Associations between major psychiatric disorder polygenic risk scores and blood-based markers in UK biobank

Michael D.E. Sewell a,*, Lorena Jiménez-Sánchez a,1, Xueyi Shen b, Amelia J. Edmondson-Stait a, Claire Green b, Mark J. Adams b,c, Olivia M. Rifai a, Andrew M. McIntosh b,c, Donald M. Lyall d, Heather C. Whalley b,2, Stephen M. Lawrie b,2

a Translational Neuroscience PhD Programme, Centre for Clinical Brain Sciences, University of Edinburgh, Kennedy Tower, Royal Edinburgh Hospital, Edinburgh EH10 5HF, UK
b Division of Psychiatry, University of Edinburgh, Kennedy Tower, Royal Edinburgh Hospital, Edinburgh EH10 5HF, UK
c Centre for Cognitive Ageing and Cognitive Epidemiology, Department of Psychology, University of Edinburgh, Edinburgh EH8 9JZ, UK
d Institute of Health & Wellbeing, University of Glasgow, Glasgow G12 8RZ, UK

ARTICLE INFO

Keywords:
Bipolar disorder
Blood markers
Major depressive disorder
Polygenic risk scores
Schizophrenia
UK Biobank

ABSTRACT

Major depressive disorder (MDD), schizophrenia (SCZ), and bipolar disorder (BD) have both shared and discrete genetic risk factors, and are associated with peripheral abnormalities. The relationships between such genetic architectures and blood-based markers are, however, unclear.

We investigated relationships between polygenic risk scores (PRS) for these disorders and peripheral markers in the UK Biobank cohort. We calculated polygenic risk scores for \( n = 367,329 \) (MDD PRS), \( n = 366,465 \) (SCZ PRS), and \( n = 366,383 \) (BD PRS) UK Biobank cohort subjects. We then examined associations between disorder PRS and 58 inflammatory/immune, hematological, bone, cardiovascular, hormone, liver, renal and diabetes-associated blood markers using two generalized linear regression models: ‘minimally adjusted’ controlling for variables such as age and sex, and ‘fully adjusted’ including additional lifestyle covariates: BMI, alcohol and smoking status, and medication intake.

There were 38/58 MDD PRS, 32/58 SCZ PRS, and 20/58 BD PRS-blood marker associations detected for our minimally adjusted model. Of these, 13/38 (MDD PRS), 14/32 (SCZ PRS), and 10/20 (BD PRS) associations remained significant after controlling for lifestyle factors. Many were disorder-specific, with 8/13 unique MDD PRS associations identified. Several disorder-specific associations for MDD and SCZ were immune-related, with mostly positive and negative associations identified for MDD and SCZ PRS respectively.

This study suggests that MDD, SCZ and BD have both shared and distinct peripheral markers associated with disorder-specific genetic risk. The results also implicate inflammatory dysfunction in MDD and SCZ, albeit with differences in patterns between the two conditions, and enrich our understanding of potential underlying pathophysiological mechanisms in major psychiatric disorders.

1. Introduction

Major depressive disorder (MDD), schizophrenia (SCZ) and bipolar disorder (BD) are potentially debilitating psychiatric disorders with significant morbidity and economic impact worldwide (Vigo et al., 2016). Despite this, the pathogenesis of these disorders remains elusive (Soliman et al., 2017), with diagnosis reliant upon symptom presentation (Garcia-Gutierrez et al., 2020). All three disorders are heritable and of polygenic nature (Wray et al., 2014), with strong evidence suggesting that they share some genetic risk (Amare et al., 2020; Cross-Disorder Group of the Psychiatric Genomics Consortium, 2013; Musliner et al., 2019; Wen et al., 2016) and phenotypic traits (Goodkind et al., 2015;
Lee et al., 2015; Pinto et al., 2017). Conversely, other studies have identified disorder-specific genetic associations (Bipolar Disorder and Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2018; Li et al., 2011; Reay and Cairns, 2020) and neurobiological characteristics (Koutsouleris et al., 2015; Maggioni et al., 2017; Schnack et al., 2014; Yang et al., 2019), alluding to variations in the genetic pathophysiology between conditions.

There is also compelling evidence demonstrating that MDD, SCZ, and BD are accompanied by genetically and environmentally influenced alterations in peripheral blood-based marker levels (Hayashi-Takagi et al., 2014; Lai et al., 2016; Strawbridge et al., 2017). Similarities and differences in the nature of these, such as elevated inflammatory markers, have been observed in patients across the three disorders (Goldsmith et al., 2016; Horsdal et al., 2017; Lu et al., 2019; Yuan et al., 2019). Although these peripheral perturbations, including immune system measures, have also differentiated cases from healthy controls (Goldsmith et al., 2016; Modabbernia et al., 2013; Osimo et al., 2020), their potential role in disorder pathophysiology has not yet been ascertained, and currently there are no markers available that can predict the risk of these disorders (Peedecayil, 2019). Conflicting data concerning associations in case-control studies, such as those investigating the relationship between inflammatory marker C-reactive protein (CRP) and schizophrenia (Fond et al., 2018), may be attributable to limitations such as selection bias, small sample sizes, and illness-related confounding factors such as medication.

Progress in identifying blood-based markers for MDD, SCZ, and BD may be advanced by examining peripheral profiles of large samples of individuals according to their genetic predisposition for each disorder, as denoted by their polygenic risk score (PRS) (Lewis and Vassos, 2020). PRS are constructed by summing the number of risk alleles, weighted by effect sizes derived from genome-wide association studies (GWAS) of independent samples (Duncan et al., 2019). Previous work by our group and others has demonstrated neurobiological associations for individuals with higher PRS for all three disorders (Barbu et al., 2019; Neilson et al., 2019; Whalley et al., 2012, 2013). Here, we extend this to compare blood-based markers.

We have employed a data-driven exploratory approach to examine relationships between PRS for all three disorders (MDD, SCZ, BD) with peripheral markers using data from the UK Biobank (UKB) cohort in a total of over 500,000 participants (Sudlow et al., 2015). We were primarily interested in identifying associations with inflammatory/immune markers, as it is increasingly clear that inflammatory and immune dysregulation are characteristics of MDD, SCZ, and BD (Pape et al., 2019; Yuan et al., 2019). We therefore hypothesized that higher disorder PRS for MDD, SCZ, and BD would be related to the levels of these markers. However, we did not restrict our analysis to inflammatory and immune system components. Markers of bone, cardiovascular, liver, and renal function markers, as well as hormone, diabetes-associated and hematological (non-immune) indices were also incorporated, as the high statistical power generated by our large population-based sample could also yield novel associations with other markers not yet implicated in disorder pathology.

2. Materials and Methods

2.1. UK biobank

The UKB study involved 502,617 community-based individuals recruited between 2006 and 2010 in the United Kingdom (https://www.ukbiobank.ac.uk/) (Bycroft et al., 2018; Sudlow et al., 2015). UKB received ethical approval from the research ethics committee (REC reference 11/NW/0382), under application #4844, with all participants providing written consent.

2.2. Study population

The study was conducted using the latest release of UKB blood marker data including 502,536 participants ($n_{\text{male}} = 229,134$, $n_{\text{female}} = 273,402$, mean age at time of blood sampling $= 56.53 \pm 8.10$ years, age range $= 37–73$ years). We firstly removed participants who had supplied no biomarker data ($n = 9,833$). Subsequently, as part of the genetic quality checking procedure described in Howard et al. (2018) we performed prior to PRS generation, we excluded participants who had either non-white British ancestry based on self-reported ethnicity, gender mismatch, relatedness using kinship coefficients calculated from the KING tool ($r > 0.044$) or genotype missingness greater than 2%. For both SCZ and BD PRS analyses, individual identifier information was not available from the Psychiatric Genomics Consortium (PGC) GWAS summary statistics used to compute SCZ and BD PRS (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014; Stahl et al., 2019). To minimize any potential overlap between UKB and PGC SCZ and BD cohorts, we excluded subjects who had received SCZ or BD diagnoses, identified through self-declaration and hospital records, and defined using international classification of disease (ICD10) codes F20-F29 for SCZ, and F30-F31 for BD (11th revision) (http://biobank.ctsu.ox.ac.uk/ryst/field.cgi?id=41202). The same measures were not implemented for MDD PRS analysis, as we were able to directly identify UKB subjects who had also participated in MDD PGC studies in order for them to be removed in the current analysis (See Genotyping and PRS Generation in Methods). This gave us final sample sizes of $n = 367,329$ for MDD PRS, $n = 366,465$ for SCZ PRS and $n = 366,383$ for BD PRS analyses respectively (see Supplementary Fig. S1, Table 1).

2.3. Genotyping and PRS generation

Genotyping of 488,377 blood samples from UKB participants was performed using either the UK BiLEVE http://biobank.ctsu.ox.ac.uk/crystal/um refers.cgi?id=149600 or UKB Axiom arrays http://biobank.ctsu.ox.ac.uk/crystal/um refers.cgi?id=149601 (Bycroft et al., 2018). Further information about the quality control procedures applied to participant genomic data are described in http://www.ukbiobank.ac.uk/scientist s-3/genetic-data, https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/genotyping.qc.pdf, Hagaarnaars et al. (2016) and Bycroft et al. (2018).

Prior to MDD PRS derivation, we removed any subjects who had participated in the UKB study from the MDD GWAS summary statistics used to construct MDD PRS (Howard et al., 2019; Wray et al., 2018). Individual PRS for all disorders at five p value thresholds (0.01, 0.05, 0.1, 0.5 and 1.0) were computed using PRSice software (Euesden et al., 2015) and PLINK (version 1.9; https://www.cog-genomics.org/p link/1.9/), by summing risk alleles weighted on the strength of their association with a certain trait. Association strength has previously been estimated through GWAS conducted by PGC working groups (https://www.med.unc.edu/pgc/pgc-workgroups/). In this manuscript, a SNP inclusion threshold was applied at a significance level of $p \leq 0.5$, in line with previous work showing this threshold is the most predictive of case-control status (International Schizophrenia Consortium et al., 2009; Major Depressive Disorder Working Group of the Psychiatric GWAS Consortium et al., 2013; Psychiatric GWAS Consortium Bipolar Disorder Working Group, 2011; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014), containing 141,802 SNPs for MDD, 138,305 SNPs for SCZ and 145,975 SNPs for BD after clump-based linkage-disequilibrium (LD) pruning ($R^2 = 0.25$, 250-kb window). In addition, SNPs with a minor allele frequency of <1%, Hardy-Weinberg equilibrium ($p < 1 \times 10^{-6}$) or variant call rate (<98%) were removed prior to analysis. The numbers of unpruned and pruned SNPs, and results for other SNP inclusion thresholds $p \leq 0.01$, $p \leq 0.05$, $p \leq 0.1$, and $p \leq 1.0$ are reported in Supplementary Tables S1 and S12-17. We did not exclude the major histocompatibility complex region in any of our analyses.
Table 1
Descriptive statistics for demographic and lifestyle variables for each sample used to compute disorder PRS.

|                      | MDD PRS | SCZ PRS | BD PRS |
|----------------------|---------|---------|--------|
| N                    | 366,383 | 367,329 | 366,465|
| Sex                  |         |         |        |
| Male                 | 197,503 | 198,050 | 197,669|
| (53.9%)              | (53.9%) | (53.9%) |        |
| Female               | 168,880 | 169,279 | 168,796|
| (46.1%)              | (46.1%) | (46.1%) |        |
| Mean Age (Years) (S.D.) | 56.7 (8.02) | 56.7 (8.02) | 56.7 (8.02) |
| Mean BMI (kg/m²) (S.D.) | 27.4 (4.76) | 27.4 (4.76) | 27.4 (4.76) |
| Drinking Status      |         |         |        |
| Never                | 11,471 (3.1%) | 11,535 (3.1%) | 11,474 (3.1%) |
| Previous             | 12,574 (3.4%) | 12,699 (3.5%) | 12,570 (3.4%) |
| Current              | 341,786 | 342,536 | 341,872|
| (93.3%)              | (93.3%) | (93.3%) |        |
| Smoking Status       |         |         |        |
| Never                | 197,356 | 197,753 | 197,434|
| (53.9%)              | (53.8%) | (53.8%) |        |
| Previous             | 129,734 | 130,033 | 129,770|
| (35.4%)              | (35.4%) | (35.4%) |        |
| Current              | 37,806  | 38,051  | 37,782 |
| (10.3%)              | (10.4%) | (10.3%) |        |
| Cholesterol-Lowering Medication |         |         |        |
| N                    | 301,299 | 302,025 | 301,831|
| Y                    | 62,635  | 62,849  | 62,645 |
| (17.1%)              | (17.1%) | (17.1%) |        |
| Blood Pressure Medication |       |         |        |
| N                    | 289,225 | 289,934 | 289,290|
| Y                    | 74,709  | 74,940  | 74,738 |
| (20.4%)              | (20.4%) | (20.4%) |        |
| Insulin              |         |         |        |
| N                    | 360,133 | 361,050 | 360,230|
| Y                    | 3,801 (1.0%) | 3,824 (1.0%) | 3,798 (1.0%) |
| (98.3%)              | (98.3%) | (98.3%) |        |
| Hormone Replacement Therapy |       |         |        |
| N                    | 349,143 | 350,033 | 349,211|
| Y                    | 14,791 (4.0%) | 14,841 (4.0%) | 14,817 (4.0%) |
| (95.3%)              | (95.3%) | (95.3%) |        |
| Oral Contraceptive Pill |       |         |        |
| N                    | 358,732 | 359,661 | 358,824|
| Y                    | 5,020 (1.4%) | 5,213 (1.4%) | 5,204 (1.4%) |
| (97.9%)              | (97.9%) | (97.9%) |        |

BD, bipolar disorder; BMI, body mass index; MDD, major depressive disorder; PRS, polygenic risk scores; SCZ, schizophrenia; S.D., standard deviation; Y/N ‘yes’ or ‘no’ responses.

2.4. Blood-Based markers

UKB performed quality checking procedures for blood sampling data, including assays, which along with instrumentation and analysis methods, are described in: https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/biomarker_issues.pdf https://biobank.ndph.ox.ac.uk/showcase/showcase/docs/haematoiology.pdf and https://www.ukbiobank.ac.uk/media/oiudpjqa/bcm023.ukb_biomarker_panel_website_v1.0-aug-2015-edit-2018.pdf. Details of the raw data available concerning 55 peripheral markers and how they were subdivided into eight groups: ‘inflammatory/immune’, ‘hematological (non-immune)’, ‘cardiovascular’, ‘liver’, ‘renal’, ‘bone’, ‘hormone’ and ‘diabetes mellitus (DM)-associated’, are reported in the Supplementary Methods and Supplementary Table S2. In addition, we calculated the monocyte-to-lymphocyte (MLR), neutrophil-to-lymphocyte (NLR) and platelet-to-lymphocyte (PLR) ratios for subjects where possible (Supplementary Table S3). Previous studies have demonstrated the use of these ratios in predicting inflammatory status in several conditions (Guragac and Demirer, 2016; Mazza et al., 2018), showing higher levels in neuropsychiatric disorder patients (Demir et al., 2015; Mazza et al., 2018, 2019; Ozdin et al., 2017; Semiz et al., 2014).

Baseline data for each marker were normalized using inverse normal transformation (see Supplementary Methods, and then standardized prior to testing for associations using the ‘scale’ function in R (version 3.6.1; R Foundation, Vienna, Austria; https://cran.r-project.org/bin/windows/base/old/3.6.1/). Subsequently, values that were further than five standard deviations away from the mean standardized value (“zero”) were removed (Supplementary Table S4).

2.5. Statistics

Statistical analysis was performed using R (version 3.6.1) in a Linux environment. To investigate associations between disorder PRS at each SNP inclusion threshold and standardized blood marker values, we employed generalized linear regression ‘minimally adjusted’ (MA) and ‘fully adjusted’ (FA) models (glm function in R; “stats” package version 3.6.1). The following covariates were included for our MA model: age, age², sex, genotyping array used, UKB assessment center the participant attended for blood sample collection, and 15 genetic principal components (see Supplementary Methods). Our FA model additionally controlled for body mass index (BMI), alcohol and smoking status, and certain medications including use of cholesterol-lowering, blood pressure, insulin, hormone replacement therapy, and oral contraceptive or minipill medications (see Supplementary Methods). This approach enabled us to distinguish biomarker associations attributed to a higher disorder genetic risk from those more likely related to these particular lifestyle factors. We also performed an additional sensitivity analysis for MDD PRS, by testing for peripheral marker associations for our UKB cohort excluding participants who had received an MDD diagnosis (ICD10 classification codes F32-F39, 11th revision, http://biobank.ctsu.ox.ac.uk/crystal/field.cgi?id=41202), (n = 10,518) (see Supplementary Fig. S1). Multiple testing correction was applied to all tests conducted across all traits at each individual SNP inclusion threshold, using the Bonferroni multiple comparison test (Armstrong, 2014), utilizing the ‘p.adjust’ function in R (“stats” package version 3.6.1). Significant associations were determined using a threshold of p corr < 0.05. Associations are reported as standardized Beta (ß) values throughout this manuscript.

3. Results

3.1. Associations between MDD PRS and peripheral markers

We initially observed significant associations between MDD PRS and 38/58 blood-based markers, reducing to 12/58 upon adjustment for lifestyle factors (Fig. 1, Supplementary Tables S5 & S6). Many markers that were significant for both MA and FA models were immune- inflammatory-related, including positive associations with: CRP (MA: ß = 0.027, p corr = 1.46 × 10−5, FA: ß = 0.0077, p corr = 4.70 × 10−4), white blood cell (MA: ß = 0.028, p corr = 2.47 × 10−6, FA: ß = 0.011, p corr = 1.01 × 10−5) and neutrophil (MA: ß = 0.025, p corr = 6.93 × 10−4, FA: ß = 0.011, p corr = 1.50 × 10−10) counts, and neutrophil percentage (MA: ß = 0.007, p corr = 7.26 × 10−4, FA: ß = 0.006, p corr = 0.016). However, several immune associations failed to remain significant after lifestyle factor adjustment, including monocyte count and percentage, lymphocyte percentage, NLR and PLR, and platelet crit.

Other markers that survived Bonferroni correction in both models include several renal function markers, with 3 out of 4 renal marker MA model associations remaining significant after lifestyle factor adjustment. These included negative associations identified with creatinine
(MA: ß = −0.008, \( p_{corr} = 6.21 \times 10^{-8} \), FA: ß = −0.009, \( p_{corr} = 9.61 \times 10^{-5} \)) and urea (MA: ß = −0.007, \( p_{corr} = 0.003 \), FA: ß = −0.007, \( p_{corr} = 9.87 \times 10^{-5} \)). A positive association was also initially detected with urate (MA: ß = 0.007, \( p_{corr} = 2.50 \times 10^{-5} \)), which became negative after lifestyle factor adjustment (FA: ß = 0.005, \( p_{corr} = 0.006 \)). Only the association with cystatin C became insignificant for our FA model.

Positive associations were also identified with cardiovascular function marker triglycerides (MA: ß = 0.024, \( p_{corr} = 1.71 \times 10^{-4} \), FA: ß = 0.006, \( p_{corr} = 0.003 \)). Conversely, negative associations were detected with bone function marker vitamin D (MA: ß = −0.020, \( p_{corr} = 2.06 \times 10^{-28} \), FA: ß = −0.011, \( p_{corr} = 3.93 \times 10^{-9} \)), hormone insulin-like growth factor 1 (IGF-1) (MA: ß = 0.017, \( p_{corr} = 4.23 \times 10^{-24} \), FA: ß = −0.009, \( p_{corr} = 1.21 \times 10^{-4} \)) and liver function marker total bilirubin (MA: ß = −0.017, \( p_{corr} = 4.26 \times 10^{-24} \), FA: ß = −0.009, \( p_{corr} = 7.32 \times 10^{-7} \)).

All of these associations continued to be significant, with the nature of their fully adjusted effect sizes unchanged, after conducting a sensitivity analysis in which we excluded MDD patients from our sample, with the exception of neutrophil percentage (Supplementary Table S7). A positive association with lymphocyte count, which was initially only significant for our MA model (ß = 0.018, \( p_{corr} = 8.47 \times 10^{-26} \)), was also detected after patient removal (ß = 0.006, \( p_{corr} = 0.020 \), taking the total number of associations detected for both models to 13.

Additionally, we noted significant associations for our MA model with 6/15 hematological (non-immune) markers: high light scatter reticulocyte count and percentage, immature reticulocyte fraction, reticulocyte count and percentage, and red blood cell distribution width, were no longer significant after controlling for lifestyle factors. This was also applicable to four liver function markers: albumin, alanine and aspartate aminotransferases, and direct bilirubin, as well as three cardiovascular function markers: apolipoprotein A, HDL and total cholesterol, hormones SHBG and testosterone, both diabetes-associated parameters glucose and glycated hemoglobin, and bone function marker alkaline phosphatase.
3.2. Associations between SCZ PRS and peripheral markers

32/58 markers were significantly associated with SCZ PRS for the MA model, with 14/58 remaining significant after lifestyle factor correction (Fig. 1, Supplementary Tables S8 & S9). We also observed associations between SCZ PRS and seven immune parameters across both models, including positive associations for white blood cell (MA: $\beta = 0.010$, $p_{corr} = 6.37 \times 10^{-5}$, FA: $\beta = 0.010$, $p_{corr} = 2.53 \times 10^{-5}$) and lymphocyte (MA: $\beta = 0.016$, $p_{corr} = 2.80 \times 10^{-7}$, FA: $\beta = 0.016$, $p_{corr} = 2.07 \times 10^{-19}$) counts, and lymphocyte percentage (MA: $\beta = 0.009$, $p_{corr} = 2.27 \times 10^{-5}$, FA: $\beta = 0.009$, $p_{corr} = 1.49 \times 10^{-5}$).

In contrast to MDD PRS, a negative association was detected for neutrophil percentage (MA: $\beta = -0.008$, $p_{corr} = 9.08 \times 10^{-5}$, FA: $\beta = -0.009$, $p_{corr} = 3.87 \times 10^{-5}$), and inflammatory markers MLR (MA: $\beta = -0.009$, $p_{corr} = 1.17 \times 10^{-5}$, FA: $\beta = -0.008$, $p_{corr} = 3.50 \times 10^{-5}$), NLR (MA: $\beta = -0.009$, $p_{corr} = 5.25 \times 10^{-5}$, FA: $\beta = -0.009$, $p_{corr} = 3.66 \times 10^{-5}$) and PLR (MA: $\beta = -0.012$, $p_{corr} = 2.90 \times 10^{-10}$, FA: $\beta = -0.012$, $p_{corr} = 3.41 \times 10^{-11}$). A greater proportion of immune parameter MA associations with SCZ PRS (78%) survived lifestyle factor adjustment compared to MDD PRS (45%). Out of the nine immune associations detected for our MA model with SCZ PRS, only basophil percentage correction was no longer significant after lifestyle factor adjustment compared to MDD PRS (45%). Out of the nine immune associations detected for our MA model with SCZ PRS, only basophil percentage correction was no longer significant after lifestyle factor adjustment compared to MDD PRS (45%).

Other markers that exhibited positive associations with SCZ PRS included liver function marker aspartate aminotransferase (MA: $\beta = 0.013$, $p_{corr} = 1.56 \times 10^{-12}$, FA: $\beta = 0.017$, $p_{corr} = 1.70 \times 10^{-21}$), and renal function markers total protein (MA: $\beta = 0.011$, $p_{corr} = 9.33 \times 10^{-8}$, FA: $\beta = 0.014$, $p_{corr} = 4.32 \times 10^{-12}$) and phosphate (MA: $\beta = 0.008$, $p_{corr} = 2.14 \times 10^{-14}$, FA: $\beta = 0.006$, $p_{corr} = 0.044$). Negative associations were identified with renal function markers creatinine (MA: $\beta = -0.014$, $p_{corr} = 3.52 \times 10^{-2}$, FA: $\beta = -0.011$, $p_{corr} = 6.23 \times 10^{-13}$), cystatin C (MA: $\beta = -0.011$, $p_{corr} = 2.20 \times 10^{-9}$, FA: $\beta = -0.006$, $p_{corr} = 0.004$) and urate (MA: $\beta = -0.014$, $p_{corr} = 3.33 \times 10^{-19}$, FA: $\beta = -0.007$, $p_{corr} = 4.65 \times 10^{-5}$), and bone function marker vitamin D (MA: $\beta = -0.014$, $p_{corr} = 3.60 \times 10^{-13}$, FA: $\beta = -0.017$, $p_{corr} = 2.96 \times 10^{-26}$).

Similarly to MDD PRS, we also detected several associations with many hematological (non-immune) markers, which were no longer significant upon lifestyle factor adjustment. This included associations with high light scatter reticulocyte count and percentage, immature reticulocyte fraction, mean corpuscular hemoglobin and volume, mean reticulocyte and spherical cell volumes, and reticulocyte count and percentage. All cardiovascular and hormone marker associations also did not remain significant after controlling for lifestyle factors, including associations with apolipoprotein A, HDL cholesterol, SHBG, testosterone and triglycerides, as well as diabetes-associated marker glucose. Urea was the only renal marker to not remain significant after lifestyle factor adjustment.

3.3. Associations between BD PRS and peripheral markers

Fewer associations with peripheral markers were initially detected with BD PRS for our MA model, with 20 blood-based marker associations identified, compared to 38 for MDD PRS, and 32 for SCZ PRS (Fig. 1, Supplementary Tables S10 & S11). Similarly to MDD and SCZ PRS, we observed a negative association with bone marker vitamin D (MA: $\beta = -0.007$, $p_{corr} = 0.010$, FA: $\beta = -0.008$, $p_{corr} = 2.10 \times 10^{-3}$), and renal function marker creatinine (MA: $\beta = -0.012$, $p_{corr} = 1.49 \times 10^{-16}$, FA: $\beta = -0.011$, $p_{corr} = 1.56 \times 10^{-12}$).

Several blood-based marker associations detected for SCZ PRS were also applicable to BD PRS, including liver function marker aspartate aminotransferase (MA: $\beta = 0.007$, $p_{corr} = 0.003$, FA: $\beta = 0.009$, $p_{corr} = 6.39 \times 10^{-5}$), renal function markers cystatin C (MA: $\beta = -0.014$, $p_{corr} = 3.62 \times 10^{-16}$, FA: $\beta = -0.012$, $p_{corr} = 5.40 \times 10^{-12}$) and phosphate (MA: $\beta = -0.009$, $p_{corr} = 4.35 \times 10^{-5}$, FA: $\beta = -0.008$, $p_{corr} = 0.001$), and PLR (MA: $\beta = -0.006$, $p_{corr} = 0.020$, FA: $\beta = -0.006$, $p_{corr} = 0.016$), which was the only immune-related association identified for BD PRS.

However, other associations were unique to BD PRS, including a positive association with cardiovascular function marker HDL cholesterol (MA: $\beta = 0.010$, $p_{corr} = 1.94 \times 10^{-7}$, FA: $\beta = 0.007$, $p_{corr} = 0.001$), and negative associations for hematological parameter high light scatter reticulocyte count (MA: $\beta = -0.010$, $p_{corr} = 9.52 \times 10^{-7}$, FA: $\beta = -0.005$, $p_{corr} = 0.044$), and bone function markers calcium (MA: $\beta = -0.007$, $p_{corr} = 0.005$, FA: $\beta = -0.007$, $p_{corr} = 0.006$) and alkaline phosphatase (MA: $\beta = -0.007$, $p_{corr} = 0.002$, FA: $\beta = -0.006$, $p_{corr} = 0.050$).

As was the case for MDD and SCZ PRS, many of the associations that failed to remain significant after lifestyle factor adjustment for BD PRS were related to hematological (non-immune) markers. For BD PRS, these included associations with high light scatter reticulocyte percentage, immature reticulocyte fraction, mean corpuscular hemoglobin and volume, mean spherical volume, red blood cell count, and reticulocyte count and percentage. Cardiovascular marker apolipoprotein A and renal function marker urate were also no longer significant after controlling for lifestyle factors.

4. Discussion

Here, we have demonstrated that MDD, SCZ and BD PRS display varying associations with blood markers across multiple physiological systems. Notably, a large proportion of marker associations for MDD and SCZ PRS were immune- and inflammatory-related, supporting previous work showing that aberrant immune and inflammatory processes contribute to and possibly underlie these disorders (Muller, 2017). Moreover, the high proportion of inflammatory marker associations we detected with SCZ PRS for both our models suggests inflammatory dysfunction in SCZ may be attributed to increased genetic risk.

In keeping with their shared polygenic architecture, we reported some overlapping peripheral traits across all three disorder PRS, notably negative associations with renal marker creatinine, a by-product of creatine metabolism (Kashani et al., 2020), and vitamin D (Fig. 2). Lower vitamin D levels (Boerman et al., 2016; Milaneschi et al., 2014; Zhu et al., 2020) and abnormal creatine/creatinine levels or metabolism (Allen, 2012; Ongur et al., 2009; Peng et al., 2016) have been previously reported in cases for all three disorders. Moreover, although considered primarily as an indicator of bone health, vitamin D is a multifaceted
blood marker with proposed roles in various physiological processes, including immune regulation, which is thought to be perturbed by vitamin D deficiency (Aranow, 2011; Pludowski et al., 2019). The negative association we found between vitamin D and both MDD and SCZ PRS, in addition to other associations with canonical immune markers, may therefore also be implicative of aberrant immunity in individuals with higher PRS for these disorders. Additionally, vitamin D is also a blood–brain barrier (BBB)-permeable marker (Anjum et al., 2018). BBB impairment has been implicated in the pathophysiology for all three disorders (Kealy et al., 2020), and may also account for altered vitamin D levels in subjects with higher PRS, along with associations with other BBB-permeable markers including IGF-1 (Levada and Troyan, 2017) and triglycerides (Banks et al., 2018) for MDD PRS, and HDL cholesterol (Wang and Eckel, 2014) for BD PRS.

Overlap was more evident between BD and SCZ PRS, with 6/10 BD PRS associations also applicable to SCZ PRS, supporting previous work showing shared genetic risk between the two conditions (International Schizophrenia Consortium et al., 2009; Lichtenstein et al., 2009). However, only the associations of 17/37 detected appear to be disorder-specific, notably for MDD, with 8/13 MDD PRS-specific associations. This important finding highlights differences in peripheral markers of genetic pathophysiology, and the relevance of considering each disorder separately when attempting to identify risk biomarkers.

Moreover, we found several associations with MDD and SCZ PRS that did not mirror those of smaller-scale patient studies, notably lymphocyte count (Cai et al., 2017; Garcia-Rizo et al., 2013) and IGF-1 (Levada and Troyan, 2017) for MDD PRS. Higher lymphocyte counts, along with levels of the innate immunity-associated markers neutrophil count (Selders et al., 2017) and CRP (Sproston and Ashworth, 2018), may all point to elevated acute inflammation in individuals with higher MDD PRS, despite the fact that MDD is regarded as a chronic inflammatory disorder (Berk et al., 2013; Miller, 2020). Although lymphocytes are typically associated with adaptive immunity, roles for lymphocytes in resolving acute inflammation have been previously proposed (Buckley et al., 2013; Rajakariar et al., 2008). An association between genetic risk and acute, rather than chronic, inflammation is also supported by the lack of significant associations we found between MDD PRS and systemic inflammatory markers such as NLR (Guragac and Demirer, 2016) after lifestyle factor correction. This observation supports previous findings showing strong associations between lifestyle factors such as obesity and smoking, and inflammatory dysfunction (Berk et al., 2013), which may, in combination with genetic factors, give rise to chronic inflammation seen in patients, possibly by exacerbating any acute inflammatory response arising from increased genetic risk.

Additionally, inflammation may account for associations identified that were not originally classified as immune/inflammatory parameters, such as IGF-1, gamma glutamyltransferase and total bilirubin. Although reduced serum IGF-1 levels have been observed in MDD patients, IGF-1 is also thought to exert anti-inflammatory effects (Levada and Troyan, 2017), with lower brain levels in several in vivo models associated with a MDD phenotype (Basta-Kaim et al., 2014; Trojan et al., 2016). Classical liver function indicators gamma glutamyltransferase and total bilirubin also display associations with inflammatory markers that resemble those reported in our study (Yamada et al., 2006; Zhao et al., 2019). Renal dysfunction has also been associated with higher levels of inflammatory markers such as CRP (Stuveling et al., 2009), which may pertain to the various renal associations detected with MDD PRS.

Conclusively, we noted negative associations between inflammatory parameters such as MLR, NLR and PLR, and increased genetic risk for SCZ, in contrast to findings from SCZ patient studies. Reduced inflammation in subjects with higher SCZ PRS was corroborated by the negative association with CRP in our MA model, which may have become insignificant upon lifestyle factor corrections due to its strong association with obesity (Visser et al., 1999). CRP may also be more strongly associated with acute, rather than chronic, inflammation (Sproston and Ashworth, 2018), the latter of which has been linked to SCZ pathogenesis (Muller et al., 2015).

Associations between lowered inflammation and elevated SCZ PRS have been suggested by two Mendelian Randomization studies showing that individuals genetically predisposed to lower CRP levels have an increased risk of SCZ, potentially due to a weakened ability to fight early-life infection (Hartwig et al., 2017; Prins et al., 2016). Viral infections in particular have been hypothesized to play a causative role in SCZ pathogenesis (Fruentes and Limosin, 2008), which may explain the association between SCZ PRS and increased lymphocyte counts, a trait replicated in some SCZ patients (Miller et al., 2013). T-lymphocytes are involved in viral responses (Khanolkar et al., 2002), and can also secrete cytokines that have been shown to perturb the BBB (Huppert et al., 2010; Kebir et al., 2007), potentially also driving SCZ onset. Future studies investigating blood-based markers of psychiatric conditions should therefore seek to incorporate additional inflammatory marker measurements, such as pro-inflammatory cytokines, to possibly elicit more information surrounding the MDD and SCZ disorder-associated inflammatory response.

The various associations between SCZ PRS and renal markers are consistent with studies showing higher chronic kidney disease incidence in SCZ patients (Teeng et al., 2015; Tzur Bitan et al., 2019). Although these are regarded as canonical renal function markers, some, such as cystatin C, have also been shown to exert neurobiological influences. In several in vivo models subjected to brain injury, cystatin C was found to be neuroprotective (Fang et al., 2017; Yang et al., 2020). Potential roles for cystatin C, and several other markers associated with SCZ PRS, such as aspartate aminotransferase, in SCZ pathogenesis have yet to be thoroughly investigated. These observations demonstrate the power of big data studies like ours to identify novel markers possibly involved in disorder etiology.

Interestingly, fewer associations were detected with immune markers in subjects with higher BD PRS. Although inflammatory and immune abnormalities have been observed in BD patients (Sterzt et al., 2013), these may well be effects of the condition, rather than causative elements. However, it is also possible that fewer associations were detected due to the lower sample sizes employed in the Stahl et al. (2019) GWAS used to determine BD PRS, reducing their predictive power. Nonetheless, we still observed unique associations for BD, including with alkaline phosphatase and calcium, which along with vitamin D are bone health markers, and may relate to bone abnormalities seen in some BD patients (Chandraskar et al., 2019). The higher proportion of markers that remained significant after correcting for lifestyle factors may also suggest genetic alterations have a greater influence on the peripheral profile for BD in comparison to MDD and SCZ.

We also noted that lifestyle factor adjustment had similar influences on all disorder PRS associations, notably reducing the number of significant associations with hematological (non-immune) and cardiovascular markers, as well as diabetes-associated markers for SCZ and MDD PRS. This is possibly due to strong associations that have been previously identified between these markers and lifestyle factors, such as smoking and red blood cell parameters (Masella et al., 2017), and BMI with both cardiovascular markers such as lipids (Bays et al., 2013), and diabetes-associated indicator glucose (Hu et al., 2004). The significant impact controlling for lifestyle factors had on associations reaffirms the importance of considering both genetic and environmental influences on peripheral marker levels and disorder pathophysiology in future studies.

We additionally observed that the nature of PRS association effect sizes were relatively small for most biomarkers, with the majority of effect sizes even lower than those previously reported for neuroimaging measures (Barbu et al., 2019; Neilson et al., 2019). Small effect sizes are common in big data research due to the small amount of variance that is typically explained by each individual variable (Smith and Nichols, 2018). Although the effect sizes we report are lower than for other patient studies investigating associations with the same markers, such as those investigating relationships between CRP and MDD (Horn et al., 2018), these studies may be confounded by issues such as small sample size and potential confounders.
sizes and selection bias that can reduce their statistical power and artificially inflate effect sizes (Button et al., 2013). Our findings highlight another strength of the current study, and other big data studies, that can detect subtle, but potentially meaningful changes in biomarker levels that are associated with disorder genetic risk.

5. Limitations

Our study has several limitations, perhaps most notably that in UKB we are examining associations in a predominantly healthy population with an older age range. This could also explain why some of our results do not reflect those of patient studies, such as one showing that increased CRP levels predispose to SCZ onset in adolescent participants (Metcalfe et al., 2017), a sample less likely to be confounded by longstanding illness and disorder-related medications. Moreover, our analyses are based on single cross-sectional assessments, as repeated measurements of peripheral marker levels are planned for only a subset of the UKB cohort and were not available at the time of our analyses (Bycroft et al., 2018). Consequently, we cannot determine whether the PRS associations of the peripheral markers we observed are stable over time.

In addition, variable heritability estimates across peripheral markers have been reported (Sinnott-Armstrong et al., 2021), and markers may be significantly influenced by environmental and lifestyle factors unaccounted for by our models, which may play a role in the pathogenesis of MDD, SCZ and BD in conjunction with genetic risk variants (Uher and Zwicker, 2017). This is especially relevant to MDD, which has been reported to have a lower heritability than both schizophrenia and bipolar disorders (Wray and Gottesman, 2012), and demonstrates the importance of other approaches such as Mendelian randomization to infer genetic causality, as previously performed for several polygenic conditions using biomarker-associated SNPs as genetic instruments and UK Biobank data (Larsson et al., 2020; Sinnott-Armstrong et al., 2021), as well as laboratory models. Additionally, lifestyle factors controlled for in our FA model such as obesity are influenced by both genetics and environment (Albuquerque et al., 2017), the individual contributions of which we cannot directly determine. We also have only accounted for the use of a limited number of medications in our FA models, although confounding effects of medications such as antipsychotic drugs are likely reduced as a result of our analyses without patients. Our models also do not account for potential technical confounders such as regression dilution bias, although the effects of this on serum biomarker measurements are thought to be minimal in most cases (Allen et al., 2020).

We also have not excluded patients with other non-neuropsychiatric conditions, which could influence certain biomarker levels. However, many of those for which we have diagnostic data available, such as for diabetes, are common comorbidities of neuropsychiatric disorders, with a degree of overlap reported (Amare et al., 2017; Mizuki et al., 2020).

We also cannot exclude the possibility of some participant overlap between UKB and SCZ and BD GWAS cohorts used to determine PRS and associations, as summary statistics from the PGC SCZ and BD GWAS do not provide this information. However, the effects of this limitation may be considered minor and confined to controls, due to our efforts to minimize overlap through removal of neuropsychiatric disorder cases in our UKB sample. Finally, our results may have limited generalizability as we only included Caucasian participants. Future studies should therefore seek to independently replicate these findings in different datasets and in populations with larger sample sizes than ours. In particular, greater participation in BD GWAS may also yield more associations for BD PRS.

5.1. Conclusions

In summary, we have demonstrated varying associations between SCZ, MDD and BD PRS with inflammatory and other blood-based markers, influenced by genetic risk and robust to controlling for lifestyle factors. This reaffirms that, despite their shared genetic risk, aspects of the underlying neurobiology of these disorders can be differentiated. Our study also reinforces the potential value and importance of examining peripheral components of disorder pathogenesis. Incorporation of additional markers, laboratory models, longitudinal data and analyses of other datasets will facilitate testing for causal inferences.

6. Funding and disclosures

This study was supported by a Wellcome Trust Strategic Award (“Stratifying Resilience and Depression Longitudinally”; Grant No. 104036/Z/14/Z; principal investigator, AMM). This research was conducted using the UKB Resource under approved project 4844. MDES, LJS, AJES and OMR all receive support from the University of Edinburgh Wellcome Trust Translational Neuroscience 4-year PhD programme (Grant No. 108890/Z/15/Z). CG is supported by The Medical Research Council and The University of Edinburgh through the Precision Medicine Doctoral Training programme.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank UKB participants for their time and UKB team members for collating the data.

Data availability

The data used in the present study is available from UK Biobank which was obtained and used under a licence, and thus is not publicly available. Interested parties can apply for access through a standard protocol (https://www.ukbiobank.ac.uk/register-apply/). The summary statistics of PGC can be downloaded from https://www.med.unc.edu/pgc/.

Code availability

All code used for data preparation and analysis are available upon request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.bbi.2021.06.002.

References

Albuquerque, D., Nobrega, C., Manco, L., Padua, C., 2017. The contribution of genetics and environment to obesity. Br. Med. Bull. 123, 159–173.
Allen, N.E., Arnold, M., Parish, S., Hill, M., Sheard, S., Callen, H., Fry, D., Moffat, S., Gordon, M., Welsh, S., Elliott, P., Collin, R., 2020. Approaches to minimizing the epidemiological impact of sources of systematic and random variation that may affect biochemistry assay data in UK Biobank. Wellcome Open Res. 5, 222.
Allen, P.J., 2012. Creatine metabolism and psychiatric disorders: Does creatine supplementation have therapeutic value? Neurosci. Biobehav. Rev. 36, 1442–1462.
Amare, A.T., Schubert, K.O., Klingler-Hoffmann, M., Cohen-Woods, S., Baune, B.T., 2017. The genetic overlap between mood disorders and cardiometabolic diseases: a systematic review of genome wide and candidate gene studies. Tranl. Psychiatry 7, e1007.
Amare, A.T., Vaz, A., Hsu, Y.H., Direk, N., Kamali, Z., Howard, D.M., McIntosh, A.M., Tienmeier, H., Bulthann, U., Snieder, H., Hartman, C.A., 2020. Bivariate genome-wide association analyses of the broad depression phenotype combined with major depressive disorder, bipolar disorder or schizophrenia reveal eight novel genetic loci for depression. Mol. Psychiatry 25, 1420–1429.
