Genome-wide association study of treatment resistant depression highlights shared biology with metabolic traits

JooEun Kang[1], Victor M. Castro[2,3,4], Michael Ripperger[5], Sanan Venkatesh[6-12], David Burstein[6-12], Richard Karlsson Linnér[13,14], Daniel B. Rocha[15], Yirui Hu[16], Drew Wilimitis[5], Theodore Morley[1], Lide Han[1], Rachel Youngjung Kim[2,3], Yen-Chen Anne Feng[17], Tian Ge[2], Stephan Heckers[18], Georgios Voloudakis[6-12], Christopher Chabris[13], Panos Roussos[6-12], Thomas H McCoy Jr[2,3], Colin G. Walsh[5,18,19], Roy H. Perlis[2,3*], Douglas M. Ruderfer[1,5,18,19*]

1. Division of Genetic Medicine, Department of Medicine, Vanderbilt Genetics Institute, Vanderbilt University Medical Center, 1211 Medical Center Dr. Nashville, TN 37232 USA
2. Department of Psychiatry, Massachusetts General Hospital, Boston, MA
3. Center for Quantitative Health, Massachusetts General Hospital, Boston, MA
4. Research Information Science and Computing, Mass General Brigham, Somerville, MA
5. Department of Biomedical Informatics, Vanderbilt University Medical Center, 1211 Medical Center Dr. Nashville, TN 37232 USA
6. Department of Psychiatry, Icahn School of Medicine at Mount Sinai, New York, NY, USA
7. Center for Disease Neurogenomics, Icahn School of Medicine at Mount Sinai, New York, NY USA
8. Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA
9. Department of Genetics and Genomic Science, Icahn School of Medicine at Mount Sinai, New York, NY, USA
10. Icahn Institute for Data Science and Genomic Technology, Icahn School of Medicine at Mount Sinai, New York, NY, USA
11. Nash Family Department of Neuroscience, Icahn School of Medicine at Mount Sinai, New York, NY, USA
12. Mental Illness Research, Education, and Clinical Center (VISN 2 South), James J. Peters VA Medical Center, Bronx, NY, USA
13. Autism and Developmental Medicine Institute, Geisinger, 120 Hamm Dr., PA 17817 Lewisburg, USA
14. Department of Economics, Universiteit Leiden, Steenschuur 25, 2311ES Leiden, the Netherlands
15. Phenomic Analytics and Clinical Data Core, Geisinger, Danville, PA 17822
16. Population Health Sciences, Geisinger, Danville, PA 17822
17. Institute of Epidemiology and Preventive Medicine, National Taiwan University, Taipei, Taiwan
18. Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center, 1211 Medical Center Dr. Nashville, TN 37232 USA
19. Department of Medicine, Vanderbilt University Medical Center, 1211 Medical Center Dr. Nashville, TN 37232 USA

* Corresponding authors:
Clinical: Roy Perlis rperlis@mgh.harvard.edu
Genomic: Douglas Ruderfer douglas.ruderfer@vumc.org

Abstract

Treatment-resistant depression (TRD), often defined by absence of symptomatic remission following at least two adequate treatment trials, occurs in roughly a third of all individuals with major depressive disorder (MDD). Prior work has suggested a significant common variant genetic component of liability to TRD, with heritability estimates of 8% when comparing to non-treatment

resistant MDD. Despite this evidence of heritability, no replicated genetic loci have been identified and the genetic architecture of TRD remains unclear. A key barrier to this work has been the paucity of adequately powered cohorts for investigation, largely because of the challenge in prospectively investigating this phenotype. Using electroconvulsive therapy (ECT) as a surrogate for TRD, we applied standard machine learning methods to electronic health record (EHR) data to derive predicted probabilities of receiving ECT. We applied these probabilities as a quantitative trait in a genome-wide association study (GWAS) over 154,433 genotyped patients across four large biobanks. With this approach, we demonstrate heritability ranging from 2% to 4.2% and significant genetic overlap with cognition, attention deficit hyperactivity disorder, schizophrenia, alcohol and smoking traits and body mass index. We identify two genome-wide significant loci, both previously implicated in metabolic traits, suggesting shared biology and potential pharmacological implications. This work provides support for the utility of estimation of disease probability for genomic investigation and provides insights into the genetic architecture and biology of TRD.

Introduction

Depression is a common, disabling mental illness, with lifetime prevalence estimates ranging from 6.6% to 21% globally[1] and 16.9% in the United States[2]. Of the individuals with depression, more than 40% do not respond to 2 sequential antidepressant therapies and a third do not respond after 4 different treatments[3]. Treatment resistant depression (TRD) disproportionately accounts for the socioeconomic burden of depression, with over $25 billion spent annually in the United States[4] and is associated with a significantly increased risk for suicide[5,6]. After decades of

stasis, novel interventions for TRD have begun to emerge; however, such treatments remain costly and challenging to access, highlighting the need to better understand risk factors for this outcome[7,8].

Despite decades of investigation[9], the neurobiology of TRD is poorly understood. Prior studies have suggested a significant genetic component of TRD, with heritability estimates from common genetic variation ranging from 17% to 25% when compared to healthy controls[10,11] and 8% when compared to non TRD MDD[11]. However, no identified genetic risk locus has been replicated in genome-wide association studies (GWAS). This likely reflects two key barriers to discovery: first, the challenges of attaining sufficient power for a heterogeneous phenotype and second, variable criteria for treatment responsiveness in defining TRD[12–15]. Limited consensus exists on the exact measures of remission, length of adequate treatment trial duration, and adequate treatment dose needed to define TRD[16].

To address these challenges, we adopted two strategies. First, to increase power, we used large-scale clinical data to build risk prediction models where quantitative phenotypes can be generated for genetic samples in associated biobanks[17]. Secondly, to bypass the problems associated with categorizing treatment responsivity to specific medication classes or to specific number of antidepressant trials, we defined TRD on the basis of whether an individual with MDD had received a gold standard treatment indicated for TRD, electroconvulsive therapy (ECT)[18]. Together, these approaches enable the large-scale genetic analyses needed to characterize the genetic architecture of TRD and improve our understanding of its biological etiology.

Specifically, using ECT as a surrogate for TRD, we applied prediction models to electronic health record (EHR) data to derive posterior probabilities of receiving ECT, as absolute numbers of ECT cases in individual health systems were modest. After internally and externally validating these models, we used the probabilities as quantitative phenotypes to perform GWAS on over

154,000 patients with MDD across four large biobanks. We quantified the genetic contribution to TRD and the genetic overlap with both psychiatric and non-psychiatric phenotypes and other definitions of TRD. Finally, we implicated specific loci associated to increased likelihood of having TRD and needing ECT.

Methods

Study Settings

Clinical and genetic data were used from the EHRs and biobanks of Mass General Brigham (MGB), Vanderbilt University Medical Center (VUMC), Geisinger Health System (Geisinger; GHS), and the Million Veteran Program (MVP). MGB consists of 2 academic medical centers and 4 community and psychiatric hospitals in Eastern Massachusetts that serve over 6.5 million patients, and electronic health data were extracted from the Mass General Brigham Research Patient Data Registry (RPDR)[19] and the Enterprise Data Warehouse. VUMC is an academic medical center in Nashville, Tennessee that manages over 2 million patient visits across Tennessee and its neighboring states each year. Its deidentified clinical EHR data is stored in the Synthetic Derivative (SD)[20]. Geisinger Health System is an academic medical center in Danville Pennsylvania and serve over 3 million patients in Pennsylvania. Deidentified electronic health data for consenting patients is extracted and stored by the Geisinger MyCode Community Health Initiative[21]. The Million Veteran Program[22] study is based on the largest integrated health care system in the United States (Veterans Health Administration; 9 million people) and has over 825,000 US veteran participants with genetic information available for over 650,000 individuals.

Clinical prediction model of TRD

We extracted de-identified clinical data of individuals with current ages of 18-90 years from the VUMC SD and MGB RPDR (**Figure 1A**). Individuals with MDD were identified using International Classification of Diseases, version 9 (ICD-9) codes 311.*, 296.2*, 296.3*, and 300.4 and ICD-10 codes F32.**, F33.**, and F34.1 with * as wildcard digits 0-9. Individuals with one or more ICD-9 or ICD-10 codes for bipolar disorders, schizophrenia, and psychotic disorders were excluded from analyses. Of the remaining individuals, TRD cases were defined using the CPT code for ECT (90870). All data 24 hours before the date of ECT for cases or the last depression code for controls were right-censored to avoid surrogates for the outcome (for example, the possibility that particular laboratory studies or pre-anesthesia procedures would directly proxy ECT). A minimum of at least two unique visit dates over four weeks before censoring date was required for study inclusion.

Structured clinical data were included as predictors for the clinical model, including: demographics (age in years, categorical sex [Male, Female, Unknown], categorical race [White, Black, Asian, Hispanic, Other]), area deprivation index (ADI), diagnostic codes (log-transformed counts of historical CCS counts)[23], and medication (log-transformed counts of RXNORM-mapped ingredients). Of note, the VUMC ADI uses six features from the American Community Survey on the census tract level[24], while MGB ADI includes 21 socioeconomic factors from the census on the zip-code level[25].

The VUMC dataset was split into training and test sets where the test sample was comprised of only patients in the biobank. The remaining sample was then randomly split into 80% for training and 20% for validation. In MGB, the dataset was randomly split into 80% for training

and 20% for testing regardless of biobank status. A LASSO model[26] was trained separately at each site using Glmnet[27] and hyperparameters were trained via a 10-fold cross-validation on the training data set (**Figure 1B**).

Each prediction model was validated internally and then externally at the other partner site (**Figure 1B**). Both the MGB and VUMC TRD models were further validated at Geisinger and MVP. Model performance was evaluated with discrimination metrics: Area Under the Receiver Operating Characteristic (AUROC) and Area Under the Precision-Recall Curve (AUPRC). Each genotyped individual from VUMC, MGB, Geisinger, and MVP had two predicted probabilities of ECT from either the MGB or VUMC TRD model, and these probabilities were used as quantitative phenotypes for genetic association analyses (**Figure 1C**).

Phenome-wide association study (PheWAS)

ICD-9 and ICD-10 codes were mapped to phecodes using version 1.2[28]. Cases were defined as those having more than 2 instances of the same phecode. Controls were patients never having documentation of the phecode. Patients were excluded if they had 1 instance of the phecode or were a case from a predefined list of excluded phecodes. Logistic regression was performed with presence or absence of phecode as the outcome. All phecodes with more than 100 cases and at least 1 case having received ECT were analyzed. The PheWAS R package[29] was used to visualize results.

Genotyping quality control, imputation and GWAS meta-analysis across four sites

Genotyping quality control (QC), imputation and GWAS were performed separately at each site. Standard QC protocols were applied removing variants based on high levels of genotype

missingness or failing Hardy-Weinberg Equilibrium and removing individuals for excessive genotype missingness, high heterozygosity or sex discrepancies. Only individuals of European ancestries were retained as defined by the 1000 Genomes reference. Imputation and GWAS were performed using comparable reference panels and approaches. Site specific details are provided in the Supplementary Methods.

GWAS of the VUMC and MGB TRD phenotypes in the individuals of European ancestries in 4 different clinical sites (VUMC, MGB, GHS, MVP) were meta-analyzed using inverse variance-weighted fixed effects model in METAL[30]. The weighted mean allele frequency was calculated weighted by the effective sample size per cohort. SNPs with a weighted minor allele frequency of < 1% or SNPs present in < 80% of total effective sample size were removed from the meta-analysis results.

Heritability estimates and genetic correlation

LD score regression[31] was used to estimate the phenotypic variance in TRD explained by common SNPs (SNP-heritability, h^2_{SNP}) from GWAS summary statistics. h^2_{SNP} was calculated on the observed scale. LDSC bivariate genetic correlations attributable to genome-wide SNPs (rg) were estimated between GWAS of the TRD phenotypes and previously published GWAS of ECT[32] or medication-defined TRD[11], as well as other psychiatric and non-psychiatric risk factors. For previously noted epidemiological risk factors of TRD, the rg of TRD GWAS summary statistics with 29 other human diseases and traits was calculated using publicly available summary statistics (PMID listed in **Table S6**). The Bonferroni corrected significance threshold was $P < 1.72 \times 10^{-3}$, adjusting for 29 traits tested. Differences in rg between VUMC TRD meta-analysis and MGB TRD meta-analysis and differences in heritability between TRD meta-analyses before and after mtCOJO

conditioning for BMI were tested for deviation from 0, using the block jackknife method, implemented in LDSC software[33].

Conditional GWAS using mtCOJO

The results of the GWAS of TRD were conditioned on the genetics of BMI using mtCOJO (multi-trait-based conditional & joint analysis using GWAS summary data)[34], implemented in GCTA software[35]. mtCOJO estimates the effect size of a SNP on an outcome trait (eg. TRD) conditioned on exposure trait(s) (eg. BMI), using the genome-wide significant SNPs for the exposure trait as instruments to estimate the effect of the exposure on the outcome. It then performs a genome-wide conditioning of the estimated effect from the exposure, which provides conditioned effect sizes and P values for the outcome trait. We conditioned TRD on BMI, since higher BMI among TRD cases has been previously reported[11]. mtCOJO analysis was performed on the TRD meta-analyses as the outcome traits with the GIANT European ancestries GWAS summary statistics[36] as the exposure trait since mtCOJO requires an ancestry-matched LD reference panel. In the selection of SNPs as instruments, independence was defined as SNPs more than 1 megabase (Mb) apart or with an LD r2 value < 0.05 based on the 1000 Genomes Project Phase 3 European reference panel[37].

Polygenic risk scoring (PRS)

PRS of the TRD meta-analyses were tested for association with ECT CPT code as well as the TRD phenotypes generated in independent target cohorts. The target cohorts were patients with MDD at VUMC and MGB. The meta-analysis of the TRD phenotypes was repeated excluding each cohort in turn to create independent discovery and target datasets. PRS analyses were

performed using PRS-CS which places a continuous shrinkage prior on SNP effect sizes using a Bayesian regression framework[38]. The continuous shrinkage priors adapt the amount of shrinkage applied to each SNP to the strength of the associated GWAS signal based on the LD structure estimated from an external reference panel. Posterior SNP effects were generated in each cohort using PRS-CS and the 1000 Genomes European reference panel was used to estimate LD between SNPs. The PRS were calculated for each individual of the target cohort using Plink 1.9. PRS was tested for association with ECT cases vs control status in the target cohort using logistic regression model, covarying for PC1-PC10, sex, and age. PRS was also tested for association with MGB and VUMC TRD phenotypes using linear regression, covarying for PC1-PC10.

Results

Demographic and phenotypic characterization of patients with MDD receiving ECT across two healthcare systems

Leveraging longitudinal clinical data from EHRs at MGB and VUMC (see Methods), we identified 185,409 patients (MGB: 78,620, VUMC: 106,789) with a diagnostic code of MDD or depressive disorder. Depressive disorder was included as prior work in these health systems and others indicated that it is commonly applied by non-psychiatrists to capture MDD. Among those patients, 467 (MGB: 242, VUMC: 225) had at least one procedural code for ECT. The prevalence of ECT among individuals with MDD was 0.26% (MGB: 0.31%, VUMC: 0.21%) which is similar to the published prevalence of ECT of ~0.25% among individuals with mood disorders[39]. The mean age at which VUMC cases received their first ECT CPT code was 53.8 ± 17.4 years, with a

median ECT trial number of 15 (SD = 16); at MGB the mean age was 57 ± 17 years with a mean ECT trial number of 16 (SD = 19).

We identified multiple significant demographic differences between MDD patients receiving ECT and those who have not (**Table 1**). Consistent across the VUMC and MGB cohorts, ECT cases were 5 years older on average, 12% more likely to be male (although ECT was still more common in women), and 8.8% more likely to be white. However, while the VUMC ECT cases had lower mean body mass index (BMI) on average and at first ECT visit, BMI variables in the MGB cohort were comparable between cases and controls.

		VUMC patients			*MGB patients*		
Demographic		**ECT case (N=225)**	**MDD control (N=106,564)**	**P-value**	**ECT case (N=242)**	**MDD control (N=78,378)**	**P-value**
Age	Mean at earliest MDD	51.2 ± 17.5	46.6 ± 19.3	**1.23E-04**	56.7 ± 16.3	51.2 ± 17.1	**6.70E-07**
	Mean at earliest ECT	53.8 ± 17.4			57 ± 17		
Gender	Female	121 (53.8%)	70979 (66.6%)	**6.01E-05**	145 (60%)	55,334 (71%)	**0.002**
	Male	104 (46.2%)	35548 (33.4%)	**6.00E-05**	97 (40%)	22,819 (29%)	**0.002**
	Unknown	0 (0%)	1 (0%)	1	0 (0%)	3 (<0.1%)	1
Race	White	200 (88.9%)	88741 (83.3%)	**0.031**	220 (91%)	62,013 (79%)	**9.03E-06**
	Black	9 (4.0%)	11552 (10.8%)	**1.41E-03**	5 (2.1%)	4,141 (5.3%)	**0.025**
	Asian	3 (1.3%)	1064 (1.0%)	0.87	5 (2.1%)	1,664 (2.1%)	> 0.9
	Other	13 (5.8%)	5171 (4.8%)	0.63	12 (5.0%)	10,338 (13.2%)	**2.17E-04**
Ethnicity	Hispanic	3 (1.3%)	2189 (2.1%)	0.60	4 (1.7%)	6,883 (8.8%)	**1.38E-04**
	Non-Hispanic	213 (94.7%)	99769 (93.7%)	0.63	238 (98%)	71,273 (91%)	**1.38E-04**
	Unknown	9 (4.0%)	4570 (4.3%)	0.96	0 (0%)	0 (0%)	NA
BMI	Mean	26.9 ± 6.81	28.2 ± 7.81	**4.77E-04**	28.1 ± 6.3	28.5 ± 6.6	0.4
	Closest to index date	27.2 ± 6.81	29.8 ± 8.31	**2.10E-07**	28.0 ± 6.7	28.5 ± 6.8	0.3

Table 1: Demographic characteristics of MGB and VUMC cohorts
In parentheses are percentages, standard deviations are reported after ±. Index time is 24 hours prior to first ECT for cases or last MDD code for controls. Age is defined as years between birth date and last EHR event. All non-listed races including unknown, and combinations are including in Other. BMI uses BMI values were cleaned for extreme outliers (> 80), unit mismatch and exclude individual measurements of age < 18. For ECT cases, the closest BMI measurement must have been within six months of the earliest ECT CPT code. Significance testing included t-test for quantitative values and two proportions Z-test for categorical variables

As TRD is associated with increased suicidality and higher burden of comorbid psychiatric illnesses, we performed a phenome-wide association study (PheWAS) to test the relationship of ECT with other phenotypes (see Methods). PheWAS association results in the VUMC and MGB cohorts were highly correlated (r = 0.70. P = 4.19×10^{-54}), with suicidality being the most significantly associated phenotype in both VUMC and MGB (VUMC: beta = 3.62, SE = 0.15, P =

2.67×10^{-128}; MGB: beta = 2.57, SE = 0.18, P = 2.58×10^{-46}). Other significantly associated phenotypes include psychiatric diseases like major depressive disorder, generalized anxiety disorder and other suicide-related traits (**Supplementary Figures 1-2**).

Internal and external validation of TRD phenotype

Models built to predict ECT separately at MGB and VUMC were tested both internally and externally (**Figure 1**). Internal validation of the TRD models showed discrimination performance metrics on the held-out test sets at MGB and VUMC of AUROC 0.81 and 0.9, respectively, and AUPRC of 0.03 and 0.04, respectively (**Table 2**). External validation of the MGB TRD model on VUMC data showed AUROC of 0.83, AUPRC of 0.03 and applying the VUMC TRD model on MGB data showed AUROC of 0.83 and AUPRC of 0.03.

To increase sample size and power for genetic analysis, both models were applied to samples at two additional sites (**Table 2**), the Geisinger Health System (GHS: 353 cases, 190,841 controls) and the Million Veteran Program (MVP: 600 cases, 259,925 controls). Prediction performance remained consistently high for both models at GHS (VUMC model: AUROC: 0.84, AUPRC: 0.021; MGB model: AUROC: 0.78, AUPRC: 0.023) and MVP (VUMC model: AUROC: 0.81, AUPRC: 0.024; MGB model: AUROC: 0.81, AUPRC: 0.040). Features selected by LASSO with the highest weights included prescriptions of antipsychotics, diagnosis of mood disorders, and suicide in both models (**Table S1, Table S2**).

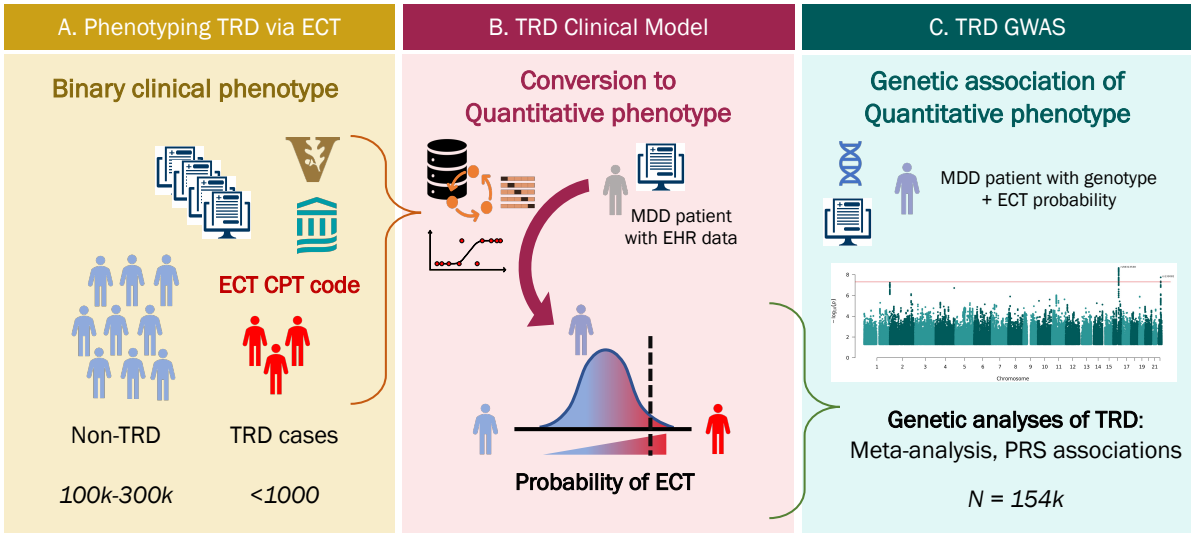

Figure 1: Schematic of the TRD clinical model generation and the genome-wide association study of the quantitative ECT prediction scores.
A. TRD cases and non-TRD MDD controls were extracted from EHRs at VUMC and MGB B. Prediction models leveraging clinical data from the EHR were trained separately at VUMC and MGB and applied to 4 independent data sets to generate probabilities C. Probabilities were used as quantitative TRD phenotypes to perform two GWAS of TRD in 154,443 individuals.

		Test Set		*VUMC model*		*MGB model*	
Site	**ECT cases**	**Controls**	**Prevalence**	**AUROC**	**AUPRC**	**AUROC**	**AUPRC**
VUMC	59	33,248	0.18%	**0.90**	**0.036**	0.83	0.030
MGB	207	43,520	0.47%	0.83	0.030	**0.91**	**0.080**
GHS	353	190,841	0.18%	0.84	0.021	0.78	0.023
MVP	600	259,925	0.23%	0.81	0.024	0.81	0.040

Table 2: **Performance metrics of VUMC and MGB models across 4 sites.** Test set is the independent sample both of the models were applied to. AUROC: area under the receiver operator curve; AUPRC: area under the precision recall curve. Bolded numbers are performance measures of internal validation.

TRD phenotypes show significant heritability and shared genetic architecture with each other

The posterior probabilities from the two TRD prediction models (VUMC and MGB) were rank normalized to generate two quantitative TRD phenotypes. Linear regressions of phenotype on genotype were performed separately on samples of European ancestries at VUMC (n = 15,305), MGB (n = 2,216), GHS (n = 39,353) and MVP (n = 97,649). Summary statistics from the four

GWAS were then meta-analyzed across 154,433 samples using a variance-weighted fixed effect model. Significant heritability estimates of 0.042 (SE = 0.004, P = 4.90×10^{-23}) for the MGB TRD meta-analysis and 0.020 (SE = 0.0036, P = 1.38×10^{-8}) for the VUMC TRD meta-analysis were estimated using LD-score regression[31] (**Table 3**). Significant genetic correlation between the VUMC TRD meta-analysis and the MGB TRD meta-analysis was observed (rg = 0.66, SE 0.05, P=1.3×10^{-34}) (**Table S3**). The rg value reflects highly overlapping but non-identical phenotypes.

		VUMC TRD			*MGB TRD*		
Cohort	**N**	**SNP h2**	**SE**	**p**	**SNP h2**	**SE**	**p**
GHS	39,353	0.042	0.0117	**1.66E-04**	0.029	0.012	**0.008**
VUMC	15,305	0.017	0.025	0.248	0.0226	0.029	0.218
MGB	2,126	0.043	0.188	0.410	0.259	0.1939	0.447
MVP	97,649	0.026	0.0054	**7.37E-07**	0.050	0.0063	**1.04E-15**
Meta-analysis	*154,433*	*0.020*	*0.0036*	***1.38E-08***	*0.0422*	*0.0043*	***4.90E-23***

Table 3: Heritability estimates of TRD GWAS within site and meta-analyses. Heritability was estimated with LD-score regression within each biobank site and the meta-analysis for both the VUMC and MGB TRD phenotypes.

We then examined the genetic correlation of our TRD phenotypes with two prior GWAS of TRD (**Table S3**). The first defined TRD based on antidepressant prescriptions in the UK Biobank (UKB TRD)[11] and the second used ECT but compared cases to healthy controls as opposed to those with non-TRD MDD (PREFECT)[32]. No significant genetic correlation was observed between the MGB TRD meta-analysis or the VUMC TRD meta-analysis with either PREFECT TRD (VUMC: rg = 0.14, SE = 0.12, P = 0.23; MGB: rg = 0.09, SE = 0.10, P = 0.31) or UKB TRD (VUMC: rg = 0.07, SE = 0.18, P = 0.65; MGB: rg = -0.05, SE = 0.13, P = 0.69). UKB TRD and PREFECT TRD were significantly correlated with each other (rg = 0.75, SE = 0.24, P = 0.003). Further, genome-wide significant loci from the previously published TRD GWAS[32,40] did not replicate at nominal significance in either of our TRD model meta-analyses (**Table S4).**

Two novel genome-wide significant loci associated to TRD

Two genome-wide significant loci were identified in the MGB TRD meta-analysis (**Figure 2**). The most significant locus was located on chromosome 16 in the intronic region of *FTO* (index SNP = rs56313538, beta for G allele = -0.0220, SE = 0.0037, MAF = 0.4, P = 4.3×10^{-11}, Cochran's Q: 0.35, I^2 heterogeneity index = 9.53) (**Figure 3A, Table S5**). No significant association was observed in the VUMC TRD meta-analysis (beta = -0.003, SE = 0.0037, P = 0.37). The SNP is in high LD (R^2 = 1.0) with SNP rs9939609 that is strongly associated with BMI[36] (beta=0.075, SE=2.9×10^{-3}, P = 1.95×10^{-145}) and weight[41] via its regulation of *IRX3* expression[42]. Therefore, this locus has an inverse effect on BMI and TRD.

The second genome-wide significant locus was in an intergenic region on chromosome 22 (index SNP = rs133082, beta for C allele = 0.0206, SE = 0.0037, MAF = 0.44, P = 1.82×10^{-8}, Cochran's Q: 0.71, I^2 heterogeneity index = 0) (**Figure 3B, Table S5**). Significant association was observed in the VUMC TRD meta-analysis although it did not surpass genome-wide correction (beta = 0.0158, SE = 0.0037, P = 1.83×10^{-5}). This variant is significantly associated with decreased expression of the nearest gene, melanin concentrating hormone receptor 1 *(MCHR1*, 15kb away) in the dorsal lateral prefrontal cortex and increased expression in whole blood and T cells[43]. The variant is in high linkage disequilibrium (R^2 = 0.76) with a genome-wide significant locus linked to increased risk of bipolar disorder in a recent large GWAS which implicated the same gene[44]. Neither locus is significantly associated in prior published TRD GWAS.

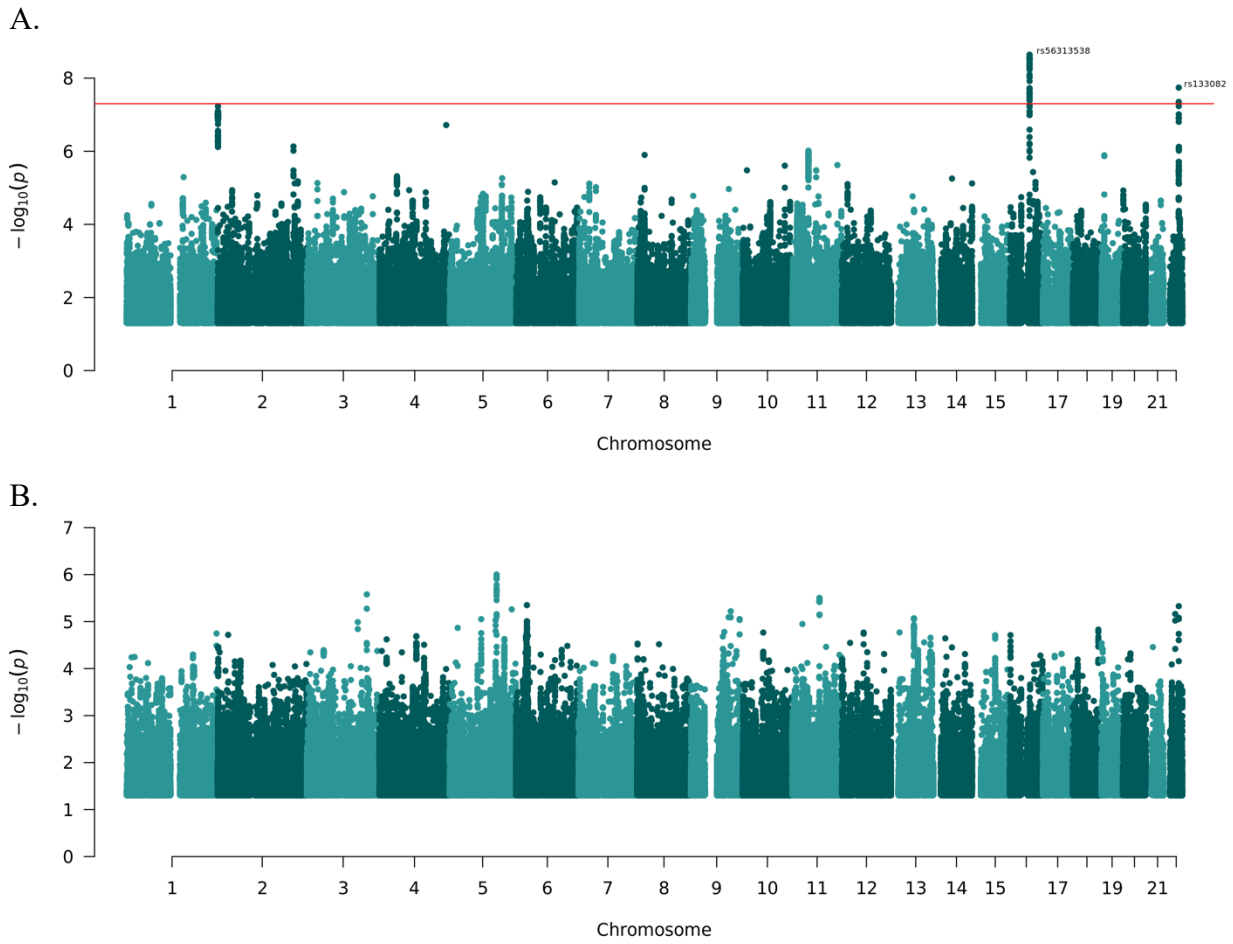

Figure 2: Manhattan plots of A. MGB TRD meta-analysis and B. VUMC TRD meta-analyses

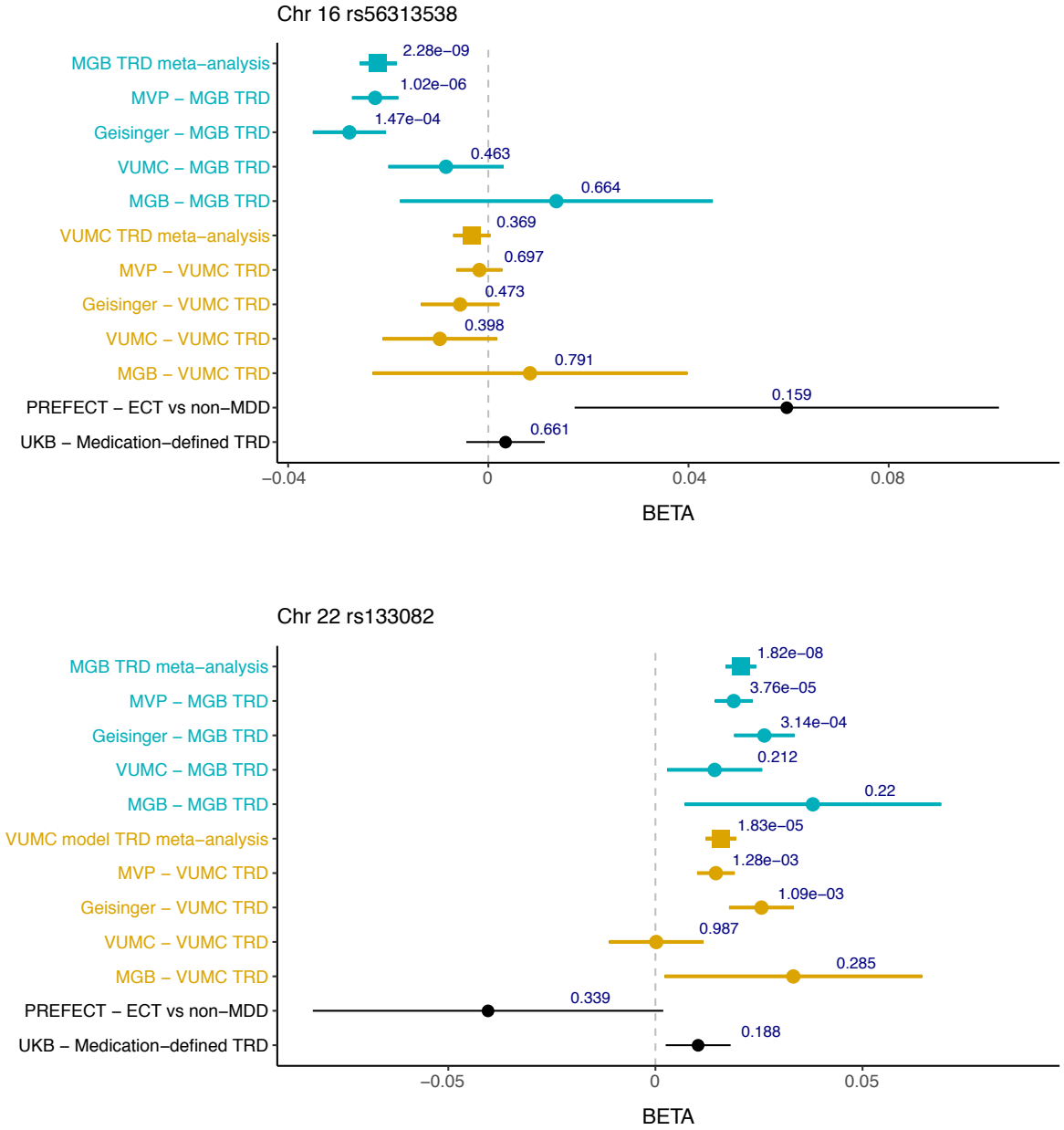

Figure 3: Forest plot of the genome-wide significant locus rs8050136 on chromosome 16 and rs133082 on chromosome 22. The points indicate the log odds ratio and the error bars show the standard error. The P value of association with each phenotype is shown above the error bars.

TRD polygenic risk scores associate with TRD phenotypes

Polygenic risk scores are a standard approach to collapsing aggregated risk from genome-wide association studies[45]. We tested for association of polygenic risk scores generated using our

TRD meta-analyses and our TRD phenotypes in the VUMC or MGB samples after excluding them from the meta-analysis (i.e., always leaving out the target sample). Among VUMC patients, PRS generated from MGB TRD meta-analysis was significantly associated with both VUMC and MGB TRD phenotypes (VUMC $P=2.15x10^{-5}$, MGB: $P=1.06x10^{-11}$) and VUMC TRD PRS was significantly associated with the MGB TRD phenotype but not the VUMC TRD phenotype (MGB $P=6.01x10^{-4}$, VUMC P=0.0619) (**Table 4**). Among the substantially smaller set of MGB patients, neither TRD PRS was significantly associated with either TRD phenotype.

We next looked at whether PRS derived from relevant psychiatric traits including depression[46], schizophrenia and bipolar disorder associated with our TRD phenotypes in VUMC or MGB patients. We identified that the depression PRS was significantly associated with the VUMC TRD phenotype ($P=4.11x10^{-5}$) and nominally associated with the MGB TRD phenotype ($P=8.96x10^{-3}$). Despite excluding patients with diagnoses of bipolar disorder or schizophrenia defined by at least one diagnostic code, we found that schizophrenia[47] PRS was significantly associated with both the MGB and VUMC TRD phenotypes (**Table 4**, MGB: linear regression $P=6.79x10^{-9}$, VUMC: $P=7.24x10^{-7}$), and bipolar disorder[44] PRS was significantly associated with MGB TRD phenotype ($P=1.99x10^{-4}$) and nominally associated with VUMC TRD phenotype (P=0.036). Among the smaller set of MDD patients in MGB, schizophrenia and bipolar disorder PRS were not significantly associated with either site's TRD phenotype.

		VUMC patients (n = 15,305)			*MGB patients (n = 2,126)*		
Discovery GWAS	**Target trait**	**BETA**	**SE**	**P**	**BETA**	**SE**	**P**
VUMC TRD	VUMC TRD phenotype	0.0159	8.51E-03	0.0619	-0.0026	2.36E-02	0.914
MGB TRD		0.0337	9.83E-03	**6.01E-04**	-0.0182	2.40E-02	0.449
Depression		0.0330	8.05E-03	**4.11E-05**	2.54E-03	2.18E-02	0.907
Bipolar Disorder		0.0172	8.19E-03	0.036	-0.0170	2.21E-02	0.443
Schizophrenia		0.0402	8.10E-03	**7.24E-07**	-0.0051	2.21E-02	0.819
VUMC TRD	MGB TRD phenotype	0.0364	8.56E-03	**2.19E-05**	0.0138	2.38E-02	0.562
MGB TRD		0.0673	9.89E-03	**1.06E-11**	0.0100	2.42E-02	0.679
Depression		0.0212	8.11E-03	8.96E-03	0.0262	2.20E-02	0.233
Bipolar Disorder		0.0307	8.24E-03	**1.99E-04**	0.0282	2.23E-02	0.206
Schizophrenia		0.0473	8.16E-03	**6.79E-09**	1.56E-02	2.23E-02	0.484
VUMC TRD	Medication based TRD	-0.0077	4.17E-02	0.853	0.0061	9.14E-02	0.947
MGB TRD		-0.0765	4.81E-02	0.112	-0.0234	9.34E-02	0.802
Depression		-0.0124	3.97E-02	0.756	0.0282	8.39E-02	0.737
Bipolar Disorder		0.0841	4.02E-02	0.036	0.1047	8.64E-02	0.226
Schizophrenia		0.0719	3.97E-02	0.070	0.0396	8.58E-02	0.644

Table 4: Polygenic risk score association results. PRS were generated using psychiatric traits and TRD meta-analyses as discovery GWAS and applied to TRD phenotypes based on models or medication in the VUMC and MGB patient samples. P-values are bolded if surpassing the multiple-test correction threshold.

Medication-defined TRD corresponds with ECT-based TRD phenotypes

To compare results from the ECT-based TRD phenotypes with a commonly used alternative definition of TRD, we defined case status based on antidepressant medication trial numbers and length. Cases were defined as having at least three unique antidepressants prescribed, requiring the time interval between the third and first antidepressant to be between 16 weeks and 2 years to account for adequate and consecutive trial for each antidepressant. In the VUMC MDD cohort, individuals with medication defined TRD (med-TRD, n = 1181) had higher ECT-based TRD phenotypes than the remaining MDD patients (n = 21,400) for both VUMC TRD (med-TRD: 0.358 ± 1.22, control: 0.015 ± 1.04, t-test $P=1.28x10^{-20}$) and MGB TRD (med-TRD: 0.056 ± 1.21, control: -0.034 ± 1.04, t-test P=0.013). In the MGB MDD cohort (n = 7443), there were no significant differences in the VUMC TRD phenotype or the MGB TRD phenotype between med-TRD patients (n = 501) compared to controls (n = 6942).

We then tested for association of the ECT-based TRD meta-analyses and psychiatric diagnoses PRS with med-TRD status. In both the MGB and VUMC MDD cohorts, neither TRD

meta-analysis PRS was significantly associated with med-TRD (MGB: P = 0.112, VUMC: P = 0.853), and there were no significant associations with PRS of the psychiatric traits.

Significant genetic overlap between TRD and psychiatric, cognitive, substance use and metabolic traits

To study the genetic overlap between TRD and psychiatric and non-psychiatric traits previously associated to TRD, genetic correlations were estimated. Both TRD meta-analyses showed significant positive genetic correlations, after multiple test correction, with cognitive traits including years of education (VUMC: rg = 0.21; MGB: rg = 0.46) and intelligence (VUMC: rg = 0.20; MGB: rg = 0.28), and significant negative genetic correlations with ADHD (VUMC: rg = -0.27; MGB: rg = -0.40), alcohol dependence (VUMC: rg = -0.50; MGB: rg = -0.38) and smoking traits (VUMC: rg = -0.26; MGB: rg = -0.42) (**Figure 4**). Both TRD meta-analyses also showed significant negative genetic correlations with BMI (VUMC: rg = -0.33; MGB: rg = -0.65). While the two TRD meta-analyses shared substantial genetic architecture, there were noticeable difference in genetic correlations across a subset of traits. Traits with significantly stronger genetic correlations in the MGB meta-analysis, based on a block jackknife approach in LD score regression[33] included negative associations with BMI ($P = 3.00x10^{-11}$) and type 2 diabetes ($P = 4.67x10^{-8}$), and positive associations with educational attainment ($P = 4.59x10^{-9}$) and marijuana use ($P = 5.03x10^{-6}$). Traits that had a significantly stronger genetic correlation with the VUMC meta-analysis were neuroticism ($P = 7.24x10^{-5}$), and multiple measures of alcohol use disorders, AUDIT-C ($P = 3.54x10^{-6}$), and AUDIT-T ($P = 3.67x10^{-6}$) (**Figure 4**).

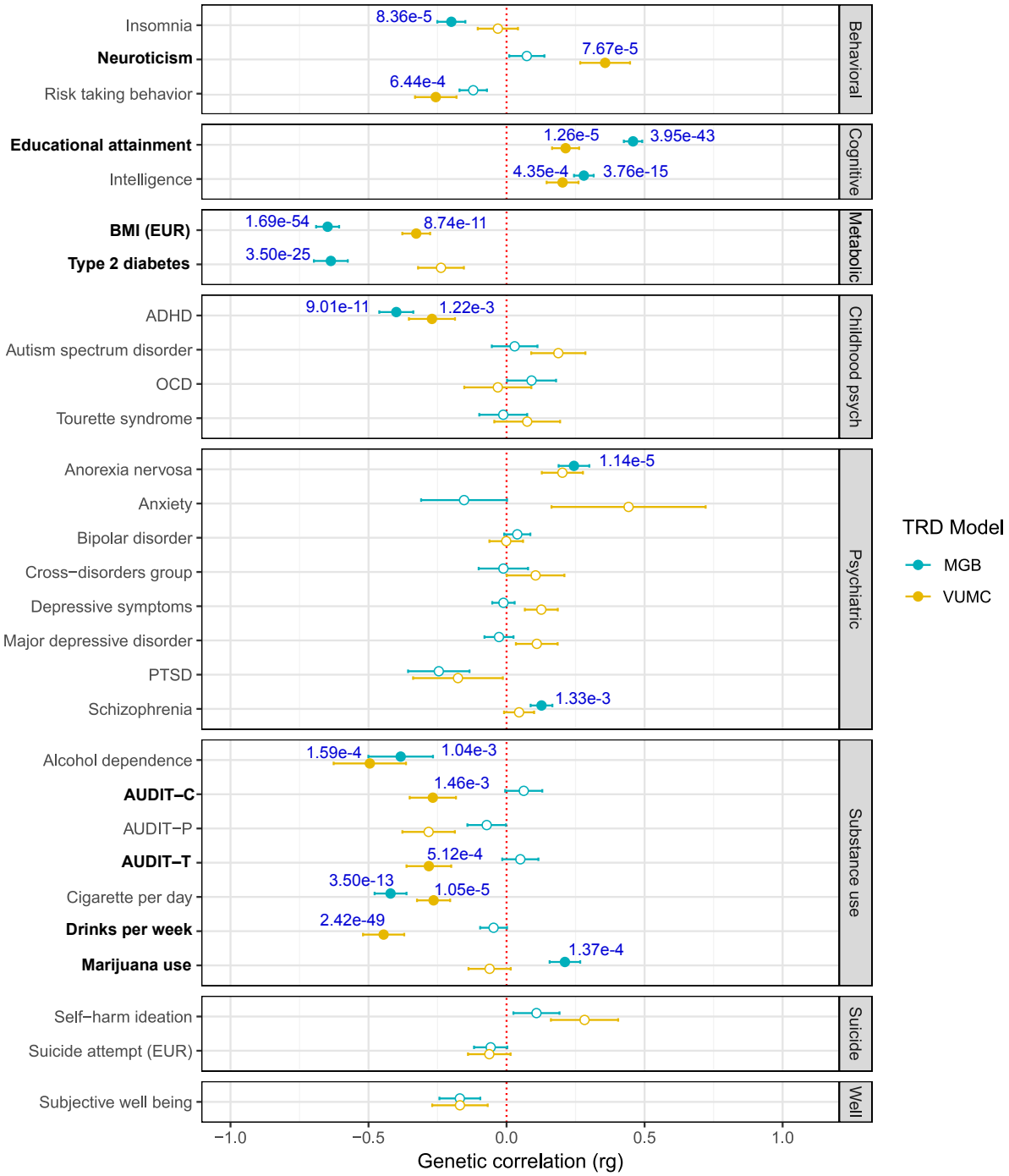

Figure 4: Genetic correlations of VUMC and MGB TRD meta-analyses with psychiatric and non-psychiatric traits. Unfilled points indicate genetic correlations that did not pass the Bonferroni-corrected significance threshold $P<1.72x10^{-3}$ (29 traits tested). Error bars represent the standard error. P values indicate significance after Bonferroni correction. Bolded traits show significant differences in genetic correlations between the two meta-analyses. BMI-body mass index, ADHD-attention-deficit/hyperactivity disorder, OCD-obsessive compulsive disorder, PTSD-post-traumatic stress disorder, AUDIT-C-Alcohol Use Disorders Identification Test-C (measure of quantity of alcohol consumption), AUDIT-P- measure of problematic consequences of drinking, AUDIT-T-total score of AUDIT.

Since we observed strong phenotypic and genetic association between TRD and BMI, we further examined the role that the genetic factors of BMI were having on our results by conditioning our TRD meta-analyses on the genetics architecture of BMI (see Methods). Significant differences in genetic correlations were confined to BMI for both meta-analyses and Type 2 diabetes for MGB meta-analysis only (**Supplementary Figures 3-4**). Among the identified loci, chromosome 16 is a known BMI locus and was used as an instrument SNP for the conditioning analysis so was removed. However, the conditioning on BMI had limited effect on the association statistics from the locus on chromosome 22 in MGB (beta = 0.0194, SE = 0.004, P = 2.02×10^{-7}) and VUMC (beta = 0.0154, SE = 0.004, P = 3.27×10^{-5}) (**Table S5**).

Discussion

In this genetic study of TRD, we found low but significant heritability of TRD, with two novel genome-wide significant loci and significant genetic overlap with schizophrenia, cognitive and substance abuse traits, as well as BMI. Application of a computed phenotype from biobank-linked electronic health records allowed detection of these effects in a total of 154,433 individuals across 4 sites. Understanding the genetic architecture of TRD is important for quantifying the role of genetics in treatment response to move beyond decades-old pharmacogenomic studies. Further, identifying risk loci could facilitate efforts to identify novel treatments considering the modest response rates observed for interventions other than ECT.

While ECT is indicative of TRD, it remains a rare occurrence with a prevalence among individuals with MDD of well below 1%. Even with 154,433 patients, a case-control approach comparing ECT cases to depression controls would have been underpowered with ECT case

numbers of only a couple thousand across all four clinical sites. Leveraging models that can predict ECT from large repositories of clinical data and assign probabilities as quantitative phenotypes allows for substantial increase in power in genetic studies. We showed that our ECT-based prediction models trained both at MGB and VUMC were robust to external validation across three independent healthcare systems. With quantitative phenotypes, we increased power across the entire genotyped cohort of 154,433 patients which resulted in a significant SNP heritability of 2-4% and significant genetic correlation between the two TRD meta-analyses. Both TRD meta-analyses showed significant positive genetic overlap with cognitive traits, and significant negative genetic correlations with ADHD, alcohol and smoking traits, and BMI. We also saw modest evidence of genetic risk of schizophrenia and bipolar disorder even after removing patients with any diagnostic evidence of schizophrenia or bipolar disorder from our population, further implicating potential genetic risk of severe psychiatric disorders to likelihood of TRD. Despite the high genetic correlation across models, genetic overlap with other traits differed significantly representing potential differences in clinical population, general population and/or clinical decision making around ECT. While our quantitative TRD phenotypes were elevated among patients with medication-defined TRD, we did not see any genetic correlation with other genetic studies of TRD. However, the comparable ECT studies used healthy controls and these two studies were highly genetically correlated with each other pointing to the potential that they are predominantly capturing depression genetic architecture as opposed to TRD genetic architecture. Our work shows there is a significant but small contribution of genetics to TRD as defined by ECT. Large studies are currently underway to collect tens of thousands of ECT cases for a case-control study[48] and the comparison to this more timely and efficient approach will be important.

We discovered two genome-wide significant loci both with prior implications to metabolic traits. The most significant locus was in the intergenic region of the obesity and BMI-related *FTO* gene on chromosome 16. The index variant is in perfect linkage disequilibrium with the known BMI variant that regulates *IRX3* and *IRX5*. This locus was not supported by the VUMC meta-analysis, consistent with MGB model having stronger genetic correlation with BMI. Of note, the *inverse* genetic association between TRD and BMI – that is, low BMI being associated with high risk of TRD – is supported by the positive correlation between TRD with anorexia nervosa, a metabopsychiatric illness characterized by extremely low body weights and a notorious lack of antidepressant responsivity in those with comorbid depression. Interestingly, BMI has previously been shown to moderate treatment response in patients with MDD receiving intravenous ketamine, where patients with higher BMI and obesity demonstrated a more robust acute antidepressant response to ketamine[49]. Our study is in line with such findings and extends this to incorporate the metabolic genetic vulnerabilities underlying differential antidepressant treatment responsivity.

The second risk gene *MCHR1*, encodes a G protein-coupled receptor protein linked to neuronal regulation of food intake as well as obesity and insulin resistance in mouse models[50]. This gene is highly expressed in the brain in regions that regulate body weight and appetite[51] and has recently been implicated in bipolar disorder[44], which is an illness associated with metabolic traits and severe depressive episodes. Moreover, inhibition of the *MCH1* signaling pathway results in anti-depressant and anorectic effects in murine models[52], which is in line with our findings that implicate *MCHR1* in depression treatment responsivity phenotype along with association to lower BMI. This locus shares substantial support from the VUMC meta-analysis, and the effect size is largely constant even after conditioning on BMI. Together, our identification of *FTO* and *MCHR1* as genetic vulnerabilities for TRD supports a direct link between depression treatment responsivity

and the complex metabolic regulatory pathways underlying energy balance including food intake and body weight homeostasis.

We note several limitations of our study, particularly the potential confounding of ECT population characteristics in our TRD clinical models. There were significant demographic differences between cases and controls where a typical ECT case tended to be an older, white, male with a lower mean BMI compared to MDD controls. These demographic differences could be driven by ascertainment in medical decisions leading to a patient receiving ECT or socioeconomic factors like access to a caregiver as patients need accompaniment after the inpatient ECT procedure. Demographic differences between the region around Nashville and eastern Massachusetts could also contribute to the differences observed in the VUMC and MGB TRD phenotypes, but the two meta-analyses showed significant genetic overlap (rg = 0.66), and both prediction models performed robustly in independent clinical sites with different demographics, especially in the Million Veteran Program cohort, which is significantly more male and older than the other cohorts. Phenotypes based on prediction models are not always representative of the original phenotype and could differ in important ways that modify genetic architecture and power. We were able to identify significant but low SNP-heritability (2-4%), meaning that even with our substantial improvements in power many more patients will be required to enable identification of additional genome-wide significant loci. Given such low genetic contribution, an important question is whether ECT represents a generalizable form of TRD such that genetic contribution to TRD broadly is likely as low or whether there is a more biologically homogenous form of TRD. We note that previous estimate of SNP-heritability of TRD within MDD patients using prescription data was only slightly higher at 8% but with a wider confidence interval (SE = 0.04)[11].

Despite these limitations, this study supports the utility of investigating a proxy for TRD that can be readily extracted and predicted from electronic health records or administrative claims data. We confirm a significant but modest genetic contribution to TRD and provide insights into its overlap with other psychiatric and non-psychiatric phenotypes, in particular metabolic traits. This effort lays the groundwork for future efforts to apply genomic data for biomarker and drug development in TRD.

Acknowledgements

This work was funded by National Institutes of Mental Health grants R01MH116269 (DMR, CGW), R01MH121455 (DMR, CGW), R01MH116270 (DMR, CGW, RHP) and R01MH123804 (RHP), and National Institute of General Medical Sciences grant 5T32GM007347 (JK). The dataset(s) used for the analyses described were obtained from Vanderbilt University Medical Center's BioVU, which is supported by institutional funding, private agencies and federal grants. These include the NIH-funded Shared Instrumentation Grant S10RR025141, and CTSA grants UL1TR002243, UL1TR000445 and UL1RR024975. Genomic data are also supported by investigator-led projects that include U01HG004798, R01NS032830, RC2GM092618, P50GM115305, U01HG006378, U19HL065962 and R01HD074711; and additional funding sources listed at https://victr.vanderbilt.edu/pub/biovu/. We thank Mass General Brigham Biobank for providing samples, genomic data, and health information data. This work was conducted in part using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University, Nashville, TN.

This research is also based on data from the Million Veteran Program, Office of Research and Development, Veterans Health Administration, and was supported by award #MVP006. This publication does not represent the views of the Department of Veteran Affairs or the United States

Government. This study was also supported by the National Institutes of Health (NIH), Bethesda, MD under award numbers K08MH122911 (GV), R01MH125246, R01AG067025, U01MH116442, R01MH109677 (PR), and by the Veterans Affairs Merit grants BX002395 and BX004189 (PR). This study has also been funded in part by the Brain & Behavior Research Foundation via the 2020 NARSAD Young Investigator Grant #29350 (GV).

References

1. Kessler, R. C. & Bromet, E. J. The epidemiology of depression across cultures. *Annu. Rev. Public Health* **34**, 119–138 (2013).
2. Andrade, L. *et al.* The epidemiology of major depressive episodes: results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys. *Int. J. Methods Psychiatr. Res.* **12**, 3–21 (2003).
3. Rush, A. J. *et al.* Acute and Longer-Term Outcomes in Depressed Outpatients Requiring One or Several Treatment Steps: A STAR*D Report. *Am. J. Psychiatry* **163**, 1905–1917 (2006).
4. Zhdanava, M. *et al.* The Prevalence and National Burden of Treatment-Resistant Depression and Major Depressive Disorder in the United States. *J. Clin. Psychiatry* **82**, (2021).
5. Hantouche, E., Angst, J. & Azorin, J.-M. Explained factors of suicide attempts in major depression. *J. Affect. Disord.* **127**, 305–308 (2010).
6. Bergfeld, I. O. *et al.* Treatment-resistant depression and suicidality. *J. Affect. Disord.* **235**, 362–367 (2018).
7. Perlis, R. H. & Fava, M. Is It Time to Try Sequenced Treatment Alternatives to Relieve Depression (STAR*D) Again? *JAMA Psychiatry* **79**, 281–282 (2022).

8. Lage, I., McCoy Jr, T. H., Perlis, R. H. & Doshi-Velez, F. Efficiently identifying individuals at high risk for treatment resistance in major depressive disorder using electronic health records. *J. Affect. Disord.* (2022) doi:10.1016/j.jad.2022.02.046.
9. Fabbri, C. *et al.* The Genetics of Treatment-Resistant Depression: A Critical Review and Future Perspectives. *Int. J. Neuropsychopharmacol.* **22**, 93–104 (2019).
10. Li, Q. S., Tian, C., Seabrook, G. R., Drevets, W. C. & Narayan, V. A. Analysis of 23andMe antidepressant efficacy survey data: implication of circadian rhythm and neuroplasticity in bupropion response. *Transl. Psychiatry* **6**, e889–e889 (2016).
11. Fabbri, C. *et al.* Genetic and clinical characteristics of treatment-resistant depression using primary care records in two UK cohorts. *Mol. Psychiatry* (2021) doi:10.1038/s41380-021-01062-9.
12. Brown, S. *et al.* Current and Common Definitions of Treatment-Resistant Depression: Findings from a Systematic Review and Qualitative Interviews. *Can. J. Psychiatry* **64**, 380–387 (2019).
13. Fekadu, A., Donocik, J. G. & Cleare, A. J. Standardisation framework for the Maudsley staging method for treatment resistance in depression. *BMC Psychiatry* **18**, 100 (2018).
14. Howes, O. D., Thase, M. E. & Pillinger, T. Treatment resistance in psychiatry: state of the art and new directions. *Mol. Psychiatry* **27**, 58–72 (2022).
15. Conway, C. R., George, M. S. & Sackeim, H. A. Toward an Evidence-Based, Operational Definition of Treatment-Resistant Depression: When Enough Is Enough. *JAMA Psychiatry* **74**, 9–10 (2017).
16. Gaynes, B. N. *et al.* Defining treatment-resistant depression. *Depress. Anxiety* **37**, 134–145 (2020).

17. Ruderfer, D. M. *et al.* Significant shared heritability underlies suicide attempt and clinically predicted probability of attempting suicide. *Mol. Psychiatry* 1 (2019) doi:10.1038/s41380-018-0326-8.
18. Lisanby, S. H. Electroconvulsive Therapy for Depression. *N. Engl. J. Med.* **357**, 1939–1945 (2007).
19. Nalichowski, R., Keogh, D., Chueh, H. C. & Murphy, S. N. Calculating the benefits of a Research Patient Data Repository. *AMIA Annu. Symp. Proc. AMIA Symp.* **2006**, 1044 (2006).
20. Roden, D. M. *et al.* Development of a large-scale de-identified DNA biobank to enable personalized medicine. *Clin. Pharmacol. Ther.* **84**, 362–369 (2008).
21. Carey, D. J. *et al.* The Geisinger MyCode Community Health Initiative: an electronic health record-linked biobank for Precision Medicine research. *Genet. Med. Off. J. Am. Coll. Med. Genet.* **18**, 906–913 (2016).
22. Gaziano, J. M. *et al.* Million Veteran Program: A mega-biobank to study genetic influences on health and disease. *J. Clin. Epidemiol.* **70**, 214–223 (2016).
23. Elixhauser, A., Steiner, C. & Palmer, L. Clinical classifications software. *US Agency Healthc. Res. Qual. URL Httpwww Hcup-Us Ahrq Govtoolssoftwareccsccs Jsp* (2015).
24. Brokamp, C. *et al.* Material community deprivation and hospital utilization during the first year of life: an urban population-based cohort study. *Ann. Epidemiol.* **30**, 37–43 (2019).
25. Singh, G. K. Area Deprivation and Widening Inequalities in US Mortality, 1969–1998. *Am. J. Public Health* **93**, 1137–1143 (2003).
26. Tibshirani, R. Regression Shrinkage and Selection Via the Lasso. *J. R. Stat. Soc. Ser. B Methodol.* **58**, 267–288 (1996).

27. Friedman, J. H., Hastie, T. & Tibshirani, R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J. Stat. Softw.* **33**, 1–22 (2010).
28. Denny, J. C. *et al.* Systematic comparison of phenome-wide association study of electronic medical record data and genome-wide association study data. *Nat. Biotechnol.* **31**, 1102–1111 (2013).
29. Carroll, R. J., Bastarache, L. & Denny, J. C. R PheWAS: data analysis and plotting tools for phenome-wide association studies in the R environment. *Bioinforma. Oxf. Engl.* **30**, 2375–2376 (2014).
30. Willer, C. J., Li, Y. & Abecasis, G. R. METAL: fast and efficient meta-analysis of genomewide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
31. Bulik-Sullivan, B. K. *et al.* LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
32. Clements, C. C. *et al.* Genome-wide association study of patients with a severe major depressive episode treated with electroconvulsive therapy. *Mol. Psychiatry* (2021) doi:10.1038/s41380-020-00984-0.
33. Hübel, C. *et al.* Genomics of body fat percentage may contribute to sex bias in anorexia nervosa. *Am. J. Med. Genet. B Neuropsychiatr. Genet.* **180**, 428–438 (2019).
34. Zhu, Z. *et al.* Causal associations between risk factors and common diseases inferred from GWAS summary data. *Nat. Commun.* **9**, 224 (2018).
35. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: A Tool for Genome-wide Complex Trait Analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
36. Pulit, S. L. *et al.* Meta-analysis of genome-wide association studies for body fat distribution in 694 649 individuals of European ancestry. *Hum. Mol. Genet.* **28**, 166–174 (2019).

37. Auton, A. *et al.* A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).

38. Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A. & Smoller, J. W. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat. Commun. 2019 101* **10**, 1–10 (2019).

39. Wilkinson, S. T., Agbese, E., Leslie, D. L. & Rosenheck, R. A. Identifying recipients of electroconvulsive therapy: Data from privately insured americans. *Psychiatr. Serv.* **69**, 542–548 (2018).

40. Li, Q. S., Tian, C. & Hinds, D. Genome-wide association studies of antidepressant class response and treatment-resistant depression. *Transl. Psychiatry* **10**, 1–12 (2020).

41. Church, C. *et al.* Overexpression of Fto leads to increased food intake and results in obesity. *Nat. Genet.* **42**, 1086–92 (2010).

42. Smemo, S. *et al.* Obesity-associated variants within FTO form long-range functional connections with IRX3. *Nature* **507**, 371–5 (2014).

43. Ochoa, D. *et al.* Open Targets Platform: supporting systematic drug–target identification and prioritisation. *Nucleic Acids Res.* **49**, D1302–D1310 (2021).

44. Mullins, N. *et al.* Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat. Genet.* **53**, 817–829 (2021).

45. The International Schizophrenia Consortium. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* **460**, 748–752 (2009).

46. Howard, D. M. *et al.* Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat. Neurosci.* **22**, 343–352 (2019).

47. Consortium, T. S. W. G. of the P. G., Ripke, S., Walters, J. T. & O'Donovan, M. C. Mapping genomic loci prioritises genes and implicates synaptic biology in schizophrenia. *medRxiv* 2020.09.12.20192922 (2020) doi:10.1101/2020.09.12.20192922.
48. Soda, T. *et al.* International Consortium on the Genetics of Electroconvulsive Therapy and Severe Depressive Disorders (Gen-ECT-ic). *Eur. Arch. Psychiatry Clin. Neurosci.* **270**, 921–932 (2020).
49. Freeman, M. P. *et al.* Body Mass Index as a Moderator of Treatment Response to Ketamine for Major Depressive Disorder. *J. Clin. Psychopharmacol.* **40**, 287–292 (2020).
50. Ludwig, D. S. *et al.* Melanin-concentrating hormone overexpression in transgenic mice leads to obesity and insulin resistance. *J. Clin. Invest.* **107**, 379–386 (2001).
51. Al-Massadi, O. *et al.* Multifaceted actions of melanin-concentrating hormone on mammalian energy homeostasis. *Nat. Rev. Endocrinol.* **17**, 745–755 (2021).
52. Borowsky, B. *et al.* Antidepressant, anxiolytic and anorectic effects of a melanin-concentrating hormone-1 receptor antagonist. *Nat. Med.* **8**, 825–830 (2002).