting antidepressant response in major depression using clinical and genetic biomarkers. Front Psychiatry. 2018;9:290 https://doi.org/10.3389/fpsyt.2018.00290.
15. Maciukiewicz M, Marshe VS, Hauschild AC, Foster JA, Rotzinger S, Kennedy JL, et al. GWAS-based machine learning approach to predict duloxetine response in major depressive disorder. J Psychiatr Res. 2018;99:62–68. https://doi.org/10.1016/j.jpsychires.2017.12.009.
16. Ji Y, Biernacka JM, Hebbring S, Chai Y, Jenkins GD, Batzler A, et al. Pharmacogenomics of selective serotonin reuptake inhibitor treatment for major depressive disorder: genome-wide associations and functional genomics. Pharmacogenomics J. 2013;13:456–63. https://doi.org/10.1038/tpj.2012.32.
17. Rush AJ, Trivedi MH, Ibrahim HM, Carmody TJ, Arnow B, Klein DN, et al. The 16-Item Quick Inventory of Depressive Symptomatology (QIDS), clinician rating (QIDS-C), and self-report (QIDS-SR): a psychometric evaluation in patients with chronic major depression. Biol Psychiatry. 2003;54:573–83. https://doi.org/10.1016/s0006-3223(02)01866-8.
18. Trivedi MH, Rush AJ, Ibrahim HM, Carmody TJ, Biggs MM, Suppes T, et al. The Inventory of Depressive Symptomatology, Clinician Rating (IDS-C) and Self-Report (IDS-SR), and the Quick Inventory of Depressive Symptomatology, Clinician Rating (QIDS-C) and Self-Report (QIDS-SR) in public sector patients with mood disorders: a psychometric evaluation. Psychol Med. 2004;34:73–82. https://doi.org/10.1017/s0033291703001107.
19. Athreya A. Making augmented human intelligence in medicine practical: a case study of treating major depressive disorder PhD thesis. University of Illinois at Urbana-Champaign (2019).
20. Habert J, Katzman MA, Oluboka OJ, McIntyre RS, McIntosh D, MacQueen GM, et al. Functional recovery in major depressive disorder: focus on early optimized treatment. Prim Care Companion CNS Disord. 2016;18, https://doi.org/10.4088/PCC.15r01926.
21. Biocrates Life Sciences. Absolute IDQ p180kit. Analytical specifications p180 (AS-p180-2 for research use). Innsbruck: Biocrates Life Sciences; 2010.
22. Bhattacharyya S, Dunlop BW, Mahmoudiandehkordi S, Ahmed AT, Louie G, Frye MA, et al. Pilot study of metabolomic clusters as state markers of major depression and outcomes to CBT treatment. Front Neurosci. 2019;13:926 https://doi.org/10.3389/fnins.2019.00926.
23. MahmoudianDehkordi S, Ahmed AT, Bhattacharyya S, Han X, Baillie RA, Arnold M, et al. Alterations in acylcarnitines, amines, and lipids inform about the mechanism of action of citalopram/escitalopram in major depression. Transl Psychiatry. 2021;11:153 https://doi.org/10.1038/s41398-020-01097-6.
24. Liu D, Ray B, Neavin DR, Zhang J, Athreya AP, Biernacka JM, et al. Beta-defensin 1, aryl hydrocarbon receptor and plasma kynurenine in major depressive disorder: metabolomics-informed genomics. Transl Psychiatry. 2018;8:10 https://doi.org/10.1038/s41398-017-0056-8.
25. Liu D, Zhuang Y, Zhang L, Gao H, Neavin D, Carrilo-Roa T, et al. ERICH3: vesicular association and antidepressant treatment response. Mol Psychiatry. 2020, https://doi.org/10.1038/s41380-020-00940-y.
26. Ho MF, Zhang C, Zhang L, Wei L, Zhou Y, Moon I, et al. TSPAN5 influences serotonin and kynurenine: pharmacogenomic mechanisms related to alcohol use disorder and acamprosate treatment response. Mol Psychiatry. 2020, https://doi.org/10.1038/s41380-020-0855-9.
27. Gupta M, Neavin D, Liu D, Biernacka J, Hall-Flavin D, Bobo WV, et al. TSPAN5, ERICH3 and selective serotonin reuptake inhibitors in major depressive disorder: pharmacometabolomics-informed pharmacogenomics. Mol Psychiatry. 2016;21:1717–25. https://doi.org/10.1038/mp.2016.6
28. Neavin D, Kaddurah-Daouk R, Weinshilboum R. Pharmacometabolomics informs Pharmacogenomics. Metabolomics. 2016;12, https://doi.org/10.1007/s11306-016-1066-x.
29. Loeuillet C, Weale M, Deutsch S, Rotger M, Soranzo N, Wyniger J, et al. Promoter polymorphisms and allelic imbalance in ABCB1 expression. Pharmacogenet Genomics. 2007;17:951–9. https://doi.org/10.1097/FPC.0b013e3282eff934.
30. Schatzberg AF, DeBattista C, Lazzeroni LC, Etkin A, Murphy GM JR, Williams LM. ABCB1 genetic effects on antidepressant outcomes: a report from the iSPOT-D trial. Am J Psychiatry. 2015;172:751–9. https://doi.org/10.1176/appi.ajp.2015.14050680.
31. Gadad BS, Raj P, Jha MK, Carmody T, Dozmorov I, Mayes TL, et al. Association of novel ALX4 gene polymorphisms with antidepressant treatment response: findings from the CO-MED trial. Mol Neuropsychiatry. 2018;4:7–19. https://doi.org/10.1159/000487321.
32. Steuer R. Review: On the analysis and interpretation of correlations in metabolomic data. Brief Bioinformatics. 2006;7:151–8. https://doi.org/10.1093/bib/bbl009.

33. Carsten F, Dormann JE, Bacher Sven, Buchmann Carsten, Carl Gudrun, Carré Gabriel, et al. Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. Ecography. 2012;36:27–46.

34. Kuhn, M. and K. Johnson, *Applied Predictive Modeling*. 1 ed. 2013: Springer-Verlag New York. 600.

35. Tidymodels: Easily install and load the'tidymodels' packages. 2019, https://CRAN.R-project.org/package=tidymodels.

36. R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna: 2020.

37. RStudio: Integrated Development for R. Boston, MA: RStudio, PBC; 2020.

38. Uppal K, Ma C, Go YM, Jones DP, Wren J. xMWAS: a data-driven integration and differential network analysis tool. Bioinformatics. 2018;34:701–2. https://doi.org/10.1093/bioinformatics/btx656.

39. Chun H, Keles S. Sparse partial least squares regression for simultaneous dimension reduction and variable selection. J R Stat Soc Ser B Stat Methodol. 2010;72:3–25. https://doi.org/10.1111/j.1467-9868.2009.00723.x.

40. Blondel VD, Guillaume J-L, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. J Stat Mech. 2008;P10008.

41. Nguyen TTL, Liu D, Ho MF, Athreya AP, Weinshilboum R. Selective serotonin reuptake inhibitor pharmaco-omics: mechanisms and prediction. Front Pharm. 2020;11:614048 https://doi.org/10.3389/fphar.2020.614048.

42. Dean J, Keshavan M. The neurobiology of depression: an integrated view. Asian J Psychiatr. 2017;27:101–11. https://doi.org/10.1016/j.ajp.2017.01.025.

43. Lee CH, Giuliani F. The role of inflammation in depression and fatigue. Front Immunol. 2019;10:1696 https://doi.org/10.3389/fimmu.2019.01696.

44. Rotroff DM, Corum DG, Motsinger-Reif A, Fiehn O, Bottrel N, Drevets WC, et al. Metabolomic signatures of drug response phenotypes for ketamine and esketamine in subjects with refractory major depressive disorder: new mechanistic insights for rapid acting antidepressants. Transl Psychiatry. 2016;6:e894 https://doi.org/10.1038/tp.2016.145.

45. Slotte JP. Biological functions of sphingomyelins. Prog Lipid Res. 2013;52:424–37. https://doi.org/10.1016/j.plipres.2013.05.001.

46. Olsen ASB, Faergeman NJ. Sphingolipids: membrane microdomains in brain development, function and neurological diseases. Open Biol. 2017;7, https://doi.org/10.1098/rsob.170069.

47. Hama H. Fatty acid 2-Hydroxylation in mammalian sphingolipid biology. Biochim Biophys Acta. 2010;1801:405–14. https://doi.org/10.1016/j.bbalip.2009.12.004.

48. Haughey NJ. Sphingolipids in neurodegeneration. Neuromolecular Med. 2010;12:301–5. https://doi.org/10.1007/s12017-010-8135-5.

49. van Kruining D, Luo Q, van Echten-Deckert G, Mielke MM, Bowman A, Ellis S, et al. Sphingolipids as prognostic biomarkers of neurodegeneration, neuroinflammation, and psychiatric diseases and their emerging role in lipidomic investigation methods. Adv Drug Deliv Rev. 2020;159:232–44. https://doi.org/10.1016/j.addr.2020.04.009.

50. Singh P, Chattopadhyay A. Removal of sphingomyelin headgroup inhibits the ligand binding function of hippocampal serotonin1A receptors. Biochem Biophys Res. Commun. 2012;419:321–5. https://doi.org/10.1016/j.bbrc.2012.02.019.

51. Jafurulla M, Bandari S, Pucadyil TJ, Chattopadhyay A. Sphingolipids modulate the function of human serotonin1A receptors: insights from sphingolipid-deficient cells. Biochim Biophys Acta Biomembr. 2017;1859:598–604. https://doi.org/10.1016/j.bbamem.2016.10.016.

52. Gulbins A, Schumacher F, Becker KA, Wilker B, Soddemann M, Boldrin F, et al. Antidepressants act by inducing autophagy controlled by sphingomyelin-ceramide. Mol Psychiatry. 2018;23:2324–46. https://doi.org/10.1038/s41380-018-0090-9.

53. Sudhof TC. The synaptic vesicle cycle. Annu Rev Neurosci. 2004;27:509–47. https://doi.org/10.1146/annurev.neuro.26.041002.131412.

54. Postila PA, Rog T. A perspective: active role of lipids in neurotransmitter dynamics. Mol Neurobiol. 2020;57:910–25. https://doi.org/10.1007/s12035-019-01775-7.

55. Athreya AP, Iyer R, Wang L, Weinshilboum RM, Bobo WV. Integration of machine learning and pharmacogenomic biomarkers for predicting response to antidepressant treatment: can computational intelligence be used to augment clinical assessments? Pharmacogenomics. 2019;20:983–8. https://doi.org/10.2217/pgs-2019-0119.

56. Gaynes BN, Warden D, Trivedi MH, Wisniewski SR, Fava M, Rush AJ. What did STAR*D teach us? Results from a large-scale, practical, clinical trial for patients with depression. Psychiatr Serv. 2009;60:1439–45. https://doi.org/10.1176/ps.2009.60.11.1439.

57. Zisook S, Rush AJ, Haight BR, Clines DC, Rockett CB. Use of bupropion in combination with serotonin reuptake inhibitors. Biol Psychiatry. 2006;59:203–10. https://doi.org/10.1016/j.biopsych.2005.06.027.

58. Köhler S, Unger T, Hoffmann S, Steinacher B, Fydrich T, Bschor T. Comparing augmentation with non-antidepressants over sticking to antidepressants after treatment failure in depression: a naturalistic study. Pharmacopsychiatry. 2013;46:69–76. https://doi.org/10.1055/s-0032-1323677.

59. Kotzan JA, Maclean R, Wade W, Martin BC, Lami H, Tadlock G, et al. Prevalence and patterns of concomitant use of selective serotonin reuptake inhibitors and other antidepressants in a high-cost polypharmacy cohort. Clin Ther. 2002;24:237–48. https://doi.org/10.1016/s0149-2918(02)85020-2.

60. Martín-López LM, Rojo JE, Gibert K, Martín JC, Sperry L, Duñó L, et al. The strategy of combining antidepressants in the treatment of major depression: clinical experience in spanish outpatients. Depress Res Treat. 2011;2011:140194 https://doi.org/10.1155/2011/140194.

61. Mischoulon D, Nierenberg AA, Kizilbash L, Rosenbaum JF, Fava M. Strategies for managing depression refractory to selective serotonin reuptake inhibitor treatment: a survey of clinicians. Can J Psychiatry. 2000;45:476–81. https://doi.org/10.1177/070674370004500509.

62. Valenstein M, McCarthy JF, Austin KL, Greden JF, Young EA, Blow FC. What happened to lithium? Antidepressant augmentation in clinical settings. Am J Psychiatry. 2006;163:1219–25. https://doi.org/10.1176/appi.ajp.163.7.1219.

63. Jha MK, Minhajuddin A, Gadad BS, Greer T, Grannemann B, Soyombo A, et al. Can C-reactive protein inform antidepressant medication selection in depressed outpatients? Findings from the CO-MED trial. Psychoneuroendocrinology. 2017;78:105–13. https://doi.org/10.1016/j.psyneuen.2017.01.023.

64. Uher R, Tansey KE, Dew T, Maier W, Mors O, Hauser J, et al. An inflammatory biomarker as a differential predictor of outcome of depression treatment with escitalopram and nortriptyline. Am J Psychiatry. 2014;171:1278–86. https://doi.org/10.1176/appi.ajp.2014.14010094.

65. Jha MK, Wakhlu S, Dronamraju N, Minhajuddin A, Greer TL, Trivedi MH. Validating pre-treatment body mass index as moderator of antidepressant treatment outcomes: Findings from CO-MED trial. J Affect Disord. 2018;234:34–37. https://doi.org/10.1016/j.jad.2018.02.089.

66. Hui Zou TH. Regularization and variable selection via the elastic net. J R Stat Soc B. 2005;67:301–20.

67. Russell SJ, Norvig P. Artificial intelligence: a modern approach. 4th ed. Pearson; 2021.

68. Kuk AY, Li J, Rush AJ. Recursive subsetting to identify patients in the STAR*D: a method to enhance the accuracy of early prediction of treatment outcome and to inform personalized care. J Clin Psychiatry. 2010;71:1502–8. https://doi.org/10.4088/JCP.10m06168blu.

69. Bicrates Life Sciences. Annotation of potential isobaric and isomeric lipid species analyzed using the MxP® Quant 500 Kit. 2020, https://biocrates.com/wp-content/uploads/2020/02/Biocrates_Q500_isomers_isobars.pdf.

70. Szegedi A, Müller MJ, Anghelescu I, Klawe C, Kohnen R, Benkert O. Early improvement under mirtazapine and paroxetine predicts later stable response and remission with high sensitivity in patients with major depression. J Clin Psychiatry. 2003;64:413–20. https://doi.org/10.4088/jcp.v64n0410.

71. Townsend MK, Clish CB, Kraft P, Wu C, Souza AL, Deik AA, et al. Reproducibility of metabolomic profiles among men and women in 2 large cohort studies. Clin Chem. 2013;59:1657–67. https://doi.org/10.1373/clinchem.2012.199133.

72. Ang JE, Revell V, Mann A, Mäntele S, Otway DT, Johnston JD, et al. Identification of human plasma metabolites exhibiting time-of-day variation using an untargeted liquid chromatography-mass spectrometry metabolomic approach. Chronobiol Int. 2012;29:868–81. https://doi.org/10.3109/07420528.2012.699122.

73. Stevens VL, Hoover E, Wang Y, Zanetti, KA. Pre-analytical factors that affect metabolite stability in human urine, plasma, and serum: a review. Metabolites. 2019;9, https://doi.org/10.3390/metabo9080156.

ACKNOWLEDGEMENTS

AUTHOR CONTRIBUTIONS

JBJ, CWG, WVB, PEC and APA designed the study. JBJ and CWG conducted analyses of the data under the supervision of APA. MT, LW, RMW, WVB, DL, MAF, TM, MS, TC, and

JMB designed and assisted with PGRN-AMPS and CO-MED studies. RKD and SM assisted with metabolomics data from PGRN-AMPS study. JBJ, CWG, APA, WVB, and PEC co-drafted the manuscript and all authors contributed to the interpretation of the findings, reviewed, and approved the final version of the manuscript.

COMPETING INTERESTS

Drs L Wang and RM Weinshilboum are co-founders and stockholders in OneOme. Dr. WV Bobo's research has been supported by the National Institute of Mental Health, the Agency for Healthcare Quality and Research, the National Science Foundation, the Myocarditis Foundation, and the Mayo Foundation for Medical Education and Research. He has contributed chapters to UpToDate concerning the pharmacological treatment of adults with bipolar major depression. Dr. PE Croarkin has received research grant support from Neuronetics, Inc., NeoSync, Inc., and Pfizer, Inc. He has received grant in-kind (equipment and supply support for research studies) from Assurex Health, Neuronetics, Inc., and MagVenture, Inc. Dr. PE Croarkin has served as a consultant for Engrail Therapeutics, Myriad Neuroscience, Procter & Gamble, and Sunovion. Dr. MH Trivedi has provided consulting services to Acadia Pharmaceuticals, Inc., Alkermes, Inc., Alto Neuroscience, Inc., Axsome Therapeutics, GH Research Limited, GreenLight VitalSign6, Inc., Janssen, Merck Sharp & Dohme Corp., Neurocrine Biosciences, Inc., Orexo US, Inc., Otsuka, SAGE Therapeutics, Signant Health, Titan Pharmaceuticals, Inc. He has received grant/research funding from NIMH, NIDA, Patient-Centered Outcomes Research Institute (PCORI), and Cancer Prevention Research Institute of Texas (CPRIT). In addition, he has received editorial compensation from Oxford University Press. Dr. T Carmody has provided consulting services to Alkermes, Inc. Dr. R Kaddurah-Daouk is an inventor on key patents in the field of metabolomics in the study of CNS diseases and holds equity in Metabolon, Inc., a biotechnology company that provides metabolic profiling capabilities. Dr. MA Frye has received research support from Assurex Health, Myriad, Mayo Foundation, and Medibio; he has served as a consultant for Actify Neurotherapies, Allergan, Intra-Cellular Therapies, Inc., Janssen, Myriad, Neuralstem, Inc., Sanofi, Takeda, Teva Pharmaceuticals; he has received CME Travel/Honoraria from American Physician Institute, CME Outfitters, Global Academy for Medical Education. All other authors declared no competing interests for this work.

ADDITIONAL INFORMATION

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41398-021-01632-z.

Correspondence and requests for materials should be addressed to Arjun P. Athreya.

Reprints and permission information is available at http://www.nature.com/reprints

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