Use of the PsycheMERGE Network to Investigate the Association Between Depression Polygenic Scores and White Blood Cell Count

Julia M. Sealock, BSc; Younga H. Lee, PhD; Arden Moscati, PhD; Sanan Venkatesh, BSc; Georgios Voloudakis, MD, PhD; Peter Straub, MS; Kritika Singh, BSc; Yen-Chen A. Feng, PhD; Tian Ge, PhD; Panos Roussos, MD, PhD; Jordan W. Smoller, MD, ScD; Guanhua Chen, PhD; Lea K. Davis, PhD

IMPORTANCE Although depression is a common psychiatric disorder, its underlying biological basis remains poorly understood. Pairing depression polygenic scores with the results of clinical laboratory tests can reveal biological processes involved in depression etiology and in the physiological changes resulting from depression.

OBJECTIVE To characterize the association between depression polygenic scores and an inflammatory biomarker, ie, white blood cell count.

DESIGN, SETTING, AND PARTICIPANTS This genetic association study was conducted from May 19, 2019, to June 5, 2021, using electronic health record data from 382,452 patients across 4 health care systems. Analyses were conducted separately in each health care system and meta-analyzed across all systems. Primary analyses were conducted in Vanderbilt University Medical Center’s biobank. Replication analyses were conducted across 3 other PsycheMERGE sites: Icahn School of Medicine at Mount Sinai, Mass General Brigham, and the Million Veteran Program. All patients with available genetic data and recorded white blood cell count measurements were included in the analyses. Primary analyses were conducted in individuals of European descent and then repeated in a population of individuals of African descent.

EXPOSURES Depression polygenic scores.

MAIN OUTCOMES AND MEASURES White blood cell count.

RESULTS Across the 4 PsycheMERGE sites, there were 382,452 total participants of European ancestry (18.7% female; median age, 57.9 years) and 12,383 participants of African ancestry (61.1% female; median age, 39.0 [range, birth-90.0 years]). A laboratory-wide association scan revealed a robust association between depression polygenic scores and white blood cell count (β, 0.03; SE, 0.004; \( P = 1.07 \times 10^{-17} \)), which was replicated in a meta-analysis across the 4 health care systems (β, 0.03; SE, 0.002; \( P = 1.03 \times 10^{-136} \)). Mediation analyses suggested a bidirectional association, with white blood cell count accounting for 2.5% of the association of depression polygenic score with depression diagnosis (95% CI, 2.2%-20.8%; \( P = 2.84 \times 10^{-70} \)) and depression diagnosis accounting for 9.8% of the association of depression polygenic score with white blood cell count (95% CI, 8.4%-11.1%; \( P = 1.78 \times 10^{-44} \)). Mendelian randomization provided additional support for an association between increased white blood count and depression risk, but depression modeled as the exposure showed no evidence of an influence on white blood cell counts.

CONCLUSIONS AND RELEVANCE This genetic association study found that increased depression polygenic scores were associated with increased white blood cell count, and suggests that this association may be bidirectional. These findings highlight the potential importance of the immune system in the etiology of depression and may motivate future development of clinical biomarkers and targeted treatment options for depression.
Depression is a common psychiatric disorder estimated to affect 264 million individuals worldwide. Diagnostic criteria for depression include clinical evaluation of self-reported psychiatric symptoms, such as depressed mood, irritability, anhedonia, or suicidal thoughts. In addition to psychiatric effects, depression is associated with increased risk for cardiovascular disease, autoimmune diseases, and diabetes. The increased risk of peripheral diseases suggests the biology of depression is not limited to the brain; nevertheless, the causes and biological effects of depression in the brain and the periphery remain poorly understood.

In a health care setting, laboratory tests aid clinicians in diagnostic and treatment decision-making. Tests that can accurately and reproducibly indicate a medical state are generally referred to as biomarkers. To date, there are no biomarkers for depression; however, consistent with the high number of common comorbidities, depression is associated with changes in a wide range of clinical laboratory values, including increased proinflammatory cytokines, altered lipids, growth factors, and decreased brain-derived neurotrophic factor. For many of these physiological quantitative values, the underlying biological mechanisms are well understood. Further understanding of the biological link between clinical depression and these laboratory values can help identify the biological processes contributing to depression and could lead to the development of more informative biomarker panels to be used in risk assessment and treatment response.

Previous studies report a bidirectional association between depression and autoimmune disease. Several immune biomarkers are increased in patients with depression compared with controls, including monocytes, neutrophils, lymphocyte ratio, and C-reactive protein. However, most immune biomarker studies of depression are limited in sample size and scope and are often unable to control for potential confounders or determine the pathway between depression and biomarkers.

Electronic health records (EHRs) store longitudinal information about the health and clinical care of individual patients. Biobanks that link EHRs to DNA information provide an opportunity to analyze clinical information along with genetic risk factors. Genetic risk for depression can be estimated using polygenic scores (PGS), which aggregate the relevant contribution to depression. In this work, we use recently developed methods to combine depression PGS with laboratory results stored in EHRs to robustly identify physiological processes affected by increased genetic liability to depression.

Whereas independent biobanks can be used to discover associations, combining multiple health record systems through consortia can validate those discoveries in broader populations. The PsycheMERGE Network consists of investigators from institutions across the US with the common goal of using EHRs and biobanks to advance the identification, biology, and treatment of psychiatric disorders. Here, we investigate the effect of polygenic risk for depression on clinically measured laboratory values leveraging data from 4 health care systems participating in the PsycheMERGE Network.

### Methods

#### Sample Description

Electronic health record and genotype information was extracted for individuals of European descent across 4 biobanks in the PsycheMERGE Network: Vanderbilt University Medical Center (VUMC), Mass General Brigham (MGB), Million Veteran Program (MVP), and Mount Sinai Icahn School of Medicine (MSSM). Biobank-specific information can be found in the eMethods in Supplement 1. All participants provided informed consent, and study procedures were approved by each institution’s respective institutional review board. This genetic association study was conducted from May 19, 2019, to June 5, 2021, and followed the Strengthening the Reporting of Genetic Association Studies (STREGA) reporting guideline.

#### Statistical Analyses

**Depression Polygenic Scoring**

Depression PGS were generated using Polygenic Risk Score–Continuous Shrinkage with single nucleotide variation (SNV) weights from the largest available depression meta-analysis. The linkage disequilibrium reference panel was constructed from 503 European samples in the 1000 Genomes Project phase 3. Polygenic scores were scaled to have a mean of 0 and a unit SD so that effect size estimates in subsequent analyses were interpreted per 1 SD increase in depression PGS. In VUMC data, the depression PGS explained 0.8% of the variance in major depressive disorder diagnosis ($P = 3.85 \times 10^{-29}$).

**Laboratory-wide Association Scan of Depression PGS in VUMC**

At VUMC, all laboratory results were extracted from the EHRs of 72 634 individuals of primarily European ancestry and 12 383...
individuals of primarily African ancestry and cleaned as previously described (eMethods in Supplement 1). Associations between the depression PGS and laboratory results were estimated with a laboratory-wide association scan (LabWAS) approach \(^3\) controlled for sex and the top 10 genetic principal components (eTable 1 in Supplement 2). In conditional analyses, the LabWAS of depression PGS was covaried for body mass index (median across each individual’s EHR) and for depression (eTable 2 in Supplement 3), anxiety (eTable 3 in Supplement 4), adjustment reaction (eTable 4 in Supplement 5 and eTable 5 in Supplement 6), and tobacco use disorder (as a proxy for smoking status) (eTable 6 in Supplement 7) diagnoses, defined by phecodes 296.2, 300.1, 304, and 318, respectively (eFigure 1 in Supplement 1). In sensitivity analyses focused on smoking, we obtained smoking data from the social history index (median across each individual's EHR) and for depression (outcome) with the remainingsubtypes as covariates in the LabWAS of depression PGS (eFigure 1 in Supplement 1; eTable 7 in Supplement 8). Genetic analyses must be conducted in ancestry-stratified samples to avoid confounding due to population stratification. Using principal component analysis, we determined the genetic ancestry of individuals in our sample. \(^3\) The largest continental ancestry group was European; therefore, we restricted the study participants to individuals of European ancestry in our primary analyses. The second largest ancestry was African. We conducted an initial LabWAS in the African sample; however, due to the small sample size, there were no significant associations. As a result, we did not continue with the sensitivity analyses in this sample. Primary analyses were restricted to individuals of European descent and repeated in individuals of African ancestry (n = 12,383) in VUMC only (eFigure 4 in Supplement 1; eTable 8 in Supplement 9).

Sensitivity Analyses in VUMC

A series of conditional and sensitivity analyses were performed to ensure the association between depression PGS and white blood cell (WBC) count was not due to a common comorbid confounder phenotype present in individuals with both an increased depression PGS and an increased WBC count. To find phenotypes associated with both depression PGS and WBC count, separate phenome-wide association scans were conducted of depression PGS and of the median, age-adjusted, inverse normal transformed WBC measurement (eMethods in Supplement 1). Next, phenotypes that were significantly associated with both depression PGS and WBC count at Bonferroni significance (WBC \(P < 3.64 \times 10^{-5}\), depression PGS \(P < 3.72 \times 10^{-5}\)) were selected and binned into 7 categories based on phenotypic similarity. Group-based case-control variables were constructed, in which an individual was considered a case if they were a case for any of the group's phecodes. Controls were required to be a control for all phecodes. To assess the effect of the comorbid phenotypes on the association between depression PGS and WBC, a series of linear regression analyses were conducted controlling for each of the groups separately and all common phenotype groups together. All analyses were controlled for sex, top 10 genetic principal components, and median age across the medical record. We also conducted a series of sensitivity analyses controlling for the impact of WBC genetics on the association between depression PGS and WBC count. Details can be found in the eResults, eFigure 5, and eTable 16 in Supplement 1.

Replication in the PsycheMERGE Network

Targeted replication analyses focused on depression PGS and WBC count were conducted in 3 external biobanks. Depression PGS were constructed and WBC quality was controlled for as in the VUMC biobank. The depression PGS and WBC counts were fitted in a linear regression model controlling for sex and top 10 genetic principal components. The associations controlling for depression and anxiety diagnoses were also replicated using the same phenotype definition as described in the discovery LabWAS at VUMC. The effect estimates from each analysis were meta-analyzed across all 4 sites using a fixed-effect inverse variance weighted model in the meta \(^4\) R package (R Core Team).

Depression PGS and WBC Mediation Analysis

Two mediation models were investigated using the mediation \(^4\) R package. First, WBC count was modeled as the mediator between depression PGS (exposure) and depression diagnosis (outcome). Second, depression diagnosis was modeled as the mediator between depression PGS (exposure) and WBC count (outcome). Details can be found in the eMethods in Supplement 1. The proportion-mediated estimates from all 4 sites were meta-analyzed using a fixed-effect inverse variance weighted model in the meta \(^4\) R package. Owing to the uniqueness of MVP (ie, combat exposed, primarily men) compared with the other sites, we also conducted meta-analyses excluding MVP (eTables 12 and 13 in Supplement 1).

Depression PGS and WBC-Differential Mediation Analysis

To determine which WBC subtypes contributed to the association between depression PGS and depression diagnosis, a series of multiple mediator analyses were conducted using the mediation \(^4\) R package. Each WBC subtype was analyzed as the main mediator between depression PGS (exposure) and depression diagnosis (outcome) with the remaining subtypes as the alternative mediators. Details can be found in the eMethods in Supplement 1.

Mendelian Randomization

We conducted bidirectional mendelian randomization between depression and WBC count using generalized summary-based mendelian randomization \(^4\) in the GCTA, or Genome-wide Complex Trait Analysis, package, version 1.92.4. \(^4\) Index SNVs were selected using the default settings in GCTA: P value threshold of \(5 \times 10^{-6}\), linkage disequilibrium \(r^2\) clumping threshold of 0.05, and a HEIDI-outlier threshold of 0.01 to remove SNVs that have pleiotropic effects on both risk factor and disease. From the depression \(^5\) and WBC \(^6\) summary statistics, 47 and 203 SNVs were selected as index SNVs, respectively. Phenome-wide association scans, LabWAS, and conditional, replication, and mediation analyses were conducted using R, version 3.4.3 (R Foundation). Mendelian randomization was
Conducted using GCTA. The code for each analysis can be found online.47

Results
LabWAS of Depression PGS
Across the 4 PsycheMERGE sites, there were 382,452 participants of European ancestry (18.7% female and 81.3% male; median age, 57.9 years). An additional 12,383 participants of primarily African ancestry (61.1% female; median age, 39.0 [range, birth-90.0 years]) were included from VUMC.

Depression PGS were screened for associations with 315 clinical laboratory measurements using a LabWAS in VUMC’s biobank (N = 72,634). After multiple testing correction, the LabWAS of depression PGS revealed significant associations with 4 elevated immune markers: WBC (P = 1.07 × 10^{-17}; β, 0.03; SE, 0.004), urinary WBC (P = 1.45 × 10^{-5}; β, 0.03; SE, 0.007), absolute monocyte count (P = 2.54 × 10^{-5}; β, 0.02; SE, 0.005), and absolute neutrophil count (P = 5.91 × 10^{-5}; β, 0.02; SE, 0.005). Significant associations also included several metabolic markers, including increased triglycerides (P = 3.14 × 10^{-18}; β, 0.05; SE, 0.006), decreased high-density lipoprotein cholesterol (P = 1.23 × 10^{-11}; β, −0.04; SE, 0.005), decreased calcitriol (P = 2.83 × 10^{-8}; β, 0.03; SE, 0.007), increased glucose (P = 2.84 × 10^{-5}; β, 0.02; SE, 0.004), decreased blood urea nitrogen (P = 5.19 × 10^{-7}; β, 0.02; SE, 0.004), decreased calcium (P = 9.74 × 10^{-7}; β, 0.02; SE, 0.004), and decreased calcitriol (P = 7.03 × 10^{-5}; β, 0.04; SE, 0.01). Depression PGS were also associated with decreased troponin I (P = 1.99 × 10^{-6}; β, 0.009; SE, 0.009), decreased urinary red blood cells (P = 1.37 × 10^{-5}; β, −0.03; SE, 0.006), decreased thyroxine (P = 1.72 × 10^{-5}; β, −0.03; SE, 0.006), and decreased blood carbon dioxide (P = 4.06 × 10^{-6}; β, −0.02; SE, 0.003) (Figure 1A; eTable 1 in Supplement 2).

In a conditional analysis, we sequentially controlled for diagnoses for depression, anxiety, adjustment reaction, and tobacco use disorder and for median body mass index across the EHR. In the analysis with all covariates, the most significant association remained WBC count (P = 1.11 × 10^{-16}; β, 0.03; SE, 0.005), followed by triglycerides (P = 1.91 × 10^{-5}; β, 0.04; SE, 0.008) (Figure 1B; eTables 2-6 in Supplements 3-7, respectively, and eFigures 1-2 in Supplement 1).

Although depression PGS remained robustly associated with WBC across all analyses, the magnitude of the association was modest (β, 0.03; SE, 0.004). Stratification of individuals in the discovery cohort (VUMC) showed that even at the highest decile of depression PGS, WBC measurements were elevated but remained within the clinical reference range (ie, 4-11 thousand cells/μL) (eFigure 3 in Supplement 1).

No laboratory results were significantly associated in the LabWAS of depression PGS in individuals of African descent, likely owing to the smaller sample size of the African ancestry sample (n = 12,383) and the low generalizability of PGS built using European summary statistics in African populations.48 However, the association with WBC count was in the same direction as in the European sample (P = .06; β, 0.02; SE, 0.01) (eFigure 4 in Supplement 1; eTable 8 in Supplement 9).

Conditional Analyses of WBC
In separate phenotype-wide association scans, depression PGS and median WBC count were significantly associated with 66 and 469 phenocodes, respectively. Of these significantly associated phenocodes, 32 were common to both depression PGS and median WBC count and were binned into 7 categories based on phenotypic similarity: cardiovascular, psychiatric, obesity, respiratory, hepatic, pain, and autoimmune conditions (Figure 2; eTable 9 in Supplement 1).

The association between depression PGS and WBC count remained significant after controlling for each group separately and controlling for all phenotype groups together (P = 4.19 × 10^{-5}; β, 0.02; SE, 0.008) with effect estimates similar to the original association despite the reduced sample size (N = 13,269) (Figure 2; eTable 10 in Supplement 1).

Replication in the PsycheMERGE Network
Given the robustness of the association with WBC count and the history of associations between depression status and proinflammatory markers, we focused on WBC count for replication and further investigation. Findings were replicated in the 3 external biobanks (MVP, MSSM, and MGB) (Table 1). In both MVP (N = 289,880) and MGB (N = 20,828), the association between depression PGS and WBC remained significant with effect estimates replicating those observed at VUMC (Figure 3). In MSSM, the effect size point estimate was similar to those observed in the 3 other sites, but did not reach statistical significance, probably due to the smaller sample size (n = 823). The meta-analyzed effect estimate from the 4 sites was robust and significant (P = 1.03 × 10^{-13}; β, 0.03; SE, 0.002), even after controlling for depression diagnosis (P = 9.52 × 10^{-12}; β, 0.03; SE, 0.002) and after controlling for depression and anxiety diagnoses (P = 8.23 × 10^{-10}; β, 0.03; SE, 0.002) (Figure 3; eTable 11 in Supplement 1).

Mediation Analysis
Two potential pathways between depression PGS, WBC count, and depression diagnosis were assessed using mediation analyses. In the first analysis, median WBC count was modeled as a mediator of the relationship between depression PGS (exposure) and depression diagnosis (outcome). Meta-analysis across all sites revealed that WBC count mediated 2.5% of the association between depression PGS and depression diagnosis (95% CI, 2.2%–20.8%; Table 2; eTable 12 in Supplement 1). When excluding MVP from the meta-analysis, WBC count mediated 0.5% of the association, although this was not a statistically significant association (95% CI, 0.03% to 0.9%; P = .06).

In the second analysis, depression diagnosis was modeled as a mediator of the association between the depression PGS (exposure) and median WBC count (outcome). Meta-analysis across all sites indicated that depression diagnosis mediated 9.8% of the association between depression PGS and WBC count (95% CI, 8.4%–11.3%; P = 1.78 × 10^{-4}) (Table 2; eTable 13 in Supplement 1). A depression diagnosis mediated 1.4% of the association when excluding MVP from the meta-analysis (95% CI, 0.6% to 3.4%; P = .17).
White bloodcellcountsarecalculatedfromthesumof5dif-
erentcellsubtypes:neutrophils,lymphocytes,monocytes,bas-
sphils, and eosinophils. These cell subtypes can be mea-
suredalongwiththetotalWBCcountusingacompleteblood
countdifferentiallaboratorypanel. Todeterminewhetherspe-
cific WBC components accounted for the associations be-
tween depression PGS and depression diagnosis, we per-
formedaseriesofmultiplemediatoranalyses.

When depression PGS was modeled as the exposure and
depressiondiagnosisastheoutcome,neutrophilswereetheonly
celltypethatexplainedasignificantproportion(1.9%;95%CI,
0.2%-3.1%) of the association between depression PGS and
depressiondiagnosis(eTable14inSupplement1).

WhenmodelingWBCcountastheexposureanddepressionas
theoutcome,mendelianrandomizationanalysisprovidedad-
ditionalevidenceforanincreaseindepressionriskwithanin-
creaseinWBC(\(P= .01\);estimatedeffectoftheexposureonthe
outcome,0.27)(eFigure6;eTable15inSupplement1).However,
depressionmodeledastheexposureshowednoevidenceofa
potential causal influence on the WBC outcome (\(P= .30\); esti-
matedeffectoftheexposureontheoutcome,0.022).

Depression PGS and WBC-Differential
Mediation Analysis
White blood cell counts are calculated from the sum of 5 dif-
erent cell subtypes: neutrophils, lymphocytes, monocytes, bas-
sphils, and eosinophils. These cell subtypes can be mea-
sured along with the total WBC count using a complete blood
count differential laboratory panel. To determine whether spe-
cific WBC components accounted for the associations be-
tween depression PGS and depression diagnosis, we per-
formed a series of multiple mediator analyses.

When depression PGS was modeled as the exposure and
depression diagnosis as the outcome, neutrophils were the only
cell type that explained a significant proportion (1.9%; 95% CI,
0.2%-3.1%) of the association between depression PGS and
depression diagnosis (eTable 14 in Supplement 1).

Mendelian Randomization
When modeling WBC count as the exposure and depression as
the outcome, mendelian randomization analysis provided ad-
ditional evidence for an increase in depression risk with an in-
crease in WBC (\(P= .01\); estimated effect of the exposure on the
outcome, 0.27) (eFigure 6; eTable 15 in Supplement 1). However,
depression modeled as the exposure showed no evidence of a
potential causal influence on the WBC outcome (\(P= .30\); esti-
mated effect of the exposure on the outcome, 0.022).

Discussion
Depression is consistently associated with increased proinflam-
ATORY biomarkers; however, the mechanisms underlying these
associations remain unclear. In this genetic association study, analysis of EHR-linked biobanks within the PsycheMERGE Network were used to examine the association between depression PGS and a variety of clinical laboratory traits, revealing a robustly replicated association with increased WBC count. Notably, several other laboratory traits were associated with depression PGS, including lipids, blood glucose, and blood urea nitrogen. The variety of associations with depression PGS suggest that multiple areas of biology are affected by depression genetics, including metabolism and inflammation. We chose to further investigate the association with WBC count given the existing literature and the robustness of the observed association with clinical confounders.

In a laboratory-wide screen, increased polygenic depression risk was associated with increased inflammatory markers, including WBC count, even after controlling for depression, anxiety, multiple comorbid phenotypes, body mass index, and smoking, thus suggesting that depression PGS was an important risk factor for the proinflammatory state observed in depression. These results suggested that genetic risk for depression, independent of depressive symptoms, was linked to a proinflammatory biomarker. The association of the depression PGS with WBC was modest across all biobanks, suggesting that individuals with high depression genetic liability may have an activated but not abnormal immune system. Nonetheless, sustained activation of the immune system could have important implications for the risk of developing depression.

Here are 2 main models that connect depression to a proinflammatory state: the neuroinflammation model and the stress response model. The neuroinflammation model hypothesizes that an activated immune system contributes to risk of developing depression. The stress response model proposes the stress of depression symptoms leads to a proinflammatory state. Importantly, these 2 models are not mutually exclusive, and some have suggested they form a feedback loop.

In support of this hypothesis, our mediation results do not distinguish either the neuroinflammation model or the stress response model as the exclusive pathway between depression and WBC count. However, mendelian randomization results supported a potential causal path from increased WBC levels to increased depression risk, consistent with the neuroinflammation model; but did not support a model of depression leading to increased WBC levels. It is important to note that only 47 SNVs met criteria to be included as depression instrument variables, limiting the statistical power of the analysis.

The notable difference in the proportions mediated between MVP and the other sites could be due to phenotypic uniqueness of the MVP sample. For example, the MVP comprises mostly male patients (92.8%), which could contribute to residual confounding by sex that is not fully accounted for in the model. Additionally, the mediated pathways could be particularly strong in MVP owing to the high prevalence of depression in the sample (MVP = 44.7%, others = 23.3%). A sensitivity analysis excluding MVP yielded marginally significant results, which indicates that an additional analysis in a larger sample size is warranted.

In the clinic, WBC measurements can be broken down into measurements of each WBC subtype. Abnormal levels of different WBC subtypes can index different immune processes. Understanding which cell types underlie the relationship between depression PGS and depression diagnosis through WBC can help narrow a specific immune process involved in depression. Neutrophil counts explained 1.9% of the association between depression PGS and depression diagnosis, and no other subtypes contributed to the association. Neutrophils are well known as responders to acute bacterial infection and are the most abundant WBC subtype in circulation (40%-60%). Recent evidence demonstrates that neutrophils have essential roles in innate and adaptive immunity, are implicated in diseases of chronic inflammation, and are experimentally shown to transmigrate into intact mouse brain to deliver interleukin 1β, resulting in depressive behavioral change.

Limitations

Our study should be interpreted in light of its limitations. First, the WBC measurements used in the study were clinically derived, with measurements reflecting a range of health states. To address this limitation, we limited to observations within 4 SDs and noted that WBC count was measured for nearly everyone in our primary and replication sample populations. However, it remains possible that individuals with multiple EHRs often contain multiple WBC measurements for the same individual. In this study, only the median values per individual were used, leaving unanswered questions about the effect of depression PGS on WBC count over time and in response to antidepressant treatment. Second, the depression PGS are based on genome-wide association statistics that are not adjusted for phenotypic comorbidities such as pro-inflammatory conditions. This situation introduces the possibility of “phenotypic hitchhiking” in which a comorbid trait is
unintentionally selected during the ascertainment of the index trait. This unintentional selection would create a situation in which 2 heritable phenotypes that share common environmental risk factors but no genetic risk factors can appear correlated in PGS analysis, even in independent samples. Therefore, we emphasize that the PGS approach is still fundamentally an association. Third, results are based on genetic studies of primarily European ancestry populations and may not generalize across diverse ancestries. Fourth, while pleiotropy was assessed in the mendelian randomization analyses, possible unknown sources of confounding (such as those mentioned in the description of “phenotypic hitchhiking”) were not assessed. Finally, even though the association between depression PGS and WBC count was robust, the effect sizes were small, making WBC count an unlikely candidate for use as a diagnostic biomarker of depression.

Table 1. Characteristics of PsycheMERGE Network Sites

<table>
<thead>
<tr>
<th>Site and group</th>
<th>No. of individuals genotyped (European)*</th>
<th>No. with WBC measurement</th>
<th>No. (%) genotyped and with WBC measurement</th>
<th>No. (%) female</th>
<th>No. (%) male</th>
<th>Mean age (SD), y</th>
<th>Mean (SD) length of record, y</th>
</tr>
</thead>
<tbody>
<tr>
<td>Icahn School of Medicine at Mount Sinai</td>
<td>9255</td>
<td>3668</td>
<td>823 (8.9)</td>
<td>429 (52.1)</td>
<td>394 (47.9)</td>
<td>59.7 (16.0)</td>
<td>11.2 (4.4)</td>
</tr>
<tr>
<td>Depression or anxiety controls</td>
<td>6722</td>
<td>2499</td>
<td>578 (8.6)</td>
<td>297 (51.4)</td>
<td>281 (48.6)</td>
<td>59.3 (16.4)</td>
<td>10.7 (4.4)</td>
</tr>
<tr>
<td>Depression or anxiety cases</td>
<td>1622</td>
<td>1169</td>
<td>245 (15.1)</td>
<td>132 (53.9)</td>
<td>113 (46.1)</td>
<td>60.5 (15.0)</td>
<td>12.5 (3.9)</td>
</tr>
<tr>
<td>Vanderbilt University Medical Center</td>
<td>72 828</td>
<td>948 590</td>
<td>70 921 (97.4)</td>
<td>39 661 (55.9)</td>
<td>31 260 (44.1)</td>
<td>48.1 (22.3)</td>
<td>8.7 (6.3)</td>
</tr>
<tr>
<td>Depression or anxiety controls</td>
<td>59 520</td>
<td>301 982</td>
<td>43 129 (72.5)</td>
<td>22 675 (52.6)</td>
<td>20 454 (47.4)</td>
<td>46.8 (23.7)</td>
<td>7.6 (6.1)</td>
</tr>
<tr>
<td>Depression or anxiety cases</td>
<td>15 985</td>
<td>71 692</td>
<td>13 371 (83.6)</td>
<td>8614 (64.4)</td>
<td>4757 (35.6)</td>
<td>50.9 (18.8)</td>
<td>11.3 (6.1)</td>
</tr>
<tr>
<td>Million Veteran Program</td>
<td>289 880</td>
<td>289 880</td>
<td>289 880 (100)</td>
<td>20 871 (7.2)</td>
<td>269 009 (92.8)</td>
<td>64.3 (12.0)</td>
<td>12.0</td>
</tr>
<tr>
<td>Depression or anxiety controls</td>
<td>150 328</td>
<td>150 328</td>
<td>150 328 (100)</td>
<td>6163 (4.1)</td>
<td>144 165 (95.9)</td>
<td>67.7 (11.2)</td>
<td>11.2</td>
</tr>
<tr>
<td>Depression or anxiety cases</td>
<td>129 552</td>
<td>129 552</td>
<td>129 552 (100)</td>
<td>14 121 (10.9)</td>
<td>115 431 (89.1)</td>
<td>61.6 (11.9)</td>
<td>12.9</td>
</tr>
<tr>
<td>Mass General Brigham</td>
<td>25 331</td>
<td>72 329</td>
<td>20 828 (82.2)</td>
<td>10 726 (51.5)</td>
<td>10 102 (48.5)</td>
<td>56.1 (16.7)</td>
<td>13.8 (8.3)</td>
</tr>
<tr>
<td>Depression or anxiety controls</td>
<td>17 879</td>
<td>51 612</td>
<td>17 098 (95.6)</td>
<td>8891 (52.0)</td>
<td>8207 (48.0)</td>
<td>59.8 (16.7)</td>
<td>11.3 (7.1)</td>
</tr>
<tr>
<td>Depression or anxiety cases</td>
<td>7452</td>
<td>20 717</td>
<td>3730 (50.1)</td>
<td>2390 (64.2)</td>
<td>1340 (35.9)</td>
<td>56.7 (16.9)</td>
<td>14.0 (6.7)</td>
</tr>
</tbody>
</table>

Abbreviations: NA, not applicable; WBC, white blood cell.
* The focused WBC analyses included only European samples, because there were no significant findings in the African ancestry LabWAS.

Figure 3. Replication Within the PsycheMERGE Network

The association between depression polygenic scores (PGS) and median white blood cell levels was replicated across the PsycheMERGE Network with sensitivity analyses controlling for depression and anxiety diagnoses. Whiskers indicate 95% CIs.
Conclusions

In this genetic association study, PGS for depression were associated with increased inflammatory markers, specifically WBC count, even in the absence of depressive symptoms. The associations described in this study highlight the importance of WBC biology in depression and demonstrate the potential use of EHR-based genomics as a tool for discovery of physiological markers in psychiatric traits.
Dr Roussos. This study was also supported by the NIH, Bethesda, Maryland, under award No. K08MH122391 (Dr Roussos) and RO1AG067025, RO1MH25246, and RO1MH09677 (Dr Roussos). This study has also been funded in part by the Brain & Behavior Research Foundation via the 2020 Brain & Behavior Research Foundation (formerly NARSAD) Young Investigator Grant No. 29350 (Dr Roussos).

Role of the Funder/Sponsor: The funders had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; preparation, review, or approval of the manuscript; and decision to submit the manuscript for publication.

Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. This publication does not represent the views of the Department of Veteran Affairs or the United States government.

Additional Contributions: We thank the Vanderbilt University Medical Center Biobank, Mass General Brigham Biobank, Million Veteran Program, and Mount Sinai School of Medicine biobank for providing genomic and health information data.

REFERENCES
31. Way NR, Lin T, Austin J, et al. From basic science to clinical application of polygenic risk


