medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

Kylie P Glanville[1], Jonathan R I Coleman[1,2], Paul F O'Reilly[3], *James Galloway[4], *Cathryn M Lewis[1,2,5]

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NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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 ($p < 5.2\times10^{-5}$, $R^2 <= 0.04\%$). PRS for depression were significantly associated with case-status for coeliac disease, inflammatory bowel disease, psoriasis, psoriatic arthritis, rheumatoid arthritis and type 1 diabetes ($p < 5.8\times10^{-5}$, R^2 range 0.06% to 0.27%).

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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 (rG) between depression and nine autoimmune diseases; the strongest correlation observed was between depression and inflammatory bowel disease (rG = .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.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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[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.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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[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, 10th Revision[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[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[13].

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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.

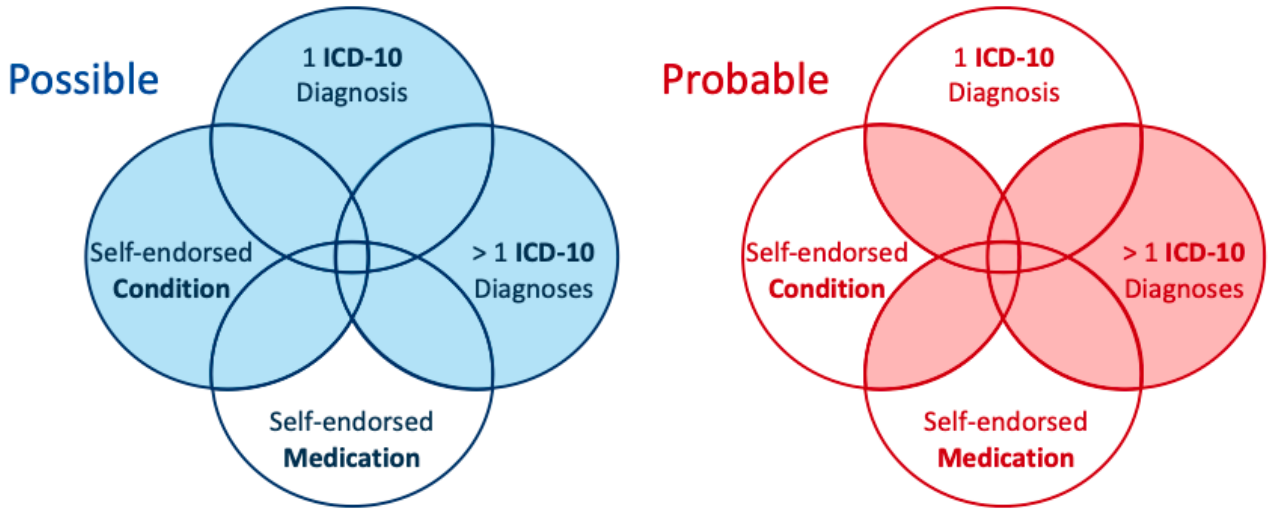

Figure 1: Autoimmune phenotyping approach. Cases are included in possible or probable if they fall within a shaded area. Autoimmune medication was used as a confirmatory, but not a primary source of information, because several medications are not disease-specific.

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[16]. We therefore classified

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

'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.[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[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)[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.[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.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Detail on the derivation of each indication of depression, schizophrenia and bipolar can be found in Supplementary Materials from our previous publication[16].

Genetic quality control (QC)

The UKB performed preliminary QC on genotype data assayed for all participants[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[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 ($p < 10e-8$), 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 $r^2 > 0.044$)[19] and the GreedyRelated[20] algorithm used to remove one individual from each pair, preferentially retaining individuals that survived QC. FlashPCA2[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[22]. We tested for significant differences in: 1) the prevalence of depression in

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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[22].

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: coeliac[23], IBD[24], MS[25], MG[26], psoriasis[27], PsA[28], RA[29], and SLE[30] (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.[11], 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[31]. 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).

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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 7.1×10^{-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%[32], ATD=2%[33], T1D=0.3%[34], MS=0.1%[25], MG=0.02%[35], coeliac=1%[23], IBD=0.5%[24], psoriasis=2%[36], AS=0.55%[37], PR/GCA= 0.85%[38], PsA=0.5%[28], RA=1%[39], SS=0.7%[40], SLE=0.1%[41] and MDD=15%[11].

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

AVENGEME[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[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.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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 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.

		Pop. Prev.	Count	UKB Prev.	Mean age (SD)	Female (%)	TDI (SD)	Current smoker (%)	Mean BMI (SD)	Count	UKB Prev.	Mean age (SD)	Female (%)	TDI (SD)	Current smoker (%)	Mean BMI (SD)
Autoimmune Diseases (system/disease)			**Possible**							**Probable**						
Blood	Pernicious Anemia	0.10%	1,555	0.48%	58.9 (7.61)	71%	-0.85 (3.19)	13%	28.0 (5.55)	423	0.13%	60.1 (7.19)	72%	-0.99 (3.23)	12%	28.5 (5.34)
Endocrine system	Autoimmune Thyroid Disease	2.00%	859	0.26%	56.8 (7.75)	85%	-1.11 (3.05)	15%	27.4 (5.11)	607	0.19%	57.0 (7.64)	86%	-1.06 (3.03)	16%	27.8 (5.21)
	Type 1 diabetes	0.30%	2,751	0.84%	58.3 (7.80)	42%	-0.46 (3.39)	13%	30.1 (6.08)	2,292	0.70%	58.0 (7.85)	43%	-0.50 (3.39)	12%	29.9 (6.07)
Nervous system	Multiple Sclerosis	0.10%	1,683	0.52%	55.4 (7.52)	73%	-1.45 (3.01)	16%	26.9 (5.08)	1,154	0.35%	55.5 (7.45)	73%	-1.33 (3.06)	17%	26.9 (5.27)
	Myasthenia Gravis	0.02%	234	0.07%	59.2 (7.39)	56%	-0.94 (3.27)	11%	29.0 (5.44)	147	0.05%	60.1 (7.09)	48%	-1.06 (3.32)	12%	29.3 (5.18)
Digestive system	Coeliac	1.00%	2,364	0.72%	57.8 (7.79)	67%	-1.49 (2.99)	7%	25.8 (4.57)	1,260	0.39%	58.4 (7.66)	68%	-1.42 (3.03)	7%	25.8 (4.60)
	Inflammatory Bowel Disease	0.50%	5,105	1.55%	57.4 (7.92)	51%	-1.27 (3.10)	10%	27.3 (4.70)	3,538	1.08%	57.3 (7.98)	50%	-1.28 (3.07)	9%	27.2 (4.63)
Skin	Psoriasis	2.00%	5,459	1.66%	56.8 (8.01)	46%	-1.05 (3.22)	15%	28.4 (5.17)	2,759	0.84%	57.1 (8.04)	42%	-0.90 (3.28)	16%	28.7 (5.26)
Musculoskeletal system and connective tissue	Ankylosing Spondylitis	0.55%	1,344	0.41%	57.9 (7.47)	38%	-1.02 (3.21)	13%	27.8 (4.80)	413	0.13%	57.4 (7.57)	26%	-1.06 (3.22)	12%	27.9 (4.85)
	Polymyalgia Rheumatica/GCA	0.85%	1,627	0.50%	63.1 (5.16)	67%	-1.70 (2.79)	9%	28.3 (5.08)	898	0.28%	63.7 (4.58)	68%	-1.67 (2.75)	9%	28.4 (5.24)
	Psoriatic Arthritis	0.50%	1,107	0.34%	56.6 (7.48)	50%	-1.16 (3.18)	11%	29.0 (5.43)	779	0.24%	56.6 (7.52)	51%	-1.06 (3.23)	11%	29.3 (5.59)
	Rheumatoid Arthritis	1.00%	6,360	1.92%	59.5 (7.06)	67%	-0.91 (3.29)	13%	28.5 (5.53)	3,451	1.05%	59.6 (7.01)	70%	-1.08 (3.18)	12%	28.2 (5.50)
	Sjögren Syndrome	0.70%	647	0.20%	59.3 (7.10)	90%	-1.20 (3.03)	6%	27.1 (5.72)	389	0.12%	59.2 (7.26)	89%	-1.19 (3.10)	7%	27.0 (5.55)
	Systemic Lupus Erythematosus	0.10%	624	0.19%	56.7 (8.18)	84%	-1.02 (3.21)	14%	27.3 (5.65)	362	0.11%	56.6 (8.00)	86%	-0.99 (3.15)	14%	27.3 (5.82)
	Any Autoimmune Disease	NA	28,479	8.08%	58.1 (7.73)	58%	-1.11 (3.17)	12%	27.9 (5.28)	16,824	4.94%	58.1 (7.73)	57%	-1.10 (3.17)	12%	28.0 (5.34)
	Autoimmune Controls	NA	324,074	NA	56.4 (8.06)	52%	-1.50 (2.98)	10%	27.2 (4.64)	324,074	NA	56.4 (8.06)	52%	-1.50 (2.98)	10%	27.2 (4.64)
Depression			**Any**							**Stringent**						
	Depression	15%	65,075	21.86%	55.4 (7.86)	67%	-1.09 (3.14)	13%	27.8 (5.32)	14,625	5.92%	56.1 (7.88)	67%	-0.58 (3.35)	19%	28.9 (5.77)
	Depression Controls	NA	232,552	NA	57.1 (8.10)	47%	-1.65 (2.89)	9%	27.2 (4.53)	232,552	NA	57.1 (8.10)	47%	-1.65 (2.89)	9%	27.2 (4.53)

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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.

		Depression prevalence in autoimmune traits			
		Any depression prevalence		Stringent depression prevalence	
Autoimmune Trait		Possible Autoimmune	Probable Autoimmune	Possible Autoimmune	Probable Autoimmune
Autoimmune Controls		**20.6%**		**5.1%**	
Any Autoimmune Disease		**28.9% ($6x10^{-177}$)**	**29.5% ($2x10^{-124}$)**	**10.8% ($3x10^{-207}$)**	**11.5% ($6x10^{-163}$)**
Blood	Pernicious Anemia	35.8% ($4x10^{-36}$)	33.7% ($4x10^{-8}$)	15.2% ($1x10^{-39}$)	16.6% ($4x10^{-15}$)
Endocrine system	Autoimmune Thyroid Disease	35.1% ($6x10^{-19}$)	34.7% ($6x10^{-13}$)	13.7% ($1x10^{-16}$)	13.4% ($3x10^{-11}$)
	Type 1 diabetes	31.1% ($3x10^{-31}$)	31.3% ($1x10^{-27}$)	14.1% ($5x10^{-59}$)	13.5% $5x10^{-43}$)
Nervous System	Multiple Sclerosis	39.3% ($6x10^{-56}$)	42.5% ($3x10^{-51}$)	16.9% ($5x10^{-54}$)	20.0% ($8x10^{-57}$)
	Myasthenia Gravis	33.3% ($5x10^{-5}$)	29.2% ($3x10^{-2}$)	14.1% ($5x10^{-6}$)	14.8% ($4x10^{-5}$)
Digestive system	Coeliac	27.3% ($3x10^{-12}$)	29.8% ($2x10^{-12}$)	8.4% ($4x10^{-8}$)	9.5% ($7x10^{-8}$)
	Inflammatory Bowel Disease	24.9% ($6x10^{-11}$)	24.9% ($3x10^{-8}$)	8.9% ($8x10^{-22}$)	9.1% ($2x10^{-17}$)
Skin	Psoriasis	26.8% ($2x10^{-22}$)	28.3% ($1x10^{-17}$)	8.7% ($4x10^{-20}$)	9.9% ($4x10^{-18}$)
Musculoskeletal system and connective tissue	Ankylosing Spondylitis	26.9% ($5x10^{-7}$)	28.4% ($5x10^{-4}$)	9.6% ($7x10^{-9}$)	10.5% ($9x10^{-5}$)
	Polymyalgia Rheumatica/GCA	29.3% ($2x10^{-13}$)	27.8% ($5x10^{-6}$)	10.8% ($3x10^{-15}$)	10.6% ($1x10^{-8}$)
	Psoriatic Arthritis	31.8% ($3x10^{-15}$)	34.3% ($5x10^{-16}$)	10.2% ($1x10^{-8}$)	13.0% ($2x10^{-13}$)
	Rheumatoid Arthritis	30.6% ($5x10^{-62}$)	29.4% ($6x10^{-28}$)	12.7% ($3x10^{-92}$)	12.1% ($1x10^{-45}$)
	Sjögren Syndrome	40.7% ($4x10^{-25}$)	41.9% ($1x10^{-17}$)	18.9% ($2x10^{-28}$)	18.9% ($7x10^{-18}$)
	Systemic Lupus Erythematosus	40.2% ($2x10^{-25}$)	44.5% ($2x10^{-22}$)	18.8% ($6x10^{-30}$)	22.2% ($8x10^{-27}$)

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The prevalence of possible cases of any autoimmune disease was significantly higher in depression cases compared to depression controls ($p = 6x10^{-177}$ for any depression versus controls, $p = 3x10^{-207}$ for stringent depression versus controls). The prevalence of probable cases of any autoimmune disease was significantly higher in depression cases compared to depression controls ($p = 2x10^{-124}$ for any depression versus controls, $p = 6x10^{-163}$ 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.

		Autoimmune prevalence in depression					
		Possible autoimmune prevalence			Probable autoimmune prevalence		
Autoimmune Trait		Any Depression	Stringent Depression	Depression Controls	Any Depression	Stringent Depression	Depression Controls
	Any Autoimmune Disease	**10.48%** ($6x10^{-177}$)	**14.27%** ($3x10^{-207}$)	**6.94%**	**6.59%** ($2x10^{-124}$)	**9.48%** ($6x10^{-163}$)	**4.19%**
Blood	Pernicious Anemia	0.77% ($4x10^{-36}$)	1.17% ($1x10^{-39}$)	0.36%	0.19% ($4x10^{-8}$)	0.35% ($4x10^{-15}$)	0.10%
Endocrine system	Autoimmune Thyroid Disease	0.41% ($6x10^{-19}$)	0.58% ($1x10^{-16}$)	0.20%	0.28% ($6x10^{-13}$)	0.39% ($3x10^{-11}$)	0.14%
	Type 1 diabetes	1.20% ($3x10^{-31}$)	2.07% ($5x10^{-59}$)	0.69%	1.02% ($1x10^{-27}$)	1.65% ($5x10^{-43}$)	0.58%
Nervous System	Multiple Sclerosis	0.87% ($6x10^{-56}$)	1.31% ($5x10^{-54}$)	0.35%	0.63% ($3x10^{-51}$)	1.02% ($8x10^{-57}$)	0.22%
	Myasthenia Gravis	0.11% ($5x10^{-5}$)	0.17% ($5x10^{-6}$)	0.05%	0.06% ($3x10^{-2}$)	0.12% ($4x10^{-5}$)	0.04%
Digestive system	Coeliac	0.92% ($3x10^{-12}$)	1.08% ($4x10^{-8}$)	0.64%	0.54% ($2x10^{-12}$)	0.64% ($7x10^{-8}$)	0.33%
	Inflammatory Bowel Disease	1.82% ($6x10^{-11}$)	2.56% ($8x10^{-22}$)	1.43%	1.28% ($3x10^{-8}$)	1.85% ($2x10^{-17}$)	1.00%
Skin	Psoriasis	2.05% ($2x10^{-22}$)	2.55% ($4x10^{-20}$)	1.46%	1.10% ($1x10^{-17}$)	1.47% ($4x10^{-18}$)	0.73%
Musculoskeletal system and connective tissue	Ankylosing Spondylitis	0.52% ($5x10^{-7}$)	0.72% ($7x10^{-9}$)	0.37%	0.17% ($5x10^{-4}$)	0.25% ($9x10^{-5}$)	0.11%
	Polymyalgia Rheumatica/GCA	0.66% ($2x10^{-13}$)	0.93% ($3x10^{-15}$)	0.42%	0.35% ($5x10^{-6}$)	0.51% ($1x10^{-8}$)	0.23%
	Psoriatic Arthritis	0.50% ($3x10^{-15}$)	0.58% ($1x10^{-8}$)	0.28%	0.38% ($5x10^{-16}$)	0.51% ($2x10^{-13}$)	0.19%
	Rheumatoid Arthritis	2.66% ($5x10^{-62}$)	4.14% ($3x10^{-92}$)	1.59%	1.42% ($6x10^{-28}$)	2.23% ($1x10^{-45}$)	0.89%
	Sjögren Syndrome	0.34% ($4x10^{-25}$)	0.55% ($2x10^{-28}$)	0.13%	0.21% ($1x10^{-17}$)	0.33% ($7x10^{-18}$)	0.08%
	Systemic Lupus Erythematosus	0.36% ($2x10^{-25}$)	0.59% ($6x10^{-30}$)	0.14%	0.23% ($2x10^{-22}$)	0.39% ($8x10^{-27}$)	0.08%

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Testing for same-trait PRS associations, PRS for MDD were significantly associated with any depression case-status ($p < 5x10^{-324}$, R^2 = 1.48%,) and stringent depression case-status ($p = 2x10^{-228}$, 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 ($p < 6x10^{-29}$), except myasthenia gravis ($p < 7x10^{-3}$), 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.

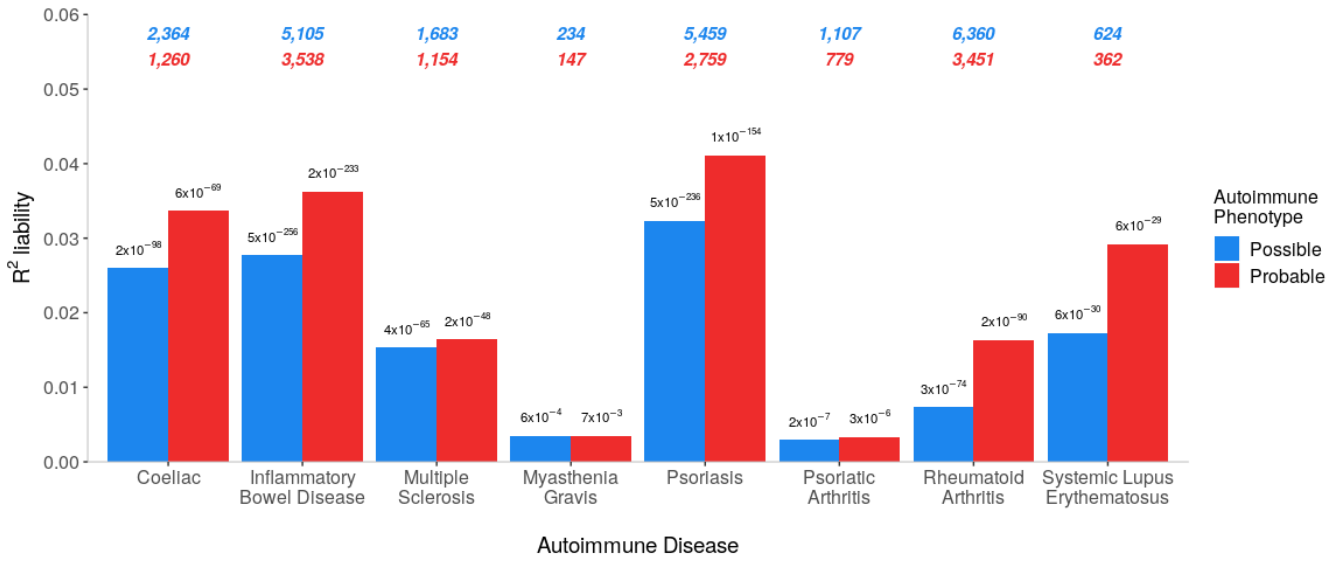

Figure 2: Variances in autoimmune liability explained by PRS for the corresponding autoimmune diseases. The number of cases are shown at the top of the plot (possible = blue, probable = red). P-values are shown atop each bar.

In the prediction of depression from autoimmune PRS (Figure 3), the PRS for myasthenia gravis were significantly associated with case-status for any depression ($p = 5.2x10^{-5}$, R^2 = 0.01%) and stringent depression ($p = 1.6x10^{-5}$, R^2 = 0.04%). PRS for psoriasis were significantly associated with case-status for

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

any depression (p = 8.7×10^{-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).

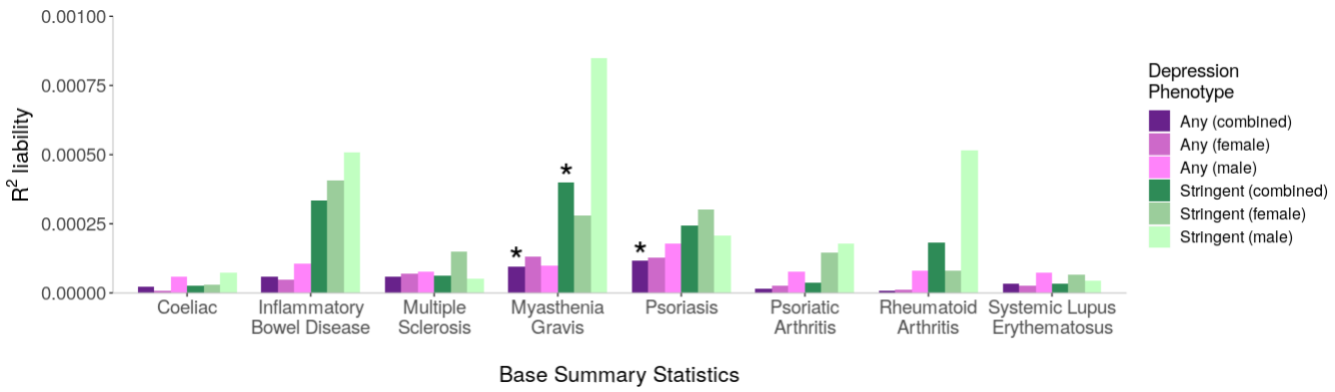

Figure 3: Variances in depression liability explained by PRS for autoimmune diseases (x-axis). Asterisks denote associations with p-values < 7.1×10^{-5}, meeting Bonferroni correction. Number of cases for depression phenotypes: Any (combined) = 65,075; Any (female) = 43,413; Any (male) = 21,662; Stringent (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 < 5.8×10^{-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,

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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 ($p = 6x10^{-8}$, $R^2 = 0.17\%$, 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 $p < 4.5x10^{-5}$). Association in males was observed in psoriasis (possible cases, $p = 5.8x10^{-5}$), and in rheumatoid arthritis (possible cases, $p = 1.6x10^{-5}$). 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 $p < 4.5x10^{-5}$, 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.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

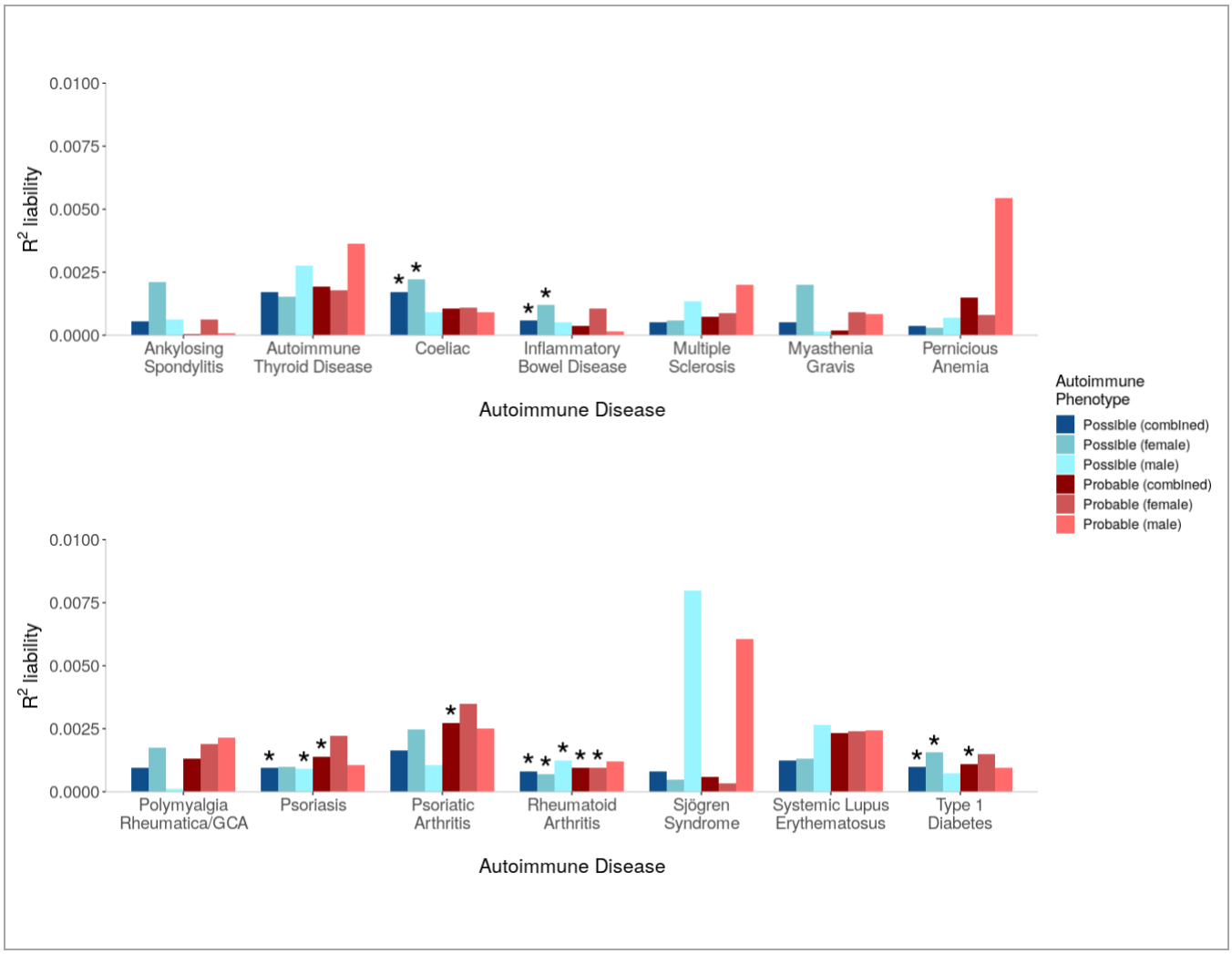

Figure 4: Variances in autoimmune liability (x-axes) explained by PRS for MDD. Asterisks denote associations with p-values < 7.1×10^{-5}, meeting Bonferroni correction. Number of cases for the autoimmune diseases are given in Table 1.

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

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

autoimmune diseases from depression PRS, there was 80% power to detect associations assuming rG <0.1, again with the exception myasthenia gravis, where an rG ~0.18 would be required to achieve 80% power (Supplementary Figure 7).

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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.[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.[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.[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[44]. Inflammatory factors underlying some cases of depression could provide a common biological pathogenesis with autoimmune diseases. Lynall, et al.[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

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

psychological impacts of autoimmune disease while remaining cognizant of the need to consider and treat the two diseases in parallel.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

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.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

References

1. Benros ME, Waltoft BL, Nordentoft M, Ostergaard SD, Eaton WW, Krogh J, et al. Autoimmune Diseases and Severe Infections as Risk Factors for Mood Disorders A Nationwide Study. JAMA Psychiatry. American Medical Association; 2013 Aug;70(8):812–20.

2. Anderson RJ, Freedberg KA, Clouse RE, Lustman PJ. The Prevalence of Comorbid Depression in Adults With Diabetes. Diabetes Care. 2001 Jun;24(6):1069–78.

3. Patten SB, Beck CA, Williams JVA, Barbui C, Metz LM. Major depression in multiple sclerosis: A population-based perspective. Neurology. 61st ed. 2003 Dec;:1524–7.

4. Kurina L, Goldacre M, Yeates D, Gill L. Depression and anxiety in people with inflammatory bowel disease. Journal of Epidemiology and Community Health. BMJ Publishing Group; 2001 Oct 1;55(10):716–20.

5. Andersson NW, Gustafsson LN, Okkels N, Taha F, Cole SW, Munk-Jorgensen P, et al. Depression and the risk of autoimmune disease: a nationally representative, prospective longitudinal study. Psychol Med. Cambridge University Press; 2015 Dec;45(16):3559–69.

6. Liu X, Nudel R, Thompson WK, Appadurai V, Schork AJ, Buil A, et al. Genetic factors underlying the bidirectional relationship between autoimmune and mental disorders – findings from a Danish population-based study. Brain, Behavior, and Immunity. Academic Press; 2020 Jun 11.

7. Glanville KP, Coleman JRI, Hanscombe KB, Euesden J, Choi SW, Purves KL, et al. Classical Human Leukocyte Antigen Alleles and C4 Haplotypes Are Not Significantly Associated With Depression. Biological Psychiatry. Elsevier; 2020 Mar 1;87(5):419–30.

8. Nudel R, Benros ME, Krebs MD, Allesøe RL, Lemvigh CK, Bybjerg-Grauholm J, et al. Immunity and mental illness: findings from a Danish population-based immunogenetic study of seven psychiatric and neurodevelopmental disorders. European Journal of Human Genetics. Nature Publishing Group; 2019 Apr 11;8(1):1.

9. Hu X, Daly M. What have we learned from six years of GWAS in autoimmune diseases, and what is next? Current Opinion in Immunology. Elsevier Current Trends; 2012 Oct 1;24(5):571–5.

10. Euesden J, Danese A, Lewis CM, Maughan B. A bidirectional relationship between depression and the autoimmune disorders - New perspectives from the National Child Development Study. PLOS ONE. Public Library of Science; 2017 Mar;12(3):1–14.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

11. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A, et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nature Genetics. Nature Publishing Group; 2018 May;50:668–81.

12. Pouget JG, Schizophrenia Working Group of the Psychiatric Genomics Consortium, Han B, Wu Y, Mignot E, Ollila HM, et al. Cross-disorder analysis of schizophrenia and 19 immune-mediated diseases identifies shared genetic risk. Human Molecular Genetics. 2019 Jun 18.

13. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. Nature. Nature Publishing Group; 2018 Oct 1;562(7726):203–9.

14. World Health Organization. International Classification of Diseases (10th Edition). Geneva: World Health Organisation; 1992.

15. Davis KAS, Coleman JRI, Adams M, Allen N, Breen G, Cullen B, et al. Mental health in UK Biobank – development, implementation and results from an online questionnaire completed by 157 366 participants: a reanalysis. BJPsych Open. Cambridge University Press; 2020 Mar 1;6(2):e18.

16. Glanville KP, Coleman JR, Howard DM, Pain O, Hanscombe KB, Jermy B, et al. Multiple measures of depression to enhance validity of Major Depressive Disorder in the UK Biobank. medRxiv. Cold Spring Harbor Laboratory Press; 2020 Sep 22;:2020.09.18.20196451.

17. Smith DJ, Nicholl BI, Cullen B, Martin D, Ul-Haq Z, Evans J, et al. Prevalence and Characteristics of Probable Major Depression and Bipolar Disorder within UK Biobank: Cross-Sectional Study of 172,751 Participants. Potash JB, editor. PLOS ONE. 2013 Nov 25;8(11):e75362.

18. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience [Internet]. 2015 Feb 25;4(1):559. Available from: www.cog-genomics.org/plink/1.9/

19. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. Robust relationship inference in genome-wide association studies. Bioinformatics. 2010 Nov 15;26(22):2867–73.

20. Choi SW. choishingwan/GreedyRelated: Function update. 2020 Mar 4.

21. Abraham G, Qiu Y, Inouye M. FlashPCA2: principal component analysis of Biobank-scale genotype datasets. Bioinformatics. 2017 Sep 1;33(17):2776–8.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

22. R Development Core Team. R: A language and environment for statistical computing [Internet]. Vienna, Austria: R Foundation for Statistical Computing; 2008. Available from: http://www.R-project.org

23. Dubois PCA, Trynka G, Franke L, Hunt KA, Romanos J, Curtotti A, et al. Multiple common variants for celiac disease influencing immune gene expression. Nature Genetics. Nature Publishing Group; 2010 Apr 1;42(4):295–302.

24. Liu JZ, van Sommeren S, Huang H, Ng SC, Alberts R, Takahashi A, et al. Association analyses identify 38 susceptibility loci for inflammatory bowel disease and highlight shared genetic risk across populations. Nature Genetics. Nature Publishing Group; 2015 Sep 1;47(9):979–86.

25. The International Multiple Sclerosis Genetics Consortium & The Wellcome Trust Case Control Consortium 2. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. Nature. Nature Publishing Group; 2011 Aug 1;476(7359):214–9.

26. Renton AE, Pliner HA, Provenzano C, Evoli A, Ricciardi R, Nalls MA, et al. A Genome-Wide Association Study of Myasthenia Gravis. JAMA neurology. JAMA Neurol; 2015 Apr 1;72(4):396–404.

27. Tsoi LC, Stuart PE, Tian C, Gudjonsson JE, Das S, Zawistowski M, et al. Large scale meta-analysis characterizes genetic architecture for common psoriasis associated variants. Nature Communications. Nature Publishing Group; 2017 May 24;8(1):1–8.

28. Aterido A, Cañete JD, Tornero J, Ferrándiz C, Pinto JA, Gratacós J, et al. Genetic variation at the glycosaminoglycan metabolism pathway contributes to the risk of psoriatic arthritis but not psoriasis. Annals of the Rheumatic Diseases. BMJ Publishing Group Ltd; 2019 Mar 1;78(3):355–64.

29. Okada Y, Di Wu, Trynka G, Raj T, Terao C, Ikari K, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. Nature. Nature Publishing Group; 2014 Feb 1;506(7488):376–81.

30. Bentham J, Morris DL, Graham DSC, Pinder CL, Tombleson P, Behrens TW, et al. Genetic association analyses implicate aberrant regulation of innate and adaptive immunity genes in the pathogenesis of systemic lupus erythematosus. Nature Genetics. Nature Publishing Group; 2015 Dec 1;47(12):1457–64.

31. Choi SW, O'Reilly PF. PRSice-2: Polygenic Risk Score software for biobank-scale data. - PubMed - NCBI. GigaScience. 2019 Jul 15;8(7):2091.

32. Andres E, Serraj K. Optimal management of pernicious anemia. Journal of blood medicine. J Blood Med; 2012;3:97–103.

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

33. Simmonds MJ, Gough SCL. Unravelling the genetic complexity of autoimmune thyroid disease: HLA, CTLA-4 and beyond. Clinical and Experimental Immunology. Wiley-Blackwell; 2004 Apr 1;136(1):1–10.

34. Bradfield JP, Qu H-Q, Wang K, Zhang H, Sleiman PM, Kim CE, et al. A Genome-Wide Meta-Analysis of Six Type 1 Diabetes Cohorts Identifies Multiple Associated Loci. McCarthy MI, editor. PLoS Genet. Public Library of Science; 2011 Sep 29;7(9):e1002293.

35. Spillane J, Higham E, Kullmann DM. Myasthenia gravis. BMJ. British Medical Journal Publishing Group; 2012 Dec 21;345(dec21 3):e8497–7.

36. Tsoi LC, Stuart PE, Tian C, Gudjonsson JE, Das S, Zawistowski M, et al. Large scale meta-analysis characterizes genetic architecture for common psoriasis associated variants. - PubMed - NCBI. Nature Communications. 2017 May 24;8:15382.

37. International Genetics of Ankylosing Spondylitis Consortium (IGAS). Identification of multiple risk variants for ankylosing spondylitis through high-density genotyping of immune-related loci. Nature Genetics. Nature Publishing Group; 2013 Jul 1;45(7):730–8.

38. Partington RJ, Muller S, Helliwell T, Mallen CD, Sultan AA. Incidence, prevalence and treatment burden of polymyalgia rheumatica in the UK over two decades: a population-based study. Annals of the Rheumatic Diseases. BMJ Publishing Group Ltd; 2018 Dec 1;77(12):1750–6.

39. Humphreys JH, Verstappen SMM, Hyrich KL, Chipping JR, Marshall T, Symmons DPM. The incidence of rheumatoid arthritis in the UK: comparisons using the 2010 ACR/EULAR classification criteria and the 1987 ACR classification criteria. Results from the Norfolk Arthritis Register. Annals of the Rheumatic Diseases. BMJ Publishing Group Ltd; 2013 Aug 1;72(8):1315–20.

40. for UK Primary Sjögren's Syndrome Registry, Lessard CJ, Li H, Adrianto I, Ice JA, Rasmussen A, et al. Variants at multiple loci implicated in both innate and adaptive immune responses are associated with Sjögren's syndrome. Nature Genetics. Nat Genet; 2013 Oct 6;45(11):1284–92.

41. Rees F, Doherty M, Grainge M, Davenport G, Lanyon P, Zhang W. The incidence and prevalence of systemic lupus erythematosus in the UK, 1999–2012. Annals of the Rheumatic Diseases. BMJ Publishing Group Ltd; 2016 Jan 1;75(1):136–41.

42. Luigi Palla FD. A Fast Method that Uses Polygenic Scores to Estimate the Variance Explained by Genome-wide Marker Panels and the Proportion of Variants Affecting a Trait. Am J Hum Genet. Elsevier; 2015 Aug 6;97(2):250–9.

43. Bulik-Sullivan BK, Loh P-R, Finucane HK, Ripke S, Yang J, Schizophrenia Working Group of the Psychiatric Genomics Consortium, et al. LD Score regression distinguishes

medRxiv preprint doi: https://doi.org/10.1101/2020.12.08.20242495; this version posted December 9, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

confounding from polygenicity in genome-wide association studies. Nature Genetics. 2015 Mar;47(3):291–5.

44. Gold SM, Köhler-Forsberg O, Moss-Morris R, Mehnert A, Miranda JJ, Bullinger M, et al. Comorbid depression in medical diseases. Nat Rev Dis Primers. Nature Publishing Group; 2020 Aug 20;6(1):1–22.

45. ME L, L T, J B, J C, de Boer P, Mondelli V, et al. Peripheral Blood Cell-Stratified Subgroups of Inflamed Depression. Biological Psychiatry. Biol Psychiatry; 2020 Jul 15;88(2):185–96.

46. Dube SR, Fairweather D, Pearson WS, Felitti VJ, Anda RF, Croft JB. Cumulative Childhood Stress and Autoimmune Diseases in Adults. Psychosomatic Medicine. 2009 Feb;71(2):243–50.

47. Hughes K, Bellis MA, Hardcastle KA, Sethi D, Butchart A, Mikton C, et al. The effect of multiple adverse childhood experiences on health: a systematic review and meta-analysis. The Lancet Public Health. Elsevier; 2017 Aug 1;2(8):e356–66.

48. A F, TJ L, C S, N D, L A, T S, et al. Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. American Journal of Epidemiology. Am J Epidemiol; 2017 Nov 1;186(9):1026–34.

49. Batty GD, Gale C, Kivimaki M, Deary I, Bell S. Generalisability of Results from UK Biobank: Comparison With a Pooling of 18 Cohort Studies. medRxiv. Cold Spring Harbor Laboratory Press; 2019 Aug 13;:19004705.

50. Duncan L, Shen H, Gelaye B, Meijsen J, Ressler K, Feldman M, et al. Analysis of polygenic risk score usage and performance in diverse human populations. Nature Communications. Nature Publishing Group; 2019 Jul 25;10(1):1–9.