Shared genetic factors do not account for the observed co-occurrence of depression and autoimmune diseases in the UK Biobank.

Kylie P Glanville¹, Jonathan R I Coleman¹,², Paul F O'Reilly³, *James Galloway⁴, *Cathryn M Lewis¹,²,⁵

*Contributed equally and share senior authorship.

1. Social Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, GB

2. NIHR Biomedical Research Centre, South London and Maudsley NHS Trust, King's College London, London, GB

3. Genetics and Genomic Sciences, Icahn School of Medicine, Mount Sinai, New York, NY, US

4. Department of Inflammation Biology, King's College London, London, GB

5. Department of Medical & Molecular Genetics, King's College London, London, GB

Address correspondence to Kylie P. Glanville, kylie.glanville@kcl.ac.uk; Social, Genetic and Developmental Psychiatry Centre; Institute of Psychiatry, Psychology & Neuroscience; King’s College London, de Crespigny Park, London SE5 8AF; UK.

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.
Abstract

**Background:** Epidemiological studies have shown increased comorbidity between depression and autoimmune diseases. The mechanisms driving the comorbidity are poorly understood, and a highly powered investigation is needed to understand the relative importance of shared genetic influences. We investigated the evidence for pleiotropy from shared genetic risk alleles between these traits in the UK Biobank (UKB).

**Methods:** We defined autoimmune and depression cases using information from hospital episode statistics, self-reported conditions and medications, and mental health questionnaires. Pairwise comparisons of depression prevalence between autoimmune cases and controls, and vice-versa, were performed. Cross-trait polygenic risk score (PRS) analyses were performed to test for pleiotropy, i.e. testing whether PRS for depression could predict autoimmune disease status, and vice-versa.

**Results:** We identified 28k cases of autoimmune diseases (pooling across 14 traits), and 65k cases of depression. The prevalence of depression was significantly higher in autoimmune cases compared to controls, and vice-versa. PRS for myasthenia gravis and psoriasis were significantly associated with depression case-status ($p < 5.2 \times 10^{-5}$, $R^2 \leq 0.04\%$). PRS for depression were significantly associated with case-status for coeliac disease, inflammatory bowel disease, psoriasis, psoriatic arthritis, rheumatoid arthritis and type 1 diabetes ($p < 5.8 \times 10^{-5}$, $R^2$ range 0.06% to 0.27%).
Conclusions: Consistent with the literature, depression was more common in individuals with autoimmune diseases compared to controls, and vice-versa, in the UKB. PRS showed some evidence for involvement of shared genetic factors, but the modest $R^2$ values suggest that shared genetic architecture accounts for only a small proportion of the increased risk across traits.
Introduction

There is substantial evidence that individuals with a history of autoimmune disease are at greater risk for developing depression\(^1-^4\), and that a history of depression increases risk for developing autoimmune diseases\(^5,^6\). The mechanisms driving the bi-directional relationship are poorly understood, but one contributory factor may be that these diseases share biological pathways.

We and others have previously shown that there is no strong evidence for the involvement of Human Leukocyte Antigen (HLA) alleles in risk for depression, suggesting that the Major Histocompatibility Complex (MHC) does not harbor shared risk for depression and autoimmune diseases\(^6^-^8\). However, genetic risk for autoimmune diseases occurs across the genome\(^9\), and pleiotropic effects outside the MHC may be involved in shared risk for depression and autoimmune diseases.

Few studies have investigated evidence for genome-wide pleiotropy between depression and autoimmune diseases. Euesden, et al.\(^10\) found no evidence for association between polygenic risk scores (PRS) for depression and risk for rheumatoid arthritis, or vice-versa. The Psychiatric Genomics Consortium (PGC) indicated no evidence for significant genetic correlations (\(r_G\)) between depression and nine autoimmune diseases; the strongest correlation observed was between depression and inflammatory bowel disease (\(r_G = .07, p = .06\)\(^11\)). Recently, Liu, et al.\(^6\) found no association between PRS for mental health disorders and risk for autoimmune diseases, and only a weak association between PRS for autoimmune diseases and risk for mental health disorders.
In this study we extend previous work, leveraging the UK Biobank (UKB) to test for pleiotropy between depression and autoimmune diseases with PRS methodology. Given the challenge of reliably defining complex disease traits using large-scale data, we take two approaches to defining autoimmune diseases and depression. We classified liberally-defined cases, based on a single item endorsing diagnosis with an autoimmune disease, and then strictly-defined cases, based on multiple items. We take a similar approach to classifying depression by requiring a greater number of endorsements in strictly-defined cases than liberally-defined cases. Liberally-defined cases increase the sample size, while strictly-defined cases will reduce the rate of misclassification. We perform cross-trait PRS analyses, testing for association between PRS for autoimmune diseases and depression, and vice-versa. Motivated by the observation of sex-dependent genetic correlations between schizophrenia and autoimmune diseases\textsuperscript{12}, and by higher prevalence in females of both depression and autoimmune diseases, we stratified PRS analyses by sex. Our study is one of the largest to explore evidence for pleiotropy between depression and autoimmune diseases and elucidates the contribution of shared genetic influences to the observed comorbidity.
Methods

Participants

The UKB is a prospective health study of 500,000 individuals in the United Kingdom. Participants were identified through NHS patient registers if they were aged 40-69 during the recruitment phase (2006-2010) and living in proximity to an assessment centre. Participants attended a baseline assessment and contributed health information via touchscreen questionnaires and verbal interviews\(^\text{13}\). Subsets of participants completed repeat assessments: instance 1) \(n = 20,335\) between 2012-2013; instance 2) \(n = 42,961\) (interview) and \(n = 48,340\) (touchscreen) in 2014; and instance 3) \(n = 2,843\) (interview) and \(n = 3,081\) (touchscreen) in 2019. Participant data are linked to Hospital Episode Statistics (HES) containing information on episodes of inpatient hospital care. Episodes are coded at admission using the International Classification of Diseases, 10\(^\text{th}\) Revision\(^\text{14}\) (ICD-10). Inpatients are assigned one primary code (reason for admission) and a variable number of secondary codes. Additional data are available for psychiatric phenotyping, including an online Mental Health Questionnaire (MHQ) completed by 157,366 participants in 2017\(^\text{15}\). The UKB received ethical approval from the North West - Haydock Research Ethics Committee (reference 16/NW/0274). Participants provided electronic signed consent at recruitment\(^\text{13}\).
Autoimmune phenotyping

Guided by studies that investigated the epidemiological relationship between autoimmune diseases and depression\(^1,5\) we identified cases for fourteen autoimmune diseases: pernicious anemia (PA), autoimmune thyroid disease (ATD), type 1 diabetes (T1D), multiple sclerosis (MS), myasthenia gravis (MG), coeliac, inflammatory bowel disease (IBD; includes Crohn’s disease and ulcerative colitis), psoriasis, ankylosing spondylitis (AS), polymyalgia rheumatica/giant cell arteritis (PR/GCA), psoriatic arthritis (PsA), rheumatoid arthritis (RA), sjögren syndrome (SS), and systemic lupus erythematosus (SLE).

Two sources of information were used to define autoimmune cases and controls. (1) HES: primary and secondary ICD-10 diagnoses recorded between April 1997 to October 2016 were identified from the UKB Data Portal Record Repository. (2) Verbal interview: we used participants’ responses at baseline or instance 1 or 2 to determine self-endorsed medical conditions (past and current) and self-endorsed prescription medications (current). ICD-10 codes, self-endorsed conditions and medications used to define each autoimmune disease are listed in the Supplementary Material.

We took two approaches to defining autoimmune cases (Figure 1). To increase sample size, we created ‘possible’ cases, comprising participants with an ICD-10 diagnosis or a self-endorsed condition. To increase validity, we used multiple observations to create ‘probable’ cases. Participants were coded as probable cases if at least two of ICD-10 diagnosis, self-endorsed condition or medication were observed. More than one ICD-10 diagnosis for the corresponding
autoimmune disease was also sufficient. A set of autoimmune controls was defined from participants with no ICD-10 diagnoses, self-endorsed conditions or medications for all fourteen autoimmune diseases. A single set of controls was used for all autoimmune diseases, given the known comorbidity between them.

Figure 1: Autoimmune phenotyping approach. Cases are included in possible or probable if they fall within a shaded area. Autoimmune medication was used as a confirmatory, but not a primary source of information, because several medications are not disease-specific.

Depression phenotyping

We created two depression case groups: strictly-defined cases termed ‘stringent depression’ and liberally-defined cases termed ‘any depression’. We have previously shown that SNP-based heritability increases with multiple endorsements of depression\textsuperscript{16}. We therefore classified
‘stringent depression’ as participants endorsing at least three of the following depression measures: ICD-10 diagnoses (F32-F33.9); self-reported depression; self-reported antidepressant usage; single or recurrent depression (defined by Smith, et al.\textsuperscript{17} from responses to a touchscreen questionnaire completed at baseline by 172,751 participants); or answered ‘yes’ to the touchscreen questionnaire: "Have you ever seen a GP/psychiatrist for nerves, anxiety, tension or depression?".

We classified ‘any depression’ as participants who endorsed two or more depression measures, or if they met criteria for lifetime depression in the Composite International Diagnostic Interview (CIDI) assessed in the MHQ\textsuperscript{15}. We classify cases defined from CIDI alone as ‘any depression’ not as ‘stringent depression’ because we previously observed lower SNP-based heritability in this group ($h^2_{SNP} = 11\%, SE = 0.008$) compared to cases defined by three or more non-CIDI measures of depression ($h^2_{SNP} = 19\%, SE = 0.018$)\textsuperscript{16}.

Depression cases were screened for schizophrenia and bipolar according to any indication: ICD-10 diagnoses (F20-29, F30-31.9, F34-39); self-endorsed conditions (schizophrenia, mania, bipolar disorder or manic depression) or self-endorsed antipsychotic usage reported at baseline or instance 1 or 2; Bipolar Type I (Mania) or Bipolar Type II (Hypomania) according to the criteria adopted by Smith, et al.\textsuperscript{17}; or indications of psychosis endorsed in the MHQ. A single set of depression controls was defined from participants who did not meet the criteria for depression, schizophrenia or bipolar.
Detail on the derivation of each indication of depression, schizophrenia and bipolar can be found in Supplementary Materials from our previous publication\textsuperscript{16}.

**Genetic quality control (QC)**

The UKB performed preliminary QC on genotype data assayed for all participants\textsuperscript{13}. Using genetic principal components (PCs) provided by the UKB, we performed 4-means clustering on the first two PCs to identify and retain individuals of European ancestry. QC was then performed using PLINK v1.9\textsuperscript{18} to remove: variants with missingness > 0.02 (before individual QC), individuals with missingness > 0.02, individuals whose self-reported sex was discordant from their genetic sex, variants with missingness > 0.02 (after individual QC), variants departing from Hardy-Weinberg Equilibrium (\(p < 10^{-8}\)), and variants with minor allele frequency (MAF) < 0.01. Relatedness kinship estimates provided by the UKB were used to identify pairs of related individuals (KING \(r^2 > 0.044\))\textsuperscript{19} and the GreedyRelated\textsuperscript{20} algorithm used to remove one individual from each pair, preferentially retaining individuals that survived QC. FlashPCA2\textsuperscript{21} was used to generate PCs for the sub-set of individuals of European ancestry surviving QC.

**Statistical analyses**

We summarised sociodemographic data taken at baseline assessment: age, sex, socio-economic status (SES), body mass index (BMI) and current smoking status. We tested for significant differences in sociodemographic variables between cases and controls using Welch Two Sample t-tests in R v3.6\textsuperscript{22}. We tested for significant differences in: 1) the prevalence of depression in
autoimmune cases compared to autoimmune controls, and 2) the prevalence of autoimmune
diseases in depression cases compared to depression controls. These tests were performed for
both probable/possible autoimmune cases and stringent/any depression, using 2-sample tests
for equality of proportions in R v3.6.

Summary statistics for autoimmune diseases and depression

We identified genome-wide association studies (GWAS) with publicly-available summary
statistics for eight of the fourteen autoimmune diseases: celiac, IBD, MS, MG, psoriasis,
PsA, RA, and SLE (Supplementary Table 1). Summary statistics from GWAS using the
Immunochip were excluded as it does not provide genome-wide coverage. For Major Depressive
Disorder (MDD), we used summary statistics from Wray, et al., excluding UKB. QC was
performed on summary statistics to remove variants within the Major Histocompatibility
Complex (28.8 to 33.7 Mb), and variants in linkage disequilibrium ($r^2 > 0.1$) with the lead variant
within a 250kb region.

Polygenic risk score (PRS) analyses

PRS analyses were conducted using the PRSice-2 software. To validate our phenotyping
approach, we tested for association between PRS for eight autoimmune diseases and case-
control status for the corresponding autoimmune diseases (possible and probable cases), and
between PRS for MDD and case-control status for depression (any and stringent).
To investigate pleiotropy between autoimmune diseases and depression we performed cross-trait analyses, testing for association between 1) PRS for eight autoimmune diseases and case-control status for depression (any and stringent cases); and 2) PRS for MDD and case-control status for fourteen autoimmune diseases (possible and probable cases). To test for sex-specific effects, we performed cross-trait analyses in males and females separately.

For each test, PRS constructed at eight p-value thresholds ($P_T$; 0.001, 0.05, 0.1, 0.2, 0.3, 0.4, 0.5 and 1.0) were regressed on case-control status using logistic regression, adjusting for the following covariates: six PCs, genotyping batch, and assessment centre (n=128 variables). We report p-values at the optimal $P_T$ for each test of association. To control for multiple testing across $P_T$ (x8), and across tests of association (autoimmune PRS (x8) predicting any/stringent depression (x2) in men and women (x2), n=32; and MDD PRS predicting possible/probable (x2) autoimmune diseases (x14) in men and women (x2), n=56), a Bonferroni correction was applied to give a p-value threshold for significance of 7.1x10^{-5} (0.05/704 tests, 704 = 8*(32+56)). Where sex-specific associations were observed, sensitivity analyses were conducted to account for different sample sizes between sexes. We tested for interactions between sex and PRS (at the optimal $P_T$ from sex-specific tests) in the full sample (Phenotype ~ sex + PRS + sex*PRS + covariates). We report $R^2$ estimates transformed to the liability scale using the following population prevalences for outcome traits: PA=0.1%, ATD=2%, T1D=0.3%, MS=0.1%, MG=0.02%, coeliac=1%, IBD=0.5%, psoriasis=2%, AS=0.55%, PR/GCA= 0.85%, PsA=0.5%, RA=1%, SS=0.7%, SLE=0.1% and MDD=15%.
AVENGEME\textsuperscript{42} was used to estimate power to detect cross-trait PRS associations, assuming varying degrees of genetic correlation (rG) between corresponding traits (rG 0.01 to 0.5). Power was estimated for cross-trait analyses where summary statistics for both traits were publicly available, such that SNP-based heritability (required for AVENGEME) could be estimated using Linkage Disequilibrium Score Regression\textsuperscript{43} (Supplementary Figure 5). Power was estimated using PRS at the optimal $P_T$ identified in cross-trait association tests, and liberally-defined sample sizes. Parameters used to estimate power are given in Supplementary Tables 6 and 7.
Results

A total of 28,479 individuals were identified as possible cases across fourteen autoimmune diseases, and a sub-set of 16,824 (59.1%) met the stringent criteria for probable cases (refer Supplementary Material for representation of the overlap between criteria used to define cases). 65,075 individuals met the criteria for any depression, 14,625 of whom met the criteria for stringent depression. Sociodemographic characteristics for autoimmune and depression cases and controls are summarised in Table 1. Overall, autoimmune and depression case groups contained a higher proportion of females, had lower SES, higher smoking prevalence, and higher BMI than their respective control groups, (N = 324,074 autoimmune controls, N = 232,552 depression controls, all p-values < 5x10^{-21} in pairwise comparisons).
<table>
<thead>
<tr>
<th>Autoimmune Diseases (system/disease)</th>
<th>Pop. Prev.</th>
<th>Count</th>
<th>UKB Prev.</th>
<th>Mean age (SD)</th>
<th>Female (%)</th>
<th>TDI (SD)</th>
<th>Current smoker (%)</th>
<th>Mean BMI (SD)</th>
<th>Count</th>
<th>UKB Prev.</th>
<th>Mean age (SD)</th>
<th>Female (%)</th>
<th>TDI (SD)</th>
<th>Current smoker (%)</th>
<th>Mean BMI (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Autoimmune Disease</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pernicious Anemia</td>
<td>0.10%</td>
<td>1,555</td>
<td>0.48%</td>
<td>58.9 (7.61)</td>
<td>71%</td>
<td>-0.85</td>
<td>3.19</td>
<td>13%</td>
<td>28.0</td>
<td>(5.55)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Autoimmune Thyroid Disease</td>
<td>2.00%</td>
<td>859</td>
<td>0.26%</td>
<td>56.8 (7.75)</td>
<td>85%</td>
<td>-1.11</td>
<td>3.05</td>
<td>15%</td>
<td>27.4</td>
<td>(5.11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type 1 diabetes</td>
<td>0.30%</td>
<td>2,751</td>
<td>0.84%</td>
<td>58.3 (7.80)</td>
<td>42%</td>
<td>-0.46</td>
<td>3.39</td>
<td>13%</td>
<td>30.1</td>
<td>(6.08)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Multiple Sclerosis</td>
<td>0.10%</td>
<td>1,683</td>
<td>0.52%</td>
<td>55.4 (7.52)</td>
<td>73%</td>
<td>-1.45</td>
<td>3.01</td>
<td>16%</td>
<td>26.9</td>
<td>(5.08)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Myasthenia Gravis</td>
<td>0.02%</td>
<td>234</td>
<td>0.07%</td>
<td>59.2 (7.39)</td>
<td>56%</td>
<td>-0.94</td>
<td>3.27</td>
<td>11%</td>
<td>29.0</td>
<td>(5.44)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coeliac</td>
<td>1.00%</td>
<td>2,364</td>
<td>0.72%</td>
<td>57.8 (7.79)</td>
<td>67%</td>
<td>-1.49</td>
<td>2.99</td>
<td>7%</td>
<td>25.8</td>
<td>(4.57)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inflammatory Bowel Disease</td>
<td>0.50%</td>
<td>5,105</td>
<td>1.55%</td>
<td>57.4 (7.92)</td>
<td>51%</td>
<td>-1.27</td>
<td>3.10</td>
<td>10%</td>
<td>27.3</td>
<td>(4.70)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psoriasis</td>
<td>2.00%</td>
<td>5,459</td>
<td>1.66%</td>
<td>56.8 (8.01)</td>
<td>46%</td>
<td>-1.05</td>
<td>3.22</td>
<td>15%</td>
<td>28.4</td>
<td>(5.17)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ankylosing Spondylitis</td>
<td>0.55%</td>
<td>1,344</td>
<td>0.41%</td>
<td>57.9 (7.47)</td>
<td>38%</td>
<td>-1.02</td>
<td>3.21</td>
<td>13%</td>
<td>27.8</td>
<td>(4.80)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Polymyalgia Rheumatica/GCA</td>
<td>0.85%</td>
<td>1,627</td>
<td>0.50%</td>
<td>63.1 (5.16)</td>
<td>67%</td>
<td>-1.70</td>
<td>2.79</td>
<td>9%</td>
<td>28.3</td>
<td>(5.08)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Psoriatic Arthritis</td>
<td>0.50%</td>
<td>1,107</td>
<td>0.34%</td>
<td>56.6 (7.48)</td>
<td>50%</td>
<td>-1.16</td>
<td>3.18</td>
<td>11%</td>
<td>29.0</td>
<td>(5.43)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rheumatoid Arthritis</td>
<td>1.00%</td>
<td>6,360</td>
<td>1.92%</td>
<td>59.5 (7.06)</td>
<td>67%</td>
<td>-0.91</td>
<td>3.29</td>
<td>13%</td>
<td>28.5</td>
<td>(5.53)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sjögren Syndrome</td>
<td>0.70%</td>
<td>647</td>
<td>0.20%</td>
<td>59.3 (7.10)</td>
<td>90%</td>
<td>-1.20</td>
<td>3.03</td>
<td>6%</td>
<td>27.1</td>
<td>(5.72)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Systemic Lupus Erythematosus</td>
<td>0.10%</td>
<td>624</td>
<td>0.19%</td>
<td>56.7 (8.18)</td>
<td>84%</td>
<td>-1.02</td>
<td>3.21</td>
<td>14%</td>
<td>27.3</td>
<td>(5.65)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Any Autoimmune Disease</td>
<td>NA</td>
<td>28,479</td>
<td>8.08%</td>
<td>58.1 (7.73)</td>
<td>58%</td>
<td>-1.11</td>
<td>3.17</td>
<td>12%</td>
<td>27.9</td>
<td>(5.28)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Autoimmune Controls</td>
<td>NA</td>
<td>324,074</td>
<td>NA</td>
<td>56.4 (8.06)</td>
<td>52%</td>
<td>-1.50</td>
<td>2.98</td>
<td>10%</td>
<td>27.2</td>
<td>(4.64)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Depression                          |            |       |           |              |            |         |                   |              |       |           |              |            |         |                   |              |
| Any                                | 15%        | 65,075| 21.86%    | 55.4 (7.86)  | 67%        | -1.09   | 3.14              | 13%          | 27.8  | (5.32)   |              |            |         |                   |              |
| Stringent                          |            |       |           |              |            |         |                   |              |       |           |              |            |         |                   |              |
| Depression Controls                |            | 232,552| NA  | 57.1 (8.10)  | 47%        | -1.65   | 2.89              | 9%           | 27.2  | (4.53)   |              |            |         |                   |              |

Table 1: Sociodemographic information for autoimmune and depression cases and controls. Pop. Prev. = population prevalence estimate. UKB Prev. = prevalence of cases in the UKB as a proportion of autoimmune/depression controls. TDI = Townsend Deprivation Index; negative scores indicate less deprivation. SD = standard deviation.
The prevalence of any depression was significantly higher in autoimmune cases compared to autoimmune controls (p = 6x10^{-177} for possible cases of any autoimmune disease versus controls, p = 2x10^{-124} for probable cases of any autoimmune disease versus controls). The prevalence of stringent depression was significantly higher in autoimmune cases compared to autoimmune controls (p = 3x10^{-207} for possible cases of any autoimmune disease versus controls, p = 6x10^{-163} for probable cases of any autoimmune disease versus controls) (Table 2).

Table 2: Prevalence of depression within autoimmune cases compared to autoimmune controls, stratified by possible/probable for autoimmune diseases, and any/stringent for depression cases. P-values from pairwise comparisons of depression prevalence in autoimmune cases compared to autoimmune controls are shown in brackets.

<table>
<thead>
<tr>
<th>Autoimmune Trait</th>
<th>Autoimmune Controls</th>
<th>Possible Autoimmune</th>
<th>Probable Autoimmune</th>
<th>Possible Autoimmune</th>
<th>Probable Autoimmune</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blood</td>
<td>Any Autoimmune Disease</td>
<td>20.6% (6x10^{-177})</td>
<td>29.5% (2x10^{-124})</td>
<td>10.8% (3x10^{-207})</td>
<td>11.5% (6x10^{-183})</td>
</tr>
<tr>
<td>Endocrine system</td>
<td>Pernicious Anemia</td>
<td>35.8% (4x10^{-24})</td>
<td>33.7% (4x10^{-9})</td>
<td>15.2% (1x10^{-19})</td>
<td>16.6% (4x10^{-15})</td>
</tr>
<tr>
<td>Nervous System</td>
<td>Autoimmune Thyroid Disease</td>
<td>35.1% (6x10^{-19})</td>
<td>34.7% (6x10^{-13})</td>
<td>13.7% (1x10^{-16})</td>
<td>13.4% (3x10^{-13})</td>
</tr>
<tr>
<td></td>
<td>Type 1 diabetes</td>
<td>31.1% (3x10^{-11})</td>
<td>31.3% (1x10^{-7})</td>
<td>14.1% (5x10^{-10})</td>
<td>13.5% (5x10^{-7})</td>
</tr>
<tr>
<td></td>
<td>Multiple Sclerosis</td>
<td>39.3% (6x10^{-16})</td>
<td>42.5% (3x10^{-12})</td>
<td>16.9% (5x10^{-14})</td>
<td>20.0% (8x10^{-12})</td>
</tr>
<tr>
<td></td>
<td>Myasthenia Gravis</td>
<td>33.3% (5x10^{-1})</td>
<td>29.2% (3x10^{-7})</td>
<td>14.1% (5x10^{-6})</td>
<td>14.8% (4x10^{-5})</td>
</tr>
<tr>
<td>Digestive system</td>
<td>Coeliac</td>
<td>27.3% (3x10^{-7})</td>
<td>29.8% (2x10^{-12})</td>
<td>8.4% (4x10^{-9})</td>
<td>9.5% (7x10^{-9})</td>
</tr>
<tr>
<td></td>
<td>Inflammatory Bowel Disease</td>
<td>24.9% (6x10^{-12})</td>
<td>24.9% (3x10^{-6})</td>
<td>8.9% (8x10^{-13})</td>
<td>9.1% (2x10^{-13})</td>
</tr>
<tr>
<td>Skin</td>
<td>Psoriasis</td>
<td>26.8% (2x10^{-2})</td>
<td>28.3% (1x10^{-7})</td>
<td>8.7% (4x10^{-9})</td>
<td>9.9% (4x10^{-9})</td>
</tr>
<tr>
<td>Musculoskeletal system and connective tissue</td>
<td>Ankylosing Spondylitis</td>
<td>26.9% (5x10^{-1})</td>
<td>28.4% (5x10^{-9})</td>
<td>9.6% (7x10^{-9})</td>
<td>10.5% (9x10^{-9})</td>
</tr>
<tr>
<td></td>
<td>Polymyalgia Rheumatica/GCA</td>
<td>29.3% (2x10^{-12})</td>
<td>27.8% (5x10^{-6})</td>
<td>10.8% (3x10^{-13})</td>
<td>10.6% (1x10^{-12})</td>
</tr>
<tr>
<td></td>
<td>Psoriatic Arthritis</td>
<td>31.8% (3x10^{-13})</td>
<td>34.3% (5x10^{-19})</td>
<td>10.2% (1x10^{-14})</td>
<td>13.0% (2x10^{-13})</td>
</tr>
<tr>
<td></td>
<td>Rheumatoid Arthritis</td>
<td>30.6% (5x10^{-12})</td>
<td>29.4% (6x10^{-18})</td>
<td>12.7% (3x10^{-15})</td>
<td>12.1% (1x10^{-15})</td>
</tr>
<tr>
<td></td>
<td>Sjögren Syndrome</td>
<td>40.7% (4x10^{-20})</td>
<td>41.9% (1x10^{-13})</td>
<td>18.9% (2x10^{-20})</td>
<td>18.9% (7x10^{-18})</td>
</tr>
<tr>
<td></td>
<td>Systemic Lupus Erythematosus</td>
<td>40.2% (2x10^{-2})</td>
<td>44.5% (2x10^{-2})</td>
<td>18.8% (6x10^{-30})</td>
<td>22.2% (8x10^{-27})</td>
</tr>
</tbody>
</table>
The prevalence of possible cases of any autoimmune disease was significantly higher in depression cases compared to depression controls ($p = 6 \times 10^{-177}$ for any depression versus controls, $p = 3 \times 10^{-207}$ for stringent depression versus controls). The prevalence of probable cases of any autoimmune disease was significantly higher in depression cases compared to depression controls ($p = 2 \times 10^{-124}$ for any depression versus controls, $p = 6 \times 10^{-163}$ for stringent depression versus controls) (Table 3).

Table 3: Prevalence of autoimmune diseases within depression cases compared to depression controls, stratified by possible/probable for autoimmune diseases, and stringent/any for depression cases. P-values from pairwise comparisons of autoimmune prevalence in depression cases compared to depression controls are shown in brackets.

<table>
<thead>
<tr>
<th>Autoimmune Trait</th>
<th>Possible autoimmune prevalence</th>
<th>Probable autoimmune prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Any Depression</td>
<td>Stringent Depression</td>
</tr>
<tr>
<td>Any Autoimmune Disease</td>
<td>10.48% ($6 \times 10^{-177}$)</td>
<td>14.27% ($3 \times 10^{-207}$)</td>
</tr>
<tr>
<td>Blood</td>
<td></td>
<td>6.59% ($2 \times 10^{-124}$)</td>
</tr>
<tr>
<td>Pernicious Anemia</td>
<td>0.77% ($4 \times 10^{-38}$)</td>
<td>1.17% ($1 \times 10^{-39}$)</td>
</tr>
<tr>
<td>Autoimmune Thyroid Disease</td>
<td>0.41% ($6 \times 10^{-39}$)</td>
<td>0.58% ($1 \times 10^{-39}$)</td>
</tr>
<tr>
<td>Type 1 diabetes</td>
<td>1.20% ($3 \times 10^{-35}$)</td>
<td>2.07% ($5 \times 10^{-35}$)</td>
</tr>
<tr>
<td>Multiple Sclerosis</td>
<td>0.87% ($6 \times 10^{-36}$)</td>
<td>1.31% ($5 \times 10^{-36}$)</td>
</tr>
<tr>
<td>Myasthenia Gravis</td>
<td>0.11% ($5 \times 10^{-5}$)</td>
<td>0.17% ($5 \times 10^{-5}$)</td>
</tr>
<tr>
<td>Coeliac</td>
<td>0.92% ($3 \times 10^{-33}$)</td>
<td>1.08% ($4 \times 10^{-34}$)</td>
</tr>
<tr>
<td>Inflammatory Bowel Disease</td>
<td>1.82% ($6 \times 10^{-33}$)</td>
<td>2.56% ($8 \times 10^{-34}$)</td>
</tr>
<tr>
<td>Psoriasis</td>
<td>2.05% ($2 \times 10^{-33}$)</td>
<td>2.55% ($4 \times 10^{-35}$)</td>
</tr>
<tr>
<td>Ankylosing Spondylitis</td>
<td>0.52% ($5 \times 10^{-1}$)</td>
<td>0.72% ($7 \times 10^{-1}$)</td>
</tr>
<tr>
<td>Polymyalgia Rheumatica/GCA</td>
<td>0.66% ($2 \times 10^{-13}$)</td>
<td>0.93% ($3 \times 10^{-13}$)</td>
</tr>
<tr>
<td>Psoriatic Arthritis</td>
<td>0.50% ($3 \times 10^{-33}$)</td>
<td>0.58% ($1 \times 10^{-33}$)</td>
</tr>
<tr>
<td>Rheumatoid Arthritis</td>
<td>2.66% ($5 \times 10^{-40}$)</td>
<td>4.14% ($3 \times 10^{-32}$)</td>
</tr>
<tr>
<td>Sjögren Syndrome</td>
<td>0.34% ($4 \times 10^{-29}$)</td>
<td>0.55% ($2 \times 10^{-29}$)</td>
</tr>
<tr>
<td>Systemic Lupus Erythematosus</td>
<td>0.36% ($2 \times 10^{-29}$)</td>
<td>0.59% ($6 \times 10^{-30}$)</td>
</tr>
</tbody>
</table>
Testing for same-trait PRS associations, PRS for MDD were significantly associated with any depression case-status ($p < 5 \times 10^{-324}, R^2 = 1.48\%$) and stringent depression case-status ($p = 2 \times 10^{-228}, R^2 = 2.23\%$). PRS for autoimmune diseases were significantly associated with both possible and probable case-control status for the corresponding diseases (Figure 2). The variance in liability, $R^2$, explained by PRS was higher in strictly-defined compared to liberally-defined phenotypes. Most results were highly significant ($p < 6 \times 10^{-29}$), except myasthenia gravis ($p < 7 \times 10^{-3}$), which has the smallest sample size of less than 250 cases, and psoriatic arthritis ($p < 3 \times 10^{-6}$) where the discovery GWAS has only 1430 cases.

![Autoimmune Variance](image)

**Figure 2:** Variances in autoimmune liability explained by PRS for the corresponding autoimmune diseases. The number of cases are shown at the top of the plot (possible = blue, probable = red). P-values are shown atop each bar.

In the prediction of depression from autoimmune PRS (Figure 3), the PRS for myasthenia gravis were significantly associated with case-status for any depression ($p = 5.2 \times 10^{-5}, R^2 = 0.01\%$) and stringent depression ($p = 1.6 \times 10^{-5}, R^2 = 0.04\%$). PRS for psoriasis were significantly associated with case-status for
any depression \((p = 8.7 \times 10^{-6}, R^2 = 0.01\%)\). No other autoimmune disease PRS predicted depression case-control status, and no sex-specific analyses met the Bonferroni-corrected threshold for significance. The \(R^2\) values for variance explained in depression by autoimmune PRS are all very low, at <0.1%, and substantially lower than the \(R^2\) for autoimmune diseases (Figure 2).

Figure 3: Variances in depression liability explained by PRS for autoimmune diseases (x-axis). Asterisks denote associations with \(p\)-values < 7.1\(\times\)10\(^{-5}\), meeting Bonferroni correction. Number of cases for depression phenotypes: Any (combined) = 65,075; Any (female) = 43,413; Any (male) = 21,662; Stringent (combined) = 14,625; Stringent (female) = 9,738; Stringent (male) = 4,887.

Genetic liability for MDD was significantly associated with six autoimmune diseases: coeliac, inflammatory bowel disease, psoriasis, psoriatic arthritis, rheumatoid arthritis, and type 1 diabetes (all \(p\)-values < 5.8\(\times\)10\(^{-5}\), \(R^2\) range between 0.06% and 0.27%) (Figure 4). For three, the association with MDD was observed in probable and possible cases (psoriasis, rheumatoid arthritis and type 1 diabetes). For coeliac and inflammatory bowel disease, the association was only in possible cases. For psoriatic arthritis,
the association was only in probable cases. For all significant associations observed, higher PRS increased risk for the outcome phenotype, except for coeliac, where higher MDD PRS was associated with reduced risk ($p = 6 \times 10^{-8}$, $R^2 = 0.17\%$, beta $=-0.11$, SE $= 0.02$, in the combined sample).

Sex-specific associations were observed, primarily in female autoimmune cases (coeliac, inflammatory bowel disease, type 1 diabetes and rheumatoid arthritis, all $p < 4.5 \times 10^{-5}$). Association in males was observed in psoriasis (possible cases, $p = 5.8 \times 10^{-5}$), and in rheumatoid arthritis (possible cases, $p = 1.6 \times 10^{-5}$). The most consistent results were observed in rheumatoid arthritis, where the sample size was largest, with five of the six analyses reaching Bonferroni threshold for significance (all $p < 4.5 \times 10^{-5}$, $R^2$ range between 0.07% and 0.1%). However, there was no evidence for a significant interaction between sex and PRS in the combined samples of men and women (all $p > 0.02$), indicating that sex-specific associations were generally influenced by sample size.

Full results of each test are shown in Supplementary Tables 2 to 5 and Supplementary Figures 1 to 4.
Figure 4: Variances in autoimmune liability (x-axes) explained by PRS for MDD. Asterisks denote associations with p-values < 7.1x10^{-5}, meeting Bonferroni correction. Number of cases for the autoimmune diseases are given in Table 1.

Power analyses showed that, in the prediction of any depression from autoimmune PRS, there was 80% power to detect associations assuming modest levels of underlying genetic correlation (rG) below 0.15. There were two exceptions; multiple sclerosis and myasthenia gravis, where the underlying rG would need to approach ~0.3 to achieve 80% power (Supplementary Figure 6). In the prediction of possible
autoimmune diseases from depression PRS, there was 80% power to detect associations assuming $r_G < 0.1$, again with the exception myasthenia gravis, where an $r_G \sim 0.18$ would be required to achieve 80% power (Supplementary Figure 7).
Discussion

Motivated by epidemiological findings of a bi-directional relationship between depression and autoimmune diseases, we tested for evidence of pleiotropy between traits, adopting both liberal and strict phenotyping approaches to define cases in the UKB. We showed modest association of PRS from autoimmune diseases with MDD, and slightly stronger associations of MDD PRS with autoimmune diseases. These observations suggest only a minor component of observed comorbidity is due to shared genetics between depression and autoimmune diseases.

We made three key observations: 1) Phenotypic variance explained by PRS for corresponding traits was higher in strictly-defined than liberally-defined cases, indicating more rigorous phenotyping improved the validity of autoimmune and depression cases; 2) The phenotypic overlap between depression and autoimmune diseases was consistent with the literature – depression was more common in individuals with autoimmune diseases, and vice-versa; 3) Cross-trait PRS analyses identified significant associations between depression and some autoimmune diseases, but with effect sizes indicating the existence of a shared biological component of modest effect on the observed comorbidity.

Our phenotyping approach for autoimmune diseases used both strictly-defined and liberally-defined cases, integrating the multiple sources of UKB data. PRS constructed for eight autoimmune diseases predicted case-control status, increasing confidence in the robustness of case definition. The phenotypic variance explained was higher in strictly-defined cases, potentially reflecting greater specificity; identifying individuals with multiple endorsements for a disease (e.g. hospital admission and prescription
medication) reduces the probability of misclassifying controls as cases. Conversely, the criteria for liberally-defined cases increases sample size, but may induce misclassification of controls as cases.

For each of the fourteen autoimmune diseases considered, cases had higher frequencies of depression than controls, recapitulating the effect observed in epidemiological studies. Similarly, the prevalence of each autoimmune disease was significantly higher in depression cases compared to controls. Prevalence estimates reported here are cross-sectional, and we lack information on the temporal relationship between traits.

The findings from cross-trait PRS analyses identified significant associations, although observed effect sizes were small, ranging between $R^2 = 0.01\%$ and 0.27\%. For three traits, we observed significant associations in liberally-defined, but not strictly-defined cases (psoriasis PRS was associated with any depression, MDD PRS was associated with possible coeliac and inflammatory bowel disease). In contrast, MDD PRS was associated with probable, but not possible, psoriatic arthritis, suggesting misclassification in possible cases. Misclassification bias may vary across diseases; some autoimmune diseases may be more prone to misclassification with other autoimmune diseases, whilst other diagnoses may misclassify with non-inflammatory conditions. For example, osteoarthritis (non-inflammatory) may misclassify as psoriatic arthritis in the absence of multiple-item endorsement to increase diagnosis validity.

Cross-trait PRS analyses identified some sex-dependent associations. MDD PRS were associated with psoriasis in males, and MDD PRS were associated with coeliac, inflammatory bowel disease and type 1 diabetes in females. However, sensitivity analyses revealed no evidence for significant interactions between PRS and sex in the combined sample, indicating that sex-dependent associations were generally
driven by different sample sizes in sex-stratified analyses. Rheumatoid arthritis was the most common autoimmune disease and showed the most consistency in cross-trait associations; MDD PRS were significantly associated with rheumatoid arthritis in all case groups, except probable males. This is in contrast with Euesden, et al.\textsuperscript{10}, who found no evidence for association between PRS for depression and risk for rheumatoid arthritis, but in a substantially smaller sample of 226 cases. Liu, et al.\textsuperscript{6} also found no evidence for association between a composite mental health disorder PRS and risk for autoimmune diseases, but also in a smaller sample of 1,383 individuals with any of seven autoimmune diseases. However, a composite PRS for autoimmune diseases did show weak association with case-control status in a sample of 43,902 individuals with any of six mental health disorders in the Liu, et al.\textsuperscript{6} study. This highlights the importance of sample size, and our study benefits from the scale of the UKB, where power calculations indicated our investigation was able to detect even modest pleiotropic effects. However, there were exceptions for rare autoimmune diseases, where larger samples would be required to reject the presence of a weak genetic correlation with depression.

The weak genetic contribution suggests that another mechanism may be driving or contributing to the bi-directional relationship between autoimmune diseases and depression\textsuperscript{44}. Inflammatory factors underlying some cases of depression could provide a common biological pathogenesis with autoimmune diseases. Lynall, et al.\textsuperscript{45} observed increased immune cell counts in depression cases compared to controls, and identified a sub-group of cases with elevated inflammatory markers who presented with more severe depression than uninflamed cases. In our data, we considered whether $R^2$ values mirrored the inflammatory burden of autoimmune diseases, recognizing that some diagnoses have much higher levels of systemic inflammation (e.g. IBD or RA) than others (e.g. Coeliac). Although subtle differences
were present, in absolute terms variation was small and it was not possible to infer a ‘dose-response’
effect between autoimmune disease PRS and depression.

Environmental risk factors such as BMI and childhood maltreatment increase risk of both depression and
autoimmune diseases and would contribute to the bi-directional effect\(^ {46,47}\). Similarly, some treatments
for depression (antidepressants) and autoimmune diseases (steroids) are obesogenic and may increase
comorbidity. Diagnosis with autoimmune disease increases risk of depression due to psychological
factors in adjusting to a chronic disorder and changes in behaviour such as reduced exercise. Health
related behaviours that are elevated in depression (smoking, poor diet and reduced physical activity)
increase risk for autoimmune diseases. These mechanisms may not be independent of joint genetic
contributors. For example, shared inflammatory mechanisms would lead to horizontal pleiotropy, where
genetic variants directly affect both disorders, and vertical pleiotropy can arise through environmental
risk factors where genetic variation influences one trait through mediation on another trait.

Limitations

A healthy volunteer bias has been observed in the UKB\(^ {48}\), and is a noted limitation of the study. However,
it has been proposed that this bias may attenuate, but not invalidate, exposure-outcome relationships\(^ {49}\).
A further limitation of the ability to extrapolate our results is the lack of representation in individuals of
diverse ancestries. The literature has demonstrated attenuation in PRS analyses where training and
target samples are drawn from different ancestral populations\(^ {50}\), highlighting the need to perform GWAS
in diverse ancestries. This limitation may have broader implications than would otherwise be the case
for some conditions, such as SLE, which disproportionately affect individuals of African and Asian ancestry.

Although every effort has been made to address the potential for misclassification bias through the criteria for multiple-item endorsements in strictly-defined cases, the approach remains imperfect. The impact of misclassification depends on the effect directionality, and it is possible that mis-estimation of associations has occurred. However, it is our contention that even misclassification of moderate size would have only marginal impacts on the inferences made above.

Conclusions

We identified cases and controls for depression and fourteen autoimmune diseases in the UKB, using both strict and liberal phenotyping. PRS analyses indicated that strict phenotyping improved the validity of cases, demonstrating that multiple UKB variables can be leveraged to increase specificity. Consistent with the literature, we found that depression was enriched in autoimmune cases, and vice-versa. Despite having power to detect subtle pleiotropic effects, we found little evidence that shared genetic factors have a meaningful influence on the observed co-occurrence of depression and autoimmune diseases in the UK Biobank. The limited shared genetic component will make only a modest contribution to the bi-directional disease risks, and shared environmental factors, including health-related characteristics and stressful life events, may be important. Future studies leveraging phenotypic, genetic, diagnostic, treatment and environmental risk factors may be necessary to unpick the mechanisms contributing to shared risks for autoimmune diseases and depression. In particular, future research should consider the
psychological impacts of autoimmune disease while remaining cognizant of the need to consider and treat the two diseases in parallel.
Funding

This work was supported by the UK Medical Research Council (PhD studentship to KPG; grant MR/N015746/1). This paper represents independent research part-funded by the National Institute for Health Research (NIHR) Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King’s College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care.

Acknowledgements

We thank participants and scientists involved in making the UK Biobank resource available (http://www.ukbiobank.ac.uk/). The UKB received ethical approval from the North West - Haydock Research Ethics Committee (reference 16/NW/0274). This study was conducted under application number 18177. We thank the research participants and employees of 23andMe for making this work possible. The MDD GWAS summary statistics results from 23andMe were available through a Data Transfer Agreement between 23andMe, Inc., and King’s College, London. Only summary statistics were shared with no individual level data. 23andMe participants provided informed consent and participated in the research online. The 23andMe protocol was approved by an external Association for the Accreditation of Human Research Protection Programs accredited Institutional Review Board, Ethical and Independent Review Services. Participants were included in the analysis on the basis of consent status as checked at the time data analyses were initiated. Statistical analyses were carried out on the King’s Health Partners High Performance Compute Cluster funded with capital equipment grants from the GSTT Charity (TR130505) and Maudsley Charity (980). We thank Nick Dand, Satveer Mahil and
Catherine Smith of King’s College London for their contribution in identifying medications used in the treatment of Psoriasis in the UKB.

Data Availability

Available from UK Biobank subject to standard procedures (www.ukbiobank.ac.uk). The full GWAS summary statistics for the 23andMe discovery data set will be made available through 23andMe to qualified researchers under an agreement with 23andMe that protects the privacy of the 23andMe participants. Please visit https://research.23andme.com/collaborate/#publication for more information and to apply to access the data.

Author contributions

Conceptualisation and study design: KPG, CML, JG, PFO. Analysis and manuscript: KPG. Analytical consultation and interpretation: CML, JG, PFO, JRIC. UKB data curation and management: JRIC, KPG. Genetic data preparation: JRIC, KPG. Project supervisors: CML, JG, PFO. All authors critically edited the paper.

Declaration of Interest

CML is a member of the SAB for Myriad Neuroscience. The remaining authors declare no competing interests.
References


