A higher genetic risk for depression is a greater risk factor for cardiovascular disease in women

Jiayue-Clara Jiang (PhD)[1*], Kritika Singh (BA)[2,3*], Lea K. Davis (PhD)[2,3,4,5,6], Naomi R. Wray (PhD)[1,7], Sonia Shah (PhD)[1]

[1] Institute for Molecular Bioscience, The University of Queensland, St Lucia, 4072, Australia

[2] Division of Genetic Medicine, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, 37232, USA

[3] Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, 37232, USA

[4] Department of Molecular Physiology and Biophysics, Vanderbilt University Medical Center, Nashville, TN, 37232, USA

[5] Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center, Nashville, TN, 37232, USA

[6] Departments of Medicine and Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, 37232, USA

[7] Queensland Brain Institute, The University of Queensland, St Lucia, 4072, Australia

*co first authors

Short title: Genetic risk of psychiatric disorders and cardiovascular disease

Corresponding authors:

Sonia Shah, PhD

Email: s.shah1@uq.edu.au

Address: 306 Carmody Rd, Institute for Molecular Bioscience, The University of Queensland, St Lucia, Australia, 4072

Jiayue-Clara Jiang, PhD

Email: j.jiang@uq.edu.au

Address: 306 Carmody Rd, Institute for Molecular Bioscience, The University of Queensland, St Lucia, Australia, 4072

Abstract

Background

Epidemiological studies show increased cardiovascular disease (CVD) risk amongst individuals with psychiatric disorders; however, sex differences in comorbidity have been inconsistent. This study examined the sex-specific association between higher genetic risk of three psychiatric disorders [major depression (MD), schizophrenia and bipolar disorder] and the risk of three CVDs [coronary artery disease (CAD), heart failure (HF) and atrial fibrillation].

Methods and Findings

Sex-specific association between polygenic scores (PGS) for psychiatric disorders and CVD outcomes was assessed in 345,201 European-ancestry individuals (UK Biobank) using logistic regression, with analyses replicated in an independent BioVU cohort (N = 62,641).

In the UK Biobank cohort, a 1-standard-deviation increase in PGS_{MD} was associated with a 1.11- (95% CI: 1.06 - 1.16; p = 5E-06) and 1.09-fold (95% CI: 1.04 - 1.14; p = 0.00038) higher risk of incident CAD and HF in females respectively, but only with a 1.04-fold (95% CI: 1.02 - 1.07; p = 0.0016) increased risk of incident CAD in males. These associations remained even in the absence of any psychiatric disorder diagnosis. The higher risk in females may potentially be mediated by greater BMI increase in females. The risk increase of HF was consistent between females who were pre-menopausal at baseline (mean age = 44.9 years) and post-menopausal at baseline (mean age = 60.8 years), while the risk increase of CAD was only observed in the baseline post-menopausal cohort. No significant association with CVD risks was observed for the genetic risk

of schizophrenia or bipolar disorder. The greater increase in CAD risk in females was replicated in the BioVU cohort.

Conclusions

Genetic predisposition to MD confers a greater risk of CAD and HF in females compared to males, even in the absence of any depression diagnosis. In addition to depression diagnosis, this study highlights the importance of considering predisposition to depression in heart health checks, especially in women.

Keywords

Psychiatric disorder, cardiovascular disease comorbidity, depression, polygenic score, sex-specific risks, menopause

1 Non-Standard Abbreviations and Acronyms

Abbreviation	Definition
AF	Atrial fibrillation
AHA	American Heart Association
AUC	Area under the receiver operator characteristic curve
BD	Bipolar disorder
BMI	Body mass index
CAD	Coronary artery disease
CI	Confidence interval
CPT	Current Procedural Terminology
CVD	Cardiovascular disease
EHR	Electronic health record
GWAS	Genome-wide association study
HF	Heart failure
HR	Hazard ratio
ICD	International Classification of Diseases
LD	Linkage disequilibrium
MD	Major depression
MDD	Major depressive disorder
MR	Mendelian randomisation
OR	Odds ratio
PC	Principal component
PGS	Polygenic score
PRS-CS	Polygenic Risk Score–Continuous Shrinkage
SCZ	Schizophrenia
SD	Standard deviation
SNP	Single nucleotide polymorphism
T2D	Type II diabetes
VUMC	Vanderbilt University Medical Center

2

Introduction

The idea that mental health is linked to cardiovascular health has long been recognised: William Harvey wrote in 1628 "every affection of the mind that is attended with either pain or pleasure, hope or fear, is the cause of an agitation whose influence extends to the heart" [1]. Modern epidemiological studies find that collectively, individuals with psychiatric disorders, namely major depressive disorder (MDD), schizophrenia (SCZ) and bipolar disorder (BD), have a ~50% higher odds of developing cardiovascular diseases (CVD) [2]. In particular, there is a doubling of the odds of developing coronary heart disease amongst individuals with depression, and those with SCZ are at a higher risk of coronary heart disease and congestive heart failure (HF) [2]. The increased CVD risk amongst individuals with psychiatric disorders may be attributed to a combination of genetic and non-genetic factors, with the latter including the use of prescribed psychiatric medications [3] and the adoption of adverse habits, such as smoking and social isolation [4].

There is marked sex difference in psychiatric disorders and CVDs, in both prevalence and clinical presentation. Depression is twice as prevalent in females than males (lifetime prevalence of 26.1% versus 14.7%) [5], with women showing augmented symptom severity [6]. While the prevalence of SCZ does not differ between genders, some studies have found that men show a higher incidence (incidence risk ratio = 1.42), earlier age of onset and worse premorbid functioning [7]. Men also show an earlier onset for BD, although the prevalence is reported to be equal between sexes for bipolar I disorder but higher in women for bipolar II disorder, which is characterised by depressive episodes [8]. Furthermore, although it is traditionally considered to be a male disease (lifetime risk of 60.3% in men versus 55.6% in women at an index age of 45 years) [9], CVD is a leading cause of female deaths [10].

Few studies have investigated the sex differences in the cardiovascular comorbidity of SCZ and BD, and observational studies have presented inconsistent findings on the sex-specific association between depression and CVD outcomes [11]. While some observational studies (N ≤ 3,237) have found depression or depressive symptoms to be a risk factor for HF and coronary artery disease (CAD) amongst women but not in men [12, 13], a study of a Chinese cohort (N = 512,712) found an association between depression and CVD-related mortality in men only [14]. However, observational studies cannot establish causal associations, and many of these observational studies did not account for the use of antidepressants, which are known to have an adverse cardiometabolic effect [3]. The variability in findings may be due to differences in sample sizes, unmeasured confounders, follow-up times, sex balance in the cohort, and the criteria for defining depression phenotypes and CVD outcomes.

Large genome-wide association studies (GWAS) have been very successful at identifying genetic loci associated with disease risk. Novel statistical methods applied to GWAS data are facilitating our understanding of shared biology between diseases, as well as in making causal inferences in disease comorbidity, overcoming some of the caveats of observational studies. Previous genetic analyses, for example, have used Mendelian randomisation (MR) to provide evidence to support a causal effect of depression on CAD and HF [15]. However, fewer studies have investigated sex differences using such genetic approaches, with an example being a small-scale study (N = 18,385, 50.9% female) that showed a positive association between MD and CAD in females but not males [16]. To what extent genetic factors contribute to comorbidity between psychiatric disorders and CVDs in different sexes is an outstanding question, and one that will better inform CVD prevention strategies. In this study, we used polygenic scores (PGS) derived from GWAS summary statistics (which estimate an individual's genetic liability to a disease) to examine the sex-specific causal

association between higher genetic risk of three psychiatric disorders [BD, major depression (MD) and SCZ] and the risk of three CVDs [atrial fibrillation (AF), CAD and HF] in the UK Biobank cohort of over 345,000 individuals. The findings were independently replicated in a BioVU cohort of over 62,000 individuals.

Methods

UK Biobank cohort summary

The UK Biobank cohort consists of over 500,000 participants (aged 40-69 years at baseline) from the UK, who were recruited between 2006 and 2010 [17]. Diverse phenotypic and genomic data were collected from participants, such as clinical diagnosis, physical and biochemical measurements, and answers from questionnaires. Consents were collected from all participants. The current study was conducted using the UK Biobank Resource under Application Number 12505. This research is covered by The University of Queensland Human Research Ethics Committee approval (HREC number 2020/HE002938).

Using information provided by the UK Biobank genotype data release, quality control was performed on the UK Biobank cohort to remove individuals who were outliers for heterozygosity and genotype missing rates, who were excluded from kinship inference, who showed evidence of putative sex chromosome aneuploidy, and who showed mismatching self-reported and genetically inferred sex (definitions of exclusion criteria are described in Bycroft et al. [17]). Individuals who have withdrawn from the study or have missing phenotypes [baseline body mass index (BMI) or smoking status] were also excluded from the analysis. After quality control, a total of 345,201 unrelated (genetic relatedness < 0.05) individuals of European ancestry (inferred from the genetic data) were subjected to analysis in this study (ancestry calling is described in Yengo et al. [18]).

BioVU cohort summary

Vanderbilt University Medical Center (VUMC) is a tertiary care centre that provides inpatient and outpatient care in middle Tennessee, US. The VUMC electronic health record (EHR) system, established in 1990, includes data on billing codes from the International Classification of Diseases, 9th and 10th editions (ICD-9 and ICD-10), Current Procedural Terminology (CPT) codes, laboratory values, reports, and clinical documentation. In 2007, VUMC launched a biobank, BioVU, which links a patient's DNA sample to their EHR. The BioVU consent form is provided to patients in the outpatient clinic environments at VUMC. The VUMC Institutional Review Board oversees BioVU and approved this project (IRB#172020). At VUMC, all medical record data were extracted from the EHRs of 72,634 individuals of European ancestry, determined by genetic ancestry analysis. We excluded individuals who did not have a BMI measurement or age recorded in the EHRs, leaving a total of 62,641 individuals for analysis.

UK Biobank phenotype ascertainment

The inclusion and exclusion criteria for the CVDs, CVD-associated risk factors [hypercholesterolemia, hypertension, and type II diabetes (T2D)], and psychiatric disorders are shown in Supplementary Table 1 and Supplementary Table 2. Smoking status at baseline (whether a person has smoked ever, defined as current or previous smokers), was collated from UK Biobank data-field 20116 ("Smoking status"). The baseline BMI of individuals was retrieved (data-field 21001). The recruitment ages (in years) of individuals were calculated using "Date of attending assessment centre" (data-field 53) and "Year of birth" (data-field 34). Summary of characteristics for the whole cohort and the cohort with no psychiatric diagnosis and medications are shown in

Table 1 and Supplementary Table 3 respectively. UK Biobank phenotypic data collected between 2006 and August 2020 was used for phenotype extraction.

Table 1. Summary of characteristics of the UK Biobank cohort

		Female		Male	
Total Number		186701		158500	
Average age at baseline in years (SD)		57 (7.9)		57.4 (8.1)	
Average BMI at baseline (SD)		26.9 (5.1)		27.8 (4.2)	
Current or previous smokers at baseline (%)		77185 (41.3)		81217 (51.2)	
		Case	Control*	Case	Control*
BD	N (%)	606 (0.3)	155517 (83.3)	421 (0.3)	141304 (89.2)
	Age at baseline (SD)	55.7 (8)	57.1 (7.9)	57.4 (7.8)	57.4 (8.1)
MD	N (%)	7234 (3.9)	155517 (83.3)	4207 (2.7)	141304 (89.2)
	Age at baseline (SD)	56.6 (8)	57.1 (7.9)	57.3 (8)	57.4 (8.1)
SCZ	N (%)	376 (0.2)	155517 (83.3)	447 (0.3)	141304 (89.2)
	Age at baseline (SD)	56 (8.2)	57.1 (7.9)	55.5 (8.6)	57.4 (8.1)
incident AF	N (%)	4006 (2.1)	181219 (97.1)	7256 (4.6)	147765 (93.2)
	Age at baseline (SD)	62.8 (5.8)	56.8 (7.9)	62.5 (6)	57 (8.1)
incident CAD	N (%)	2057 (1.1)	182495 (97.7)	5501 (3.5)	144435 (91.1)
	Age at baseline (SD)	61.9 (6.4)	56.9 (7.9)	61.3 (6.5)	57 (8.1)
incident HF	N (%)	1796 (1)	184373 (98.8)	3366 (2.1)	153594 (96.9)
	Age at baseline (SD)	62.5 (6.4)	56.9 (7.9)	62.6 (6)	57.3 (8.1)
incident hyper-cholesterolemia	N (%)	10684 (5.7)	154721 (82.9)	15003 (9.5)	111598 (70.4)
	Age at baseline (SD)	61.2 (6.3)	56.1 (7.9)	61.2 (6.5)	55.8 (8.2)
incident hypertension	N (%)	23719 (12.7)	132431 (70.9)	27762 (17.5)	95193 (60.1)
	Age at baseline (SD)	60.6 (6.8)	55.7 (7.9)	60.7 (6.8)	55.5 (8.2)
incident T2D	N (%)	4977 (2.7)	178927 (95.8)	7815 (4.9)	146092 (92.2)
	Age at baseline (SD)	60.3 (6.9)	56.9 (7.9)	60.7 (6.8)	57.1 (8.1)
Baseline post-menopausal (%)**		114382 (80.1)		-	
Baseline pre-menopausal (%)**		28409 (19.9)		-	

* Controls for BD, MD and SCZ were individuals who had no diagnosis of psychiatric disorders or psychiatric medication use

** Menopausal state was defined by self-reported menopausal state at baseline (baseline post-menopausal: females who answered 'yes' to having had menopause at baseline; baseline pre-menopausal: females who answered 'no' to having had menopause at baseline). Female participants who met any of the following criteria were excluded from analysis: 1) females who did not report to have had menopause at baseline but later self-reported menopause in follow-up assessments (although only ~13% completed follow-up surveys on their menopausal states); 2) females who did not report to have had menopause at baseline but were at least 50 years of age at baseline; 3) females who preferred not to answer or self-reported to be not sure about their menopausal state at baseline or any follow-up assessment. The percentages shown in the table represent percentages of total females who met the criteria for either menopausal group (N = 142,791).

UK Biobank incident CVD ascertainment

For each CVD, incident cases were defined as individuals who developed the specific CVD after recruitment into the UK Biobank. These were defined as

1. individuals whose dates of first diagnosis ["Date of first in-patient diagnosis" (data-field 41280)] of all relevant CVD-associated ICD-10 codes were later than the date of their first visit to the assessment centre ["Date of attending assessment centre" (data-field 53)], or
2. individuals who had no relevant CVD-associated ICD-10 code when they first visited the assessment centre but self-reported to have the relevant CVD-related illness in subsequent assessments (data-field 53), or
3. individuals who had no relevant CVD-associated ICD-10 code or self-reported illness when they first visited the assessment centre but subsequently had a relevant CVD-associated ICD-10 code recorded as the primary (data-field 40001) or contributory (data-field 40002) death cause.

Prevalent cases consisted of both incident cases and existing cases (full definitions in Supplementary Table 1) at baseline.

BioVU phenotype ascertainment

For the BioVU cohort, the inclusion and exclusion criteria for CVDs, CVD-associated risk factors and psychiatric disorders are shown in Supplementary Table 4. Smoking status was ascertained by "tobacco use disorder" (Phecode 318), which is a strong proxy for smoking. The median BMI and median age of individuals were selected from assessments taken at chronological visits, available

in the medical records at VUMC. Summary of phenotypes is presented in Supplementary Table 5 and Supplementary Table 6.

GWAS summary statistics

Association summary statistics derived from European-ancestry or predominantly European-ancestry cohorts were retrieved for three CVDs (AF, CAD and HF) and three psychiatric disorders (BD, MD and SCZ), which were used for generating PGS in the UK Biobank and BioVU cohorts (datasets summarised Supplementary Table 7). To avoid discovery bias, GWAS summary statistics used for single nucleotide polymorphism (SNP) selection and per-allele weight estimation were derived from studies that did not include individuals from the corresponding target cohorts (UK Biobank or BioVU) (Supplementary Table 7).

Polygenic score generation

PGSs are an estimate of the lifetime genetic liability of an individual to a trait. Briefly, we define a PGS of an individual, j, as a weighted sum of SNP allele counts: $\sum_{i=1}^{m} \hat{b}_i x_{ij}$, where m is the number of SNPs included in the predictor, $\hat{b}_i$ is the per-allele weight (based on the estimated effect of the allele on trait value) for the SNP, x_{ij} is a count of the number (0, 1, or 2) of trait-associated alleles of SNP i in individual j.

Several methods are available for optimising the choice of SNPs and the per-allele weights for PGS calculation. For analysis in the UK Biobank cohort, we used SBayesR [19] which implements a Bayesian approach. It does not require a training dataset for identifying trait-associated SNPs and estimating per-allele weights, and is computationally efficient [19, 20]. A linkage disequilibrium (LD) reference dataset was downloaded from the SBayesR website (banded LD

matrix) and default settings were used unless otherwise stated (--exclude-mhc; --chain-length 50000; --burn-in 10000; --no-mcmc-bin) [19]. For each disease, we applied PLINK1.9 [21] to calculate PGS (based on SbayesR SNPs and weights) in UK Biobank individuals. Analysis was restricted to unrelated individuals of European ancestry to avoid biases introduced by population stratification and cryptic relatedness.

PGSs in BioVU were generated using Polygenic Risk Score–Continuous Shrinkage (PRS-CS) [22] (which has been shown to have similar performance to SbayesR [20]). The LD reference panel was constructed from 503 European samples in the 1000 Genomes Project Phase 3.41.

The PGS distribution was scaled to a mean of zero and standard deviation (SD) of 1; therefore, effect size estimates in subsequent logistic or linear regression were interpreted as changes in outcome per 1-SD increase in PGS. To validate the PGS, logistic regression was used to examine the association with disease prevalent cases in the prediction cohort (UK Biobank or BioVU). PGS performance was evaluated using the Nagelkerke pseudo R^2 [R package rcompanion [23] (version 2.4.13)] and area under the receiver operator characteristic curve (AUC) [R package pROC [24] (version 1.18.0)]. The Nagelkerke pseudo R^2 compares the reduced model (containing only the covariates as predictors) and the full model (with PGS added as a predictor), and is a relative measure of how well the full model explains the data compared to the reduced model. The AUC can be interpreted as a probability of a case having a higher PGS than a control.

Sex-stratified logistic regression of PGS and incident CVD in the UK Biobank cohort

For each CVD-psychiatric disorder pair, logistic regression was used to estimate the association between incident CVD and psychiatric disorder PGS in sex-stratified cohorts, including BMI, smoking status, genotyping array, age and 20 genetic principal components (PC) as covariates. A

$p < 0.0028$ indicated statistical significance (multiple testing correction for 18 tests – association of three psychiatric PGS with three CVDs in two sexes). The regression coefficients (beta) for the PGS were compared between sexes using a Wald test (two-sided $p < 0.05$ indicated statistical significance). Due to the previously reported genetic correlations between CVD and depression [25], to assess if the MD PGS were associated with CVD risks independently of the CVD PGS, we performed sensitivity analysis where the PGS for the relevant CVD was fitted as an additional covariate. For all sensitivity analyses, a nominal significance threshold ($p < 0.05$) was applied due to the reduction in sample sizes.

Cox proportional hazards regression analysis

To account for the age differences between CVD cases and controls, we performed age-as-time-scale Cox proportional hazards regression to investigate the sex-specific association between incident CVD and psychiatric disorder PGS. For each CVD incident case, we defined their time-to-event (in months) as the number of months since birth [estimated using "Year of birth" (data-field 34) and "Month of birth" (data-field 52)], to the age they were diagnosed with CVD [estimated using "Date of first in-patient diagnosis" (data-field 41280) of the relevant CVD-associated ICD-10 code], or first self-reported CVD illness [estimated using "Date of attending assessment centre" (data-field 53)], or died of CVDs [estimated using "Date of death" (data-field 40000)]. We censored CVD controls at the age of death, age of losing contact [estimated using "Date lost to follow-up" (data-field 191)], or age at censoring (August 2020). The Cox models included BMI, smoking, genotyping array and 20 genetic PCs as covariates, and were analysed using the "Survival" package (version 3.2.11) in R.

Cox proportional hazards regression stratified by baseline menopausal state

Longitudinal studies report adverse changes in body composition, lipids, and measures of vascular health during menopause transition (perimenopause), which can increase a woman's risk of developing CVD post-menopause [26]. The mean age of natural menopause is 50 years (interquartile range = 48.0–53.0 years) [26]. We sought to determine if the association between genetic risk of MD and risk of incident CVD differed between two groups of women at different stages of menopause at baseline.

Menopause state at baseline was determined using UK Biobank data-field 2724 ("Had menopause?"). We first explored the MD-CVD association amongst females who were post-menopausal at baseline (who answered 'yes' to the survey question "Had menopause?" at the initial assessment visit). We also conducted analysis in a cohort consisting of females who at baseline were < 50 years of age and pre-menopausal (answered 'no' to "Had menopause?" at the initial assessment), and either answered 'no' or did not answer the "Had menopause" question in subsequent assessments (about 13% of females completed follow-up assessments on menopausal state). Females who did not answer, preferred not to answer, or answered "not sure" to the menopause question at baseline were excluded from both menopausal groups, while those who in any follow-up assessments reported to have had menopause, preferred not to answer, or answered "not sure" were also excluded from the baseline pre-menopausal group. The final baseline pre-menopausal cohort had a mean age at baseline of 44.9 years old (SD = 2.6) compared to a mean age at baseline of 60.8 years (SD = 5.4) in the baseline post-menopausal cohort.

A total of 142,791 females in the study met the criteria for either menopausal group, including 114,382 (80.1%) females who self-reported menopause at baseline, and 28,409 (19.9%) females in the baseline pre-menopausal group. We performed Cox proportional hazards regression analysis (as described above) between incident CVD and PGS_{MD} in the two menopausal groups. BMI,

smoking status, genotyping array, and 20 genetic PCs were included in the model as covariates. For each CVD, we applied a Wald test to compare the Cox regression coefficient (beta) for PGS_{MD} in each menopausal group against the previously estimated coefficient for the male cohort (as described above).

Sex-stratified association between CVD-associated risk factors and MD PGS

To understand the mechanisms underlying the sex differences in CVD risks, we explored the association between PGS_{MD} and various risk factors of CVD (hypercholesterolemia, hypertension, T2D, and BMI) in the UK Biobank cohort. For incident hypercholesterolemia, hypertension and T2D, we performed logistic regression analysis and included BMI, smoking status, genotyping array, age, and 20 genetic PCs as covariates. We performed linear regression to estimate the association between baseline BMI and PGS_{MD}, including smoking status, genotyping array, age and 20 genetic PCs as covariates.

Logistic regression between genetic risk of MD and CVD in individuals with no diagnosis of psychiatric disorders or psychiatric medication use

To dissociate the effects of behavioural changes or medication use as a consequence of depression diagnosis, sex-stratified associations were estimated between PGS_{MD} and CVD risks amongst UK Biobank participants who did not meet the criteria for the three psychiatric disorders, had no ICD diagnosis of personality or neurotic disorders, had no self-reported history of mood disorders, and were not on any antidepressants or antipsychotic medications (detailed phenotype information in Supplementary Table 2). A total of 296,821 individuals from the UK Biobank cohort met the above criteria and were included for analysis. To account for the reduced statistical power to detect an

association with incident CVDs due to the reduced sample size in this cohort, we also investigated the associations of PGS_{MD} with prevalent CVDs.

Logistic regression in BioVU

Findings from UK Biobank were tested for association in the independent BioVU cohort at VUMC. Logistic regression was used to estimate the sex-stratified association between prevalent CVD or CVD-associated risk factors and psychiatric disorder PGS, including median age, median BMI, smoking status and 10 genetic PCs as covariates. The sex-stratified associations were also estimated for PGS_{MD} amongst individuals that did not meet the criteria for the three psychiatric disorders, and had no ICD diagnosis of neurotic disorders and psychosis.

Results

PGS validation in UK Biobank cohort

We generated PGS for BD, MD and SCZ in 345,201 unrelated, European-ancestry individuals [186,701 (54.1%) females and 158,500 (45.9%) males] from the UK Biobank. The odds ratio (OR) for MD was 1.3 [95% confidence interval (CI): 1.27 - 1.32; $p < 2E-16$] per SD increase in PGS_{MD}, and the corresponding ORs for PGS_{BD} and PGS_{SCZ} were 2.04 (95% CI: 1.91 - 2.17; $p < 2E-16$) and 2.00 (95% CI: 1.86 - 2.15; $p < 2E-16$) respectively (Supplementary Table 8). Similarly, ORs for PGS of CVDs ranged between 1.19 and 1.63 ($p < 2E-16$) (Supplementary Table 8). The AUC for the models containing PGSs (CVD or psychiatric disorder) and covariates ranged between 0.65 and 0.8, indicating agreement between the predicted disease liability and observed prevalent cases (Supplementary Table 8). Our models demonstrate comparable performance to the corresponding PGSs generated previously using different methods [20, 27, 28, 29] (Supplementary Table 8).

Overall, the addition of PGSs improved the prediction performance of the logistic regression models, as indicated by positive Nagelkerke R^2 values (Supplementary Table 8). For the psychiatric disorders PGSs generated for the BioVU cohort, ORs ranged between 1.14 and 1.55 (Supplementary Table 8).

Increased genetic risk of MD is associated with greater increase of CAD and HF risk in females

Sex-stratified logistic regression analysis showed that PGS_{MD} was significantly associated with increased CAD risk in both sexes [$OR_{CAD\text{-}female}$ = 1.11 (95% CI: 1.06 - 1.16; p = 5E-06) and $OR_{CAD\text{-}male}$ = 1.04 (95% CI: 1.02 - 1.07; p = 0.0016)], with risk significantly higher in females compared to males (two-sided Wald test p = 0.027) (Figure 1). PGS_{MD} was significantly associated with HF risk in females only [$OR_{HF\text{-}female}$ = 1.09 (95% CI: 1.04 - 1.14; p = 0.00038)] (two-sided Wald test p = 0.025) (Figure 1). We did not observe statistically significant associations between PGS_{MD} and incident AF in the sex-stratified analysis (Figure 1). Similar results were observed in sensitivity analyses, where the corresponding CVD PGS was added to the model as a covariate, indicating that any significant association between PGS_{MD} and incident CVDs was independent of the CVD genetic risk (Supplementary Figure 1). No association between SCZ or BD PGS with CVD risks passed multiple testing correction ($p < 0.0028$) in either sex (Figure 1).

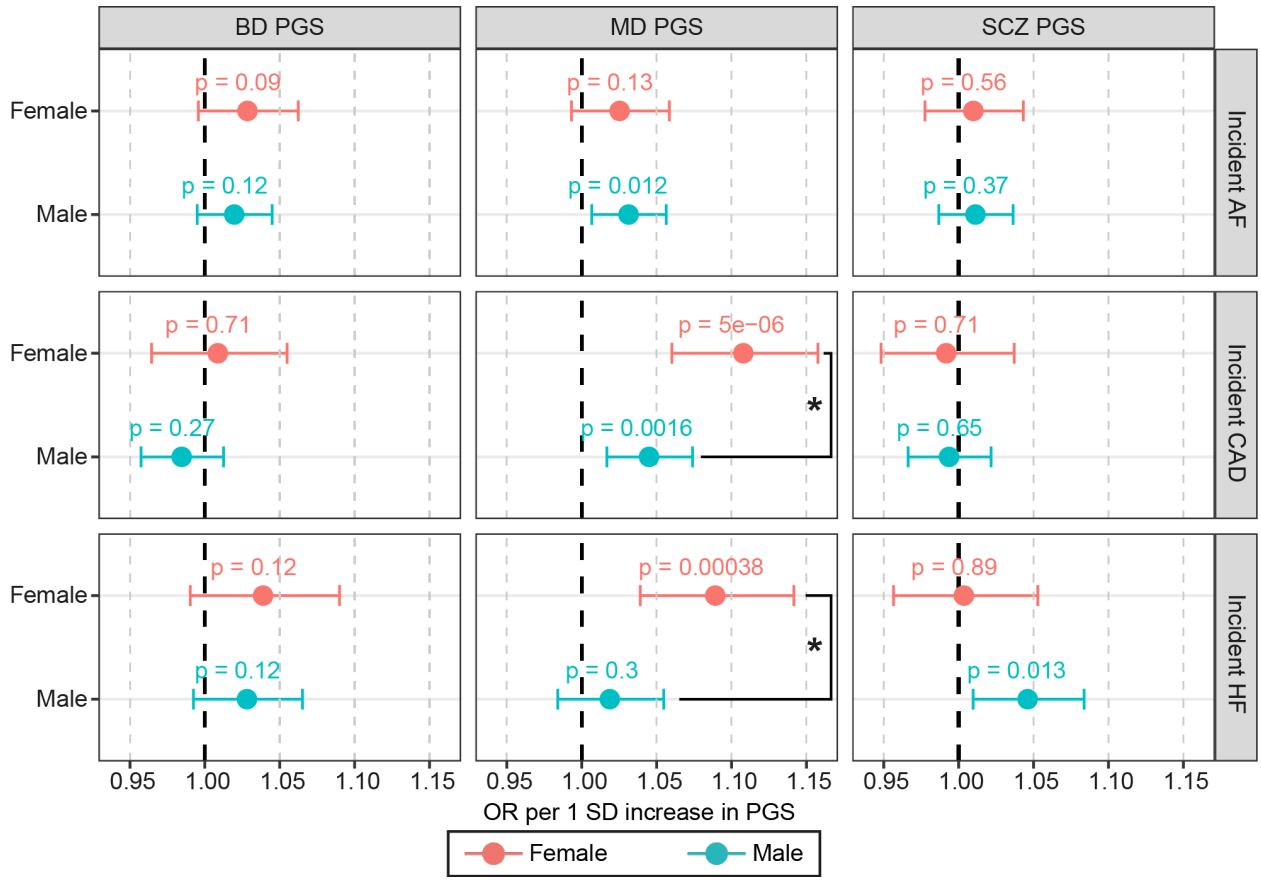


Figure 1. Association between psychiatric disorder PGSs and incident CVDs in the UK Biobank cohort. Associations for the sex-stratified cohorts (red - female; blue - male) were estimated with logistic regression. X-axis shows the OR per SD increase in PGS, with p-values labelled, and error bars indicate 95% confidence intervals. Asterisk (*) indicates a statistically significant difference (two-sided Wald test $p < 0.05$) in the beta values between females and males. Dark grey line indicates OR of 1. AF, atrial fibrillation; BD, bipolar disorder; CAD, coronary artery disease; HF, heart failure; MD, major depression; OR, odds ratio; PGS, polygenic score; SCZ, schizophrenia; SD, standard deviation.

Cox proportional hazards regression analysis, which accounts for age-specific risk of incident CVD, gave hazard ratio (HR) estimates that were similar to the OR estimated from logistic regression (Supplementary Figure 2), indicating minimal bias due to the older average age of CVD cases versus control (Table 1). We thus performed logistic regression for all subsequent stratified analyses unless otherwise stated.

Menopausal status influenced the association between genetic risk of MD and incident CVD

We further explored the association between PGS_{MD} and incident CVD in females, by stratifying based on their baseline menopausal status. As expected, we observed an overall higher incidence of CVD amongst females who were post-menopausal at baseline, compared to females who were pre-menopausal at baseline (chi square $p < 2.2E-16$) (Supplementary Table 9). We found that PGS_{MD} was significantly associated with increased incident CAD in the baseline post-menopausal females only ($HR_{CAD-post}$ = 1.11 (95% CI: 1.06 - 1.17; p = 5.4E-05), which was still significantly higher than the CAD risk increase in the male cohort (two-sided Wald test = 0.05). However, we did not observe a significant association with CAD in the baseline pre-menopausal group ($HR_{CAD-pre}$ = 1.01 (95% CI: 0.80 - 1.27; p = 0.94). In contrast, we observed consistent positive associations of PGS_{MD} with age-specific risk of HF regardless of menopausal states, reflected by comparable HR estimates [HR_{HF-pre} = 1.31 (95% CI: 1.02 - 1.68; p = 0.034); $HR_{HF-post}$ = 1.09 (95% CI: 1.03 - 1.15; p = 0.0024)] (Figure 2).

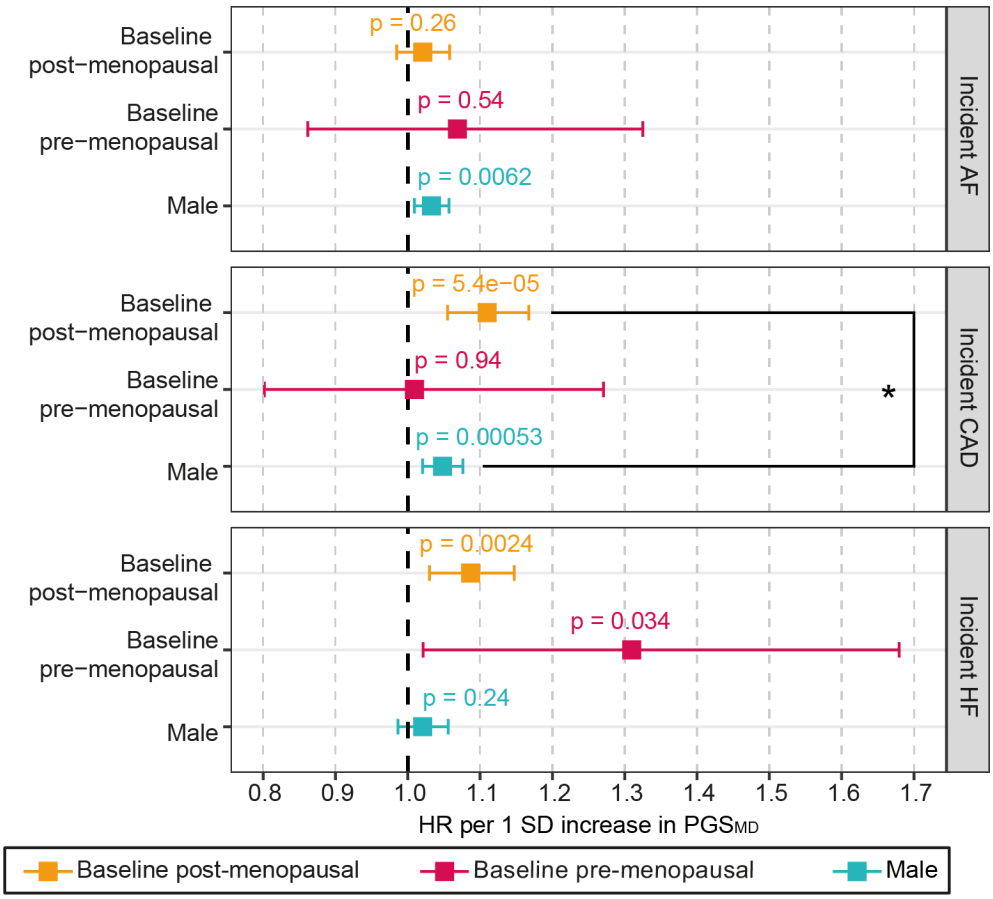


Figure 2. Age-specific change in risk of incident CVDs per SD increase in PGS_{MD} in the UK Biobank cohort, stratified by menopausal state. Associations for females who were baseline post-menopausal (orange) and baseline pre-menopausal (pink) were estimated with Cox proportional hazards regression models, and the estimates for the male cohort (blue) are shown for the purpose of comparison. X-axis shows the HR per SD increase in PGS, with p-values labelled, and error bars indicate 95% confidence intervals. Asterisk (*) indicates a statistically significant difference (two-sided Wald test $p < 0.05$) in the beta values between the corresponding female cohort and male cohort. Dark grey line indicates HR of 1. AF, atrial fibrillation; BD, bipolar disorder; CAD, coronary artery disease; AF, atrial fibrillation; CAD, coronary artery disease; HF, heart failure; HR, hazard ratio; PGS, polygenic score; SD, standard deviation.

A higher genetic risk of MD is associated with a greater increase in BMI in females compared to males

To gain a mechanistic insight into the female preponderance in the association between CVD risk and PGS_{MD}, we explored the association between CVD-associated risk factors and the genetic risk of MD in the UK Biobank cohort. An increased genetic risk of MD was significantly associated with an increase in incident hypercholesterolemia, hypertension, T2D and baseline BMI in both

sexes, though a significant sex difference was only observed for BMI, with a higher increase in BMI for females compared to males per SD increase in PGS_{MD} (two-sided Wald test p = 5E-06) (Figure 3).

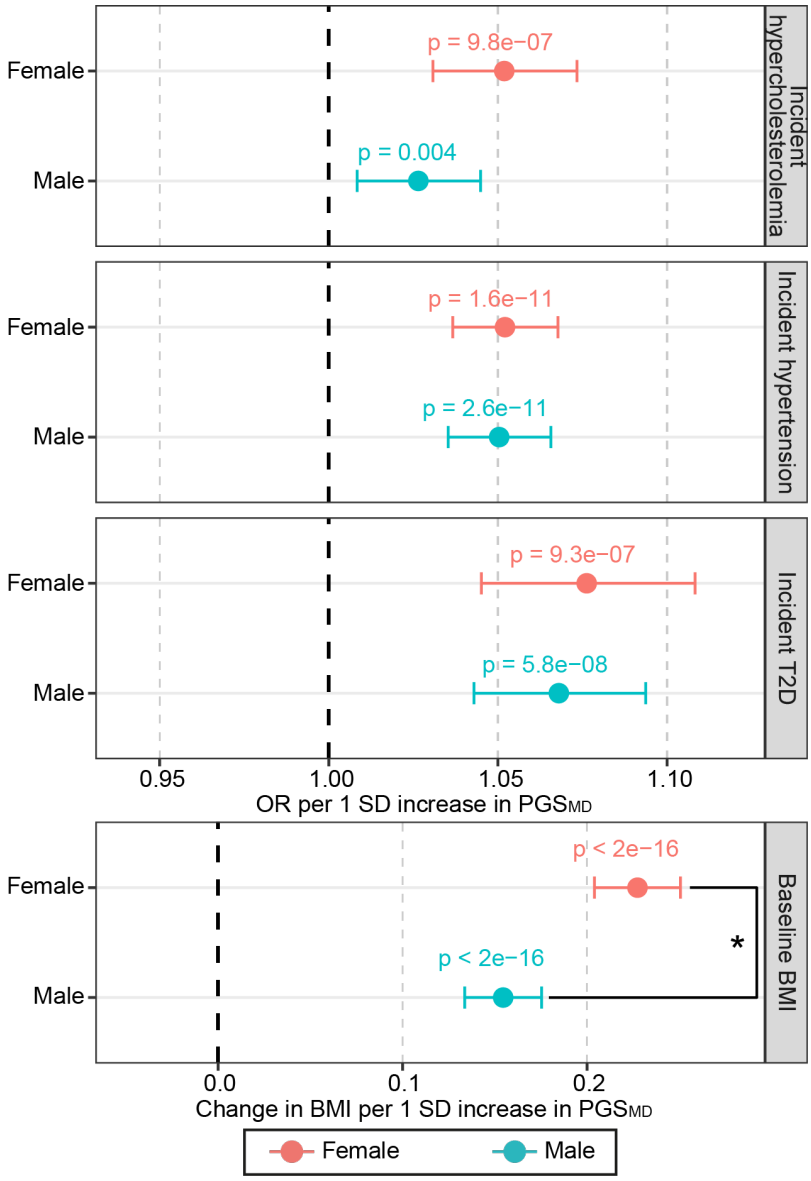


Figure 3. Sex-specific associations of PGS_{MD} with CVD-associated risk factors amongst UK Biobank individuals. Associations are shown for the sex-stratified cohorts (red - female; blue - male). Asterisk (*) indicates a statistically significant difference (two-sided Wald test $p < 0.05$) in the risk estimates between females and males. OR was estimated with logistic regression, and change in BMI was estimated with linear regression. X-axis shows the OR or change in BMI per SD increase in PGS, with p-values labelled, and error bars indicate 95% confidence intervals. Dark grey line indicates OR of 1 or change of 0. BMI, body mass index; OR, odds ratio; PGS, polygenic score; SD, standard deviation; T2D, type II diabetes.

Genetic risk of MD is associated with increased CVD risk in individuals with no psychiatric diagnosis and medications

Amongst individuals who had no psychiatric disorder diagnosis and were not on any psychiatric medication, we observed a positive association between PGS_{MD} and prevalent CAD in both sexes [$OR_{CAD\text{-}female}$ = 1.06 (95% CI: 1.03 - 1.1; p = 0.00099) and $OR_{CAD\text{-}male}$ = 1.04 (95% CI: 1.02 – 1.06; p = 1.4E-05)] (Figure 4A). We observed a positive association with prevalent HF in females [$OR_{HF\text{-}female}$ = 1.06 (95% CI: 1.01 – 1.12; p = 0.021)] and a positive association with prevalent AF in males [$OR_{AF\text{-}male}$ = 1.03 (95% CI: 1.01 – 1.05; p = 0.01)] (Figure 4A). To avoid survival bias, we repeated the association analysis for incident CVD cases. Similar risk estimates were observed between PGS_{MD} and incident CVDs in this group, but only the associations with incident CAD in both sexes and incident AF in males reached nominal significance, likely due to reduced sample sizes (as indicated by similar effect estimates to prevalent cases but larger CIs) (Figure 4B). Similar to the associations observed for the whole cohort, we observed a significant, positive association between PGS_{MD} and the CVD-associated risk factors (except hypercholesterolemia in males), including a greater increase in the association with BMI for females (Supplementary Figure 3).

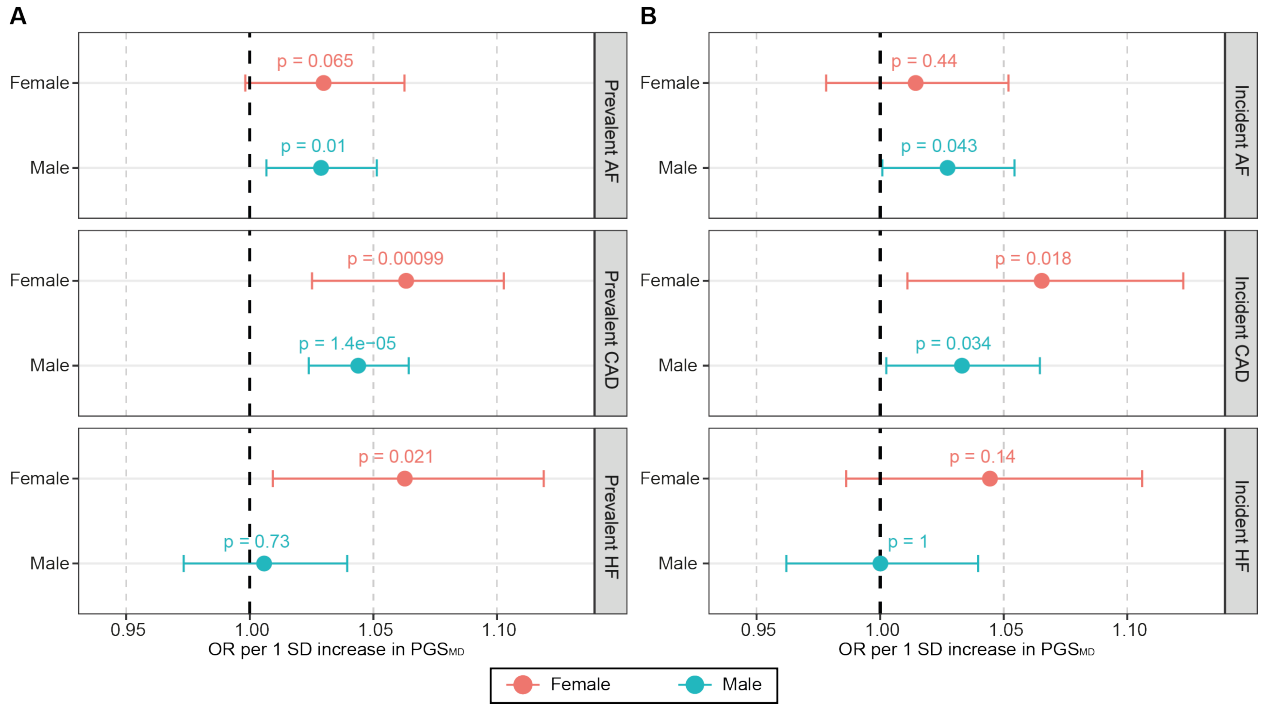


Figure 4. Association between PGS_{MD} and CVDs amongst UK Biobank individuals with no diagnosis of psychiatric disorders and psychiatric medications. The associations of PGS_{MD} with (**A**) prevalent or (**B**) incident CVDs are shown for the sex-stratified cohorts (red – female; blue – male). Associations were estimated with logistic regression. X-axis shows the OR per SD increase in PGS, with p-values labelled, and error bars indicate 95% confidence intervals. Dark grey line indicates OR of 1. AF, atrial fibrillation; CAD, coronary artery disease; HF, heart failure; OR, odds ratio; PGS, polygenic score; SD, standard deviation.

Replication of main findings in the BioVU cohort

We replicated our analysis in the BioVU cohort (N = 62,641), which was independent from the UK Biobank cohort. Similar to the UK Biobank cohort, we observed a positive association between PGS_{MD} and prevalent CAD in the females of the BioVU cohort [$OR_{CAD\text{-}female}$ = 1.09 (95% CI: 1.02 – 1.16; p = 0.012)], similar to the estimates for incident CAD from the UK Biobank cohort (Supplementary Figure 4). However, the OR estimate for males was lower than that in the UK Biobank cohort and did not reach nominal significance. Unlike the UK Biobank cohort, we did not find a positive association between PGS_{MD} and HF risk in the BioVU cohort, possibly due to the reduced sample sizes (Supplementary Figure 4). Similar to the UK Biobank cohort, we observed a significant association between PGS_{MD} and BMI in the BioVU cohort, with a marginally greater

increase in female BMI (two-sided Wald test $p > 0.05$) (Supplementary Figure 5). However, we did not observe any association with other CVD-associated risk factors in the BioVU cohort.

Amongst BioVU participants with no psychiatric disorder diagnosis, we did not observe any significant association of PGS_{MD} with prevalent CAD or HF, although we observed a marginally greater increase in female BMI (two-sided Wald test $p > 0.05$) as described above (Supplementary Figure 6). We also observed a negative association with prevalent AF and hypertension for both sexes, likely due to the lack of information on medication use in the BioVU cohort, or survival bias as only prevalent data were available (Supplementary Figure 6).

Discussion

Our study demonstrates for the first time the causal association between genetic predisposition to MD with incident HF in females but not in males, and validates a previously reported sex difference (increased risk in women) in the association with CAD. Though previous associations between depression and HF have been reported using genetic analyses (Mendelian Randomisation analysis) [15], these have not been sex-stratified and are likely to be driven by female-specific risk. Interestingly, the associations between the genetic risk of MD and CVD risk appeared to vary depending on the menopausal status, as the increased CAD risk was only observed in older females who were post-menopausal at baseline (mean age of 60.8 years at baseline) but not in females who were pre-menopausal at baseline (average age of 44.9 years at baseline). In comparison, the increased risk of HF was present in both groups, with an even higher risk in baseline-pre-menopausal women, although this cohort had significantly smaller sample sizes and thus were less statistically powered to detect any menopause-specific difference in HF risk increase across the two groups of females. Our results suggest this greater risk in females may partly be driven by a

higher causal effect on BMI, and exists even in the absence of depression diagnosis or use of antidepressants or antipsychotics, many of which are known to induce adverse cardiometabolic profiles [3]. Therefore, in addition to a diagnosis of depression which is incorporated in some CVD risk calculators (e.g QRISK3 [30]), a high genetic predisposition to depression in the absence of any diagnosis may also be an important consideration in cardiovascular disease prevention, particularly in women.

Epidemiological studies have reported notable sex differences in the prevalence and clinical manifestation of psychiatric disorders, with women particularly prone to depression [5]. Similarly, previous evidence from a small observational study (N = 2,501) in elderly individuals (≥ 65 years) found depression to be a risk factor for HF amongst elderly women but not in men [12]. However, observational studies do not provide evidence for a causal association. Previous genetic analyses that utilise GWAS summary data to assess causality have shown that an increased genetic risk for MD is associated with an increased CAD and HF risk [15]. However, these studies have not investigated sex-specific associations, except for a previously reported genetic association between MD PGS and CAD observed in females but not males in a (smaller) BioVU cohort [16]. Using a much larger dataset and sex-stratified genetic analyses, our study confirms MD as a greater risk factor for CAD in females compared males, and shows MD as a unique risk factor for HF in females only.

It has been suggested that MD and CVDs may represent manifestation of a common pathophysiological dysfunction [31]. The pathophysiology underlying the female preponderance remains unclear; however, our analysis showed that a higher genetic risk of MD was associated with a greater increase of BMI in females compared to males. This corroborates previous findings that females with depression had higher BMI compared to males with depression [32], despite a

similar BMI between sexes amongst individuals without depression, suggesting that obesity could potentially mediate the amplified risk of CVD amongst females with a higher risk of MD. We did not observe any sex differences in the association between genetic risk of MD and other risk factors, namely hypercholesterolemia, hypertension and T2D. Other possible mechanisms underlying the sex difference in the comorbidity of depression and CVDs include hormonal dysregulation and pro-inflammatory responses. Specific subtypes of depression, such as postpartum depression and post-menopausal depression, suggests the involvement of hormone fluctuations in depression amongst females [33]. Interestingly, Takotsubo cardiomyopathy, often triggered by emotional or physical stress, is the most common in post-menopausal women (90% of cases) [34]. At the same time, elevated inflammatory biomarkers, such as interleukin-6 and tumour necrosis factor alpha, are reportedly higher in depressed women relative to depressed men, suggesting a sex-differential inflammatory response to MD [33]. These pro-inflammatory cytokines are linked to incident CVD and poorer cardiovascular outcomes [35, 36]. Further studies are required to understand the role of these factors in the heightened risk of CVDs amongst females with depression.

The menopause transition involves extensive changes in sex hormones, body composition, and lipid profiles, which can increase CVD risk in women post-menopause [26]. The mechanisms underlying the lower magnitude of depression-related HF risk augmentation in post-menopausal females are unclear; however, menopausal hormone therapy, a treatment for improving menopausal symptoms, has been shown to reduce coronary heart disease by up to 48% amongst women who initiate the therapy less than 10 years after menopause or are less than 60 years old [37]. As over 47.5% (54,352 out of 114,382) of baseline post-menopausal females in this study self-reported to have used hormone replacement therapy at baseline (versus only 1.5% of the

baseline pre-menopausal cohort), this may explain the attenuated association between PGS_{MD} and HF risks in the post-menopausal females of UK Biobank. Further investigation into the impact of menopausal stage on CAD and HF risk in females with a higher genetic predisposition to MD is needed.

CVD risk calculators failing to incorporate psychiatric disorders as a predictor are reported to underestimate CVD risk by 30% and 60% in men and women respectively [38]. Currently, diagnoses of severe mental illnesses are included in the QRISK3 prediction algorithm (UK), and are recognised as a risk factor in New Zealand, where individuals with severe mental illness are advised to undergo earlier and more frequent risk assessments [30, 39]. Furthermore, the American Heart Association (AHA) has recommended incorporation of a depression diagnosis as a risk factor of adverse prognosis amongst patients with acute coronary syndrome [40]. However, our findings suggest that even in the absence of a depression diagnosis, predisposition to depression may be an important consideration in the preventative assessment of CVD risk, particularly amongst females.

Our study has several strengths. Though most existing genetic studies have used summary data from GWAS conducted in a sex-combined sample, and therefore were unable to investigate sex-specific association across different diseases, our study explores the sex differences by examining these associations in females and males separately. To our best knowledge, this is the first study that shows a female-specific causal association between higher genetic risk of depression and HF. Furthermore, our analysis examines the impacts of menopause state on the MD-CVD link, which has not been explored previously. Previous studies are also unable to dissociate effects of behaviours or medication use as a consequence of depression diagnosis, which our study was able to show through analysis of individuals with no diagnosis of psychiatric disorders and psychiatric

medications. In addition, our PGSs were derived from GWAS summary statistics with the largest sample sizes where possible, enhancing the statistical power in detecting genetic variants with small but significant effects. Previous MR studies have used GWAS summary statistics derived from cohorts with overlapping samples [15, 25, 41], which could lead to biased estimates of genetic association. To avoid any bias, our psychiatric PGSs were derived from discovery cohorts with no sample overlap with the prediction cohort (UK Biobank or BioVU).

However, limitations of our study need to be acknowledged. We derived PGSs from sex-combined summary statistics of GWAS, which might fail to capture genetic variants that confer sex-differential associations. The between-sex genetic correlations for BD, MDD and SCZ are reported to range between 0.86 and 1, indicating moderate to subtle sex differences in the genetic architecture of these psychiatric disorders [42]. Nevertheless, PGSs derived from sufficiently powered sex-stratified GWAS studies, which are becoming increasingly available for neuropsychiatric and cardiovascular traits [42, 43], will provide additional information on the sex-differential genetic factors in diseases. Moreover, due to the lower number of CVD cases amongst females in UK Biobank, our analysis in the female cohort had lower power. In addition, the analysis of CVD risk difference between different menopausal states did not account for the potential effects of any menopause treatments, and the small sample size of the baseline pre-menopausal cohort led to limited statistical power. Lastly, the BioVU data were only available on individuals who were admitted to VUMC; therefore, ascertainment bias may significantly affect observed associations. Furthermore, findings from the BioVU cohort may be confounded by medication status of participants, and potential survival bias as incident data are not available.

To our best knowledge, this is the first study that has explored sex-specific association between genetic risk of three different psychiatric disorders and risks of CAD, HF and AF, with further

stratification for menopause status. Given that these CVDs have both shared and unique underlying disease biology, future research into understanding whether an increased risk of depression increases risk of CAD and HF through the same or unique biological pathways may be important for informing prevention strategies. Our results warrant further investigation into the predictive value of depression risk even in the absence of a depression diagnosis in CVD risk calculators, particularly for females, as well as how CAD and HF risk differs across different menopausal stages.

Acknowledgments

We would like to thank the research participants and employees of 23andMe, inc. for making this work possible. The full GWAS summary statistics for the 23andMe discovery dataset 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/#dataset-access/ for more information and to apply to access the data.

The dataset(s) used for the BioVU analyses described were obtained from Vanderbilt University Medical Center's BioVU which is supported by numerous sources: institutional funding, private agencies, and federal grants. These include the NIH funded Shared Instrumentation Grant S10RR025141; and CTSA grants UL1TR002243, UL1TR000445, and UL1RR024975. Genomic data are also supported by investigator-led projects that include U01HG004798, R01NS032830, RC2GM092618, P50GM115305, U01HG006378, U19HL065962, R01HD074711; and additional funding sources listed at https://victr.vumc.org/biovu-funding/.

This publication (journal article, etc.) was supported by Grant No. R01 HG011405, partially funded by the Office of Research on Women's Health, Office of the Director, NIH and the National Human Genome Research Institute. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the Office of Research on Women's Health or the National Human Genome Research Institute.

The scripts used in the study are available on: https://github.com/CNSGenomics/CVD_psych_PRS_Jiang.

Competing interests

Authors declare no competing interests.

References

1. Graboys TB. Stress and the aching heart. N Engl J Med. 1984;311(9):594-5.
2. Correll CU, Solmi M, Veronese N, Bortolato B, Rosson S, Santonastaso P, et al. Prevalence, incidence and mortality from cardiovascular disease in patients with pooled and specific severe mental illness: a large-scale meta-analysis of 3,211,768 patients and 113,383,368 controls. World Psychiatry. 2017;16(2):163-80.
3. Marano G, Traversi G, Romagnoli E, Catalano V, Lotrionte M, Abbate A, et al. Cardiologic side effects of psychotropic drugs. J Geriatr Cardiol. 2011;8(4):243-53.
4. Larsen BA, Christenfeld NJS. Cardiovascular disease and psychiatric comorbidity: the potential role of perseverative cognition. Cardiovasc Psychiatry Neurol. 2009;2009:791017-.
5. Hasin DS, Sarvet AL, Meyers JL, Saha TD, Ruan WJ, Stohl M, et al. Epidemiology of adult DSM-5 major depressive disorder and its specifiers in the United States. JAMA Psychiatry. 2018;75(4):336-46.
6. Marcus SM, Kerber KB, Rush AJ, Wisniewski SR, Nierenberg A, Balasubramani GK, et al. Sex differences in depression symptoms in treatment-seeking adults: confirmatory analyses from the Sequenced Treatment Alternatives to Relieve Depression study. Compr Psychiatry. 2008;49(3):238-46.
7. Ochoa S, Usall J, Cobo J, Labad X, Kulkarni J. Gender differences in schizophrenia and first-episode psychosis: a comprehensive literature review. Schizophr Res Treatment. 2012;2012:916198.
8. Vega P, Barbeito S, de Azúa SR, Martínez-Cengotitabengoa M, González–Ortega I, Saenz M, et al. Bipolar disorder differences between genders: special considerations for women. Womens Health. 2011;7(6):663-76.

9. Wilkins JT, Ning H, Berry J, Zhao L, Dyer AR, Lloyd-Jones DM. Lifetime risk and years lived free of total cardiovascular disease. JAMA. 2012;308(17):1795-801.

10. Vogel B, Acevedo M, Appelman Y, Bairey Merz CN, Chieffo A, Figtree GA, et al. The Lancet women and cardiovascular disease Commission: reducing the global burden by 2030. The Lancet. 2021;397(10292):2385-438.

11. Möller-Leimkühler AM. Gender differences in cardiovascular disease and comorbid depression. Dialogues Clin Neurosci. 2007;9(1):71-83.

12. Williams SA, Kasl SV, Heiat A, Abramson JL, Krumholz HM, Vaccarino V. Depression and risk of heart failure among the elderly: a prospective community-based study. Psychosom Med. 2002;64(1):6-12.

13. Shah AJ, Ghasemzadeh N, Zaragoza-Macias E, Patel R, Eapen DJ, Neeland IJ, et al. Sex and age differences in the association of depression with obstructive coronary artery disease and adverse cardiovascular events. J Am Heart Assoc. 2014;3(3):e000741.

14. Meng R, Yu C, Liu N, He M, Lv J, Guo Y, et al. Association of depression with all-cause and cardiovascular disease mortality among adults in China. JAMA Network Open. 2020;3(2):e1921043-e.

15. Lu Y, Wang Z, Georgakis MK, Lin H, Zheng L. Genetic liability to depression and risk of coronary artery disease, myocardial infarction, and other cardiovascular outcomes. J Am Heart Assoc. 2021;10(1):e017986.

16. Dennis J, Sealock J, Levinson RT, Farber-Eger E, Franco J, Fong S, et al. Genetic risk for major depressive disorder and loneliness in sex-specific associations with coronary artery disease. Mol Psychiatry. 2021;26(8):4254-64.

17. 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. 2018;562(7726):203-9.

18. Yengo L, Sidorenko J, Kemper KE, Zheng Z, Wood AR, Weedon MN, et al. Meta-analysis of genome-wide association studies for height and body mass index in ∼700000 individuals of European ancestry. Hum Mol Genet. 2018;27(20):3641-9.

19. Lloyd-Jones LR, Zeng J, Sidorenko J, Yengo L, Moser G, Kemper KE, et al. Improved polygenic prediction by Bayesian multiple regression on summary statistics. Nat Commun. 2019;10(1):5086.

20. Ni G, Zeng J, Revez JA, Wang Y, Zheng Z, Ge T, et al. A comparison of ten polygenic score methods for psychiatric disorders applied across multiple cohorts. Biol Psychiatry. 2021;90(9):611-20.

21. Chang CC, Chow CC, Tellier LCAM, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience. 2015;4(1):s13742-015-0047-8.

22. Ge T, Chen C-Y, Ni Y, Feng Y-CA, Smoller JW. Polygenic prediction via Bayesian regression and continuous shrinkage priors. Nat Commun. 2019;10(1):1776.

23. Mangiafico S. rcompanion: Functions to support extension education program evaluation. 2022.

24. Robin X, Turck N, Hainard A, Tiberti N, Lisacek F, Sanchez J-C, et al. pROC: an open-source package for R and S+ to analyze and compare ROC curves. BMC Bioinformatics. 2011;12(1):77.

25. Zhang F, Cao H, Baranova A. Shared genetic liability and causal associations between major depressive disorder and cardiovascular diseases. Front Cardiovasc Med. 2021;8.

26. Khoudary SRE, Aggarwal B, Beckie TM, Hodis HN, Johnson AE, Langer RD, et al. Menopause transition and cardiovascular disease risk: implications for timing of early prevention: a scientific statement from the American Heart Association. Circulation. 2020;142(25):e506-e32.

27. So H-C, Sham PC. Exploring the predictive power of polygenic scores derived from genome-wide association studies: a study of 10 complex traits. Bioinformatics. 2017;33(6):886-92.

28. O'Sullivan JW, Raghavan S, Marquez-Luna C, Luzum JA, Damrauer SM, Ashley EA, et al. Polygenic risk scores for cardiovascular disease: a scientific statement from the American Heart Association. Circulation. 2022;146(8):e93-e118.

29. Sinnott-Armstrong N, Tanigawa Y, Amar D, Mars N, Benner C, Aguirre M, et al. Genetics of 35 blood and urine biomarkers in the UK Biobank. Nat Genet. 2021;53(2):185-94.

30. Hippisley-Cox J, Coupland C, Brindle P. Development and validation of QRISK3 risk prediction algorithms to estimate future risk of cardiovascular disease: prospective cohort study. BMJ. 2017;357:j2099.

31. Khandaker GM, Zuber V, Rees JMB, Carvalho L, Mason AM, Foley CN, et al. Shared mechanisms between coronary heart disease and depression: findings from a large UK general population-based cohort. Mol Psychiatry. 2020;25(7):1477-86.

32. Li L, Gower BA, Shelton RC, Wu X. Gender-specific relationship between obesity and major depression. Front Endocrinol (Lausanne). 2017;8:292.

33. Tobaldini E, Carandina A, Toschi-Dias E, Erba L, Furlan L, Sgoifo A, et al. Depression and cardiovascular autonomic control: a matter of vagus and sex paradox. Neurosci Biobehav Rev. 2020;116:154-61.

34. Schneider B, Sechtem U. Influence of age and gender in Takotsubo Syndrome. Heart Fail Clin. 2016;12(4):521-30.

35. Ridker PM, Rane M. Interleukin-6 signaling and anti-Interleukin-6 therapeutics in cardiovascular disease. Circ Res. 2021;128(11):1728-46.

36. Kaptoge S, Seshasai SRK, Gao P, Freitag DF, Butterworth AS, Borglykke A, et al. Inflammatory cytokines and risk of coronary heart disease: new prospective study and updated meta-analysis. Eur Heart J. 2013;35(9):578-89.

37. Boardman HMP, Hartley L, Eisinga A, Main C, Roqué i Figuls M, Bonfill Cosp X, et al. Hormone therapy for preventing cardiovascular disease in post-menopausal women. Cochrane Database of Systematic Reviews. 2015(3).

38. Cunningham R, Poppe K, Peterson D, Every-Palmer S, Soosay I, Jackson R. Prediction of cardiovascular disease risk among people with severe mental illness: A cohort study. PLoS One. 2019;14(9):e0221521-e.

39. Ministry of Health New Zealand. Cardiovascular disease risk assessment and management for primary care. 2018.

40. Lichtman JH, Froelicher ES, Blumenthal JA, Carney RM, Doering LV, Frasure-Smith N, et al. Depression as a risk factor for poor prognosis among patients with acute coronary syndrome: systematic review and recommendations. Circulation. 2014;129(12):1350-69.

41. Tang B, Yuan S, Xiong Y, He Q, Larsson SC. Major depressive disorder and cardiometabolic diseases: a bidirectional Mendelian randomisation study. Diabetologia. 2020;63(7):1305-11.

42. Martin J, Khramtsova EA, Goleva SB, Blokland GAM, Traglia M, Walters RK, et al. Examining sex-differentiated genetic effects across neuropsychiatric and behavioral traits. Biol Psychiatry. 2021;89(12):1127-37.

43. Deloukas P, Kanoni S, Willenborg C, Farrall M, Assimes TL, Thompson JR, et al. Large-scale association analysis identifies new risk loci for coronary artery disease. Nat Genet. 2013;45(1):25-33.

44. Christophersen IE, Rienstra M, Roselli C, Yin X, Geelhoed B, Barnard J, et al. Large-scale analyses of common and rare variants identify 12 new loci associated with atrial fibrillation. Nat Genet. 2017;49(6):946-52.

45. Stahl EA, Breen G, Forstner AJ, McQuillin A, Ripke S, Trubetskoy V, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. Nat Genet. 2019;51(5):793-803.

46. Nikpay M, Goel A, Won H-H, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive 1000 Genomes–based genome-wide association meta-analysis of coronary artery disease. Nat Genet. 2015;47(10):1121-30.

47. Shah S, Henry A, Roselli C, Lin H, Sveinbjörnsson G, Fatemifar G, et al. Genome-wide association and Mendelian randomisation analysis provide insights into the pathogenesis of heart failure. Nat Commun. 2020;11(1):163.

48. 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. Nat Genet. 2018;50(5):668-81.

49. Pardiñas AF, Holmans P, Pocklington AJ, Escott-Price V, Ripke S, Carrera N, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nat Genet. 2018;50(3):381-9.

50. Mullins N, Forstner AJ, O'Connell KS, Coombes B, Coleman JRI, Qiao Z, et al. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. Nat Genet. 2021;53(6):817-29.

51. Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirali M, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. Nat Neurosci. 2019;22(3):343-52.

52. Trubetskoy V, Pardiñas AF, Qi T, Panagiotaropoulou G, Awasthi S, Bigdeli TB, et al. Mapping genomic loci implicates genes and synaptic biology in schizophrenia. Nature. 2022;604(7906):502-8.

Supplementary Material

of

A higher genetic risk for depression is a greater risk factor for cardiovascular disease in women

Jiayue-Clara Jiang (PhD)[1*], Kritika Singh (BA)[2,3*], Lea K. Davis (PhD)[2,3,4,5,6], Naomi R. Wray (PhD)[1,7], Sonia Shah (PhD)[1]

[1] Institute for Molecular Bioscience, The University of Queensland, St Lucia, 4072, Australia

[2] Division of Genetic Medicine, Department of Medicine, Vanderbilt University Medical Center, Nashville, TN, 37232, USA

[3] Vanderbilt Genetics Institute, Vanderbilt University Medical Center, Nashville, TN, 37232, USA

[4] Department of Molecular Physiology and Biophysics, Vanderbilt University Medical Center, Nashville, TN, 37232, USA

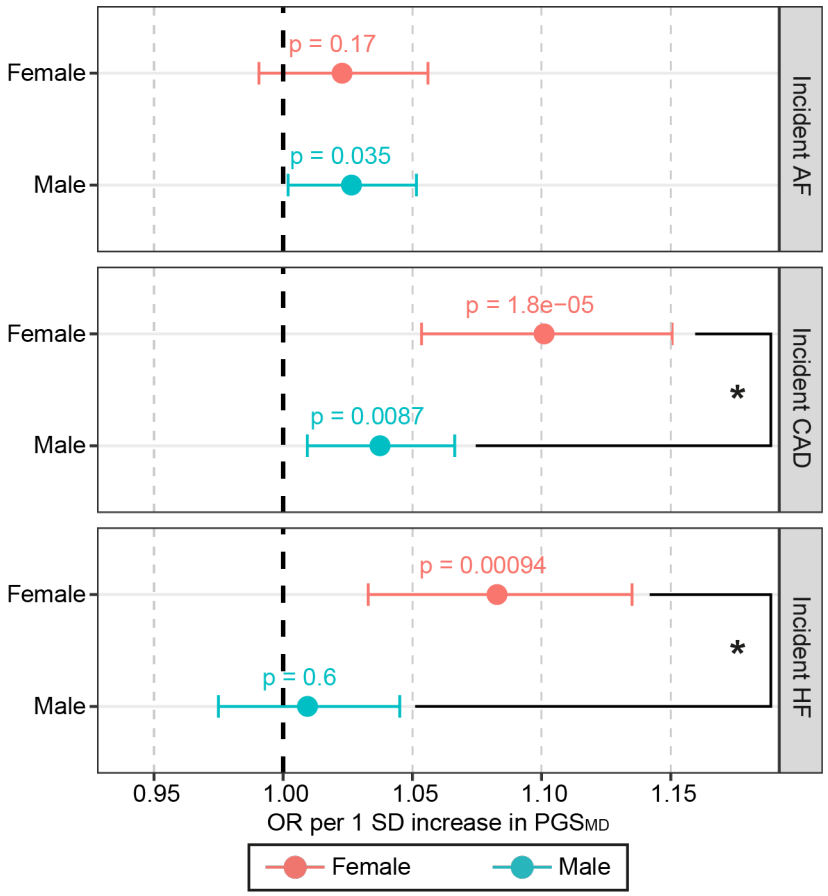
[5] Department of Psychiatry and Behavioral Sciences, Vanderbilt University Medical Center, Nashville, TN, 37232, USA

[6] Departments of Medicine and Biomedical Informatics, Vanderbilt University Medical Center, Nashville, TN, 37232, USA

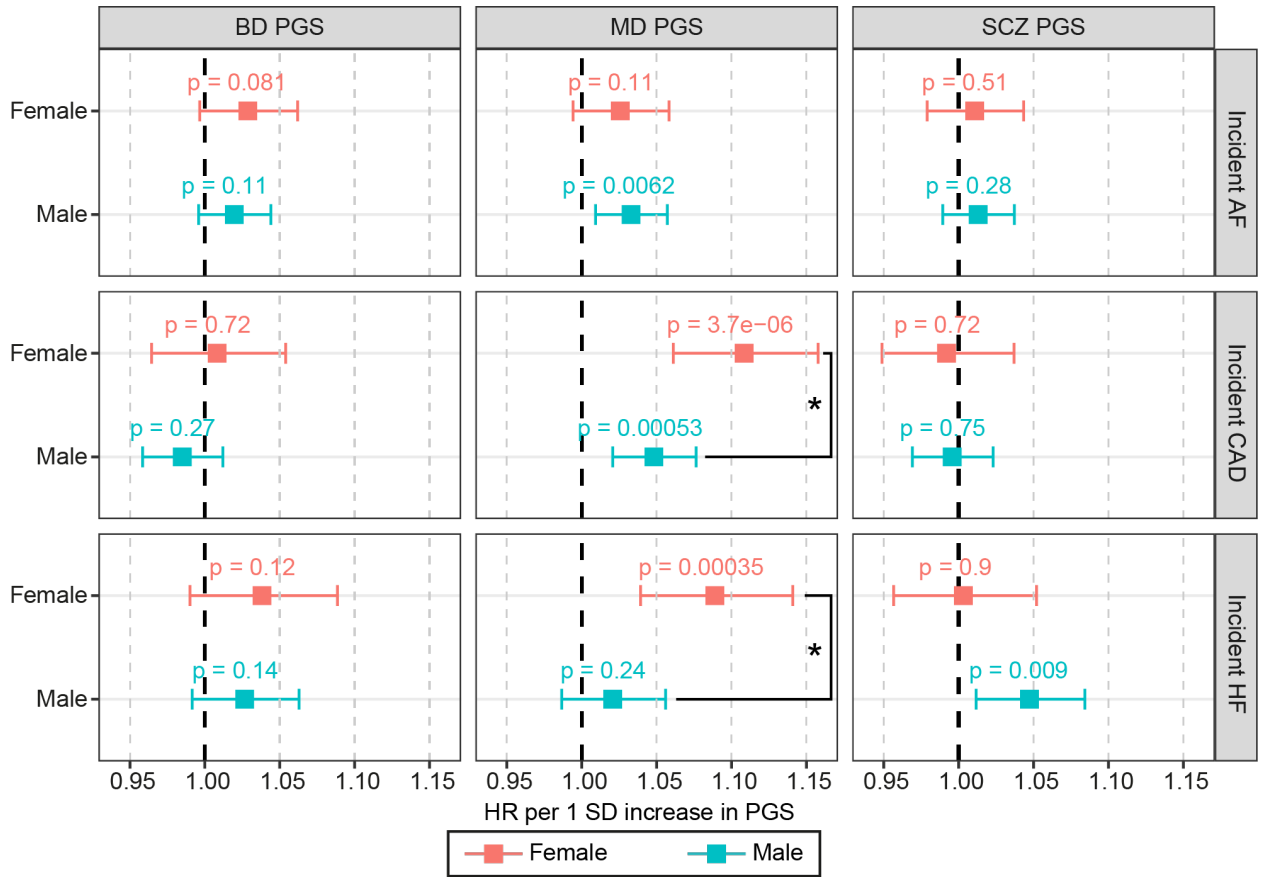
[7] Queensland Brain Institute, The University of Queensland, St Lucia, 4072, Australia

*co first authors

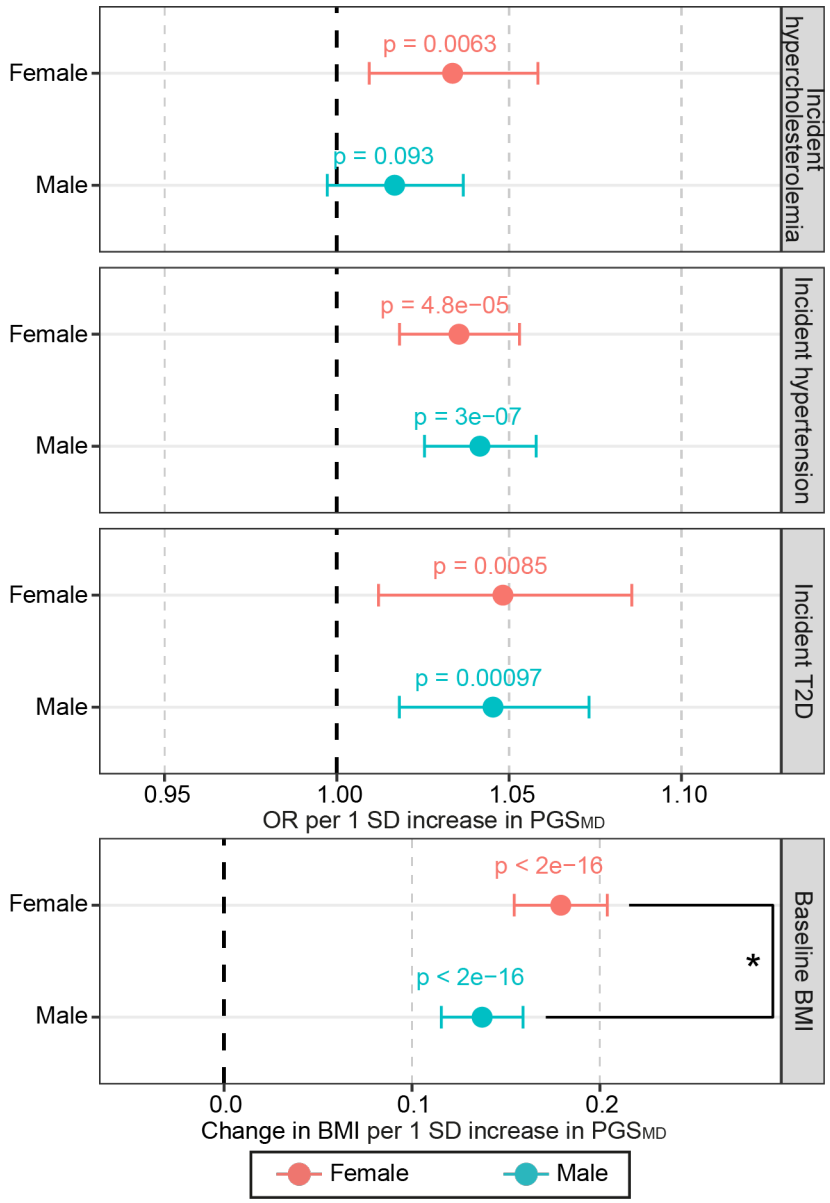
Supplementary Figures



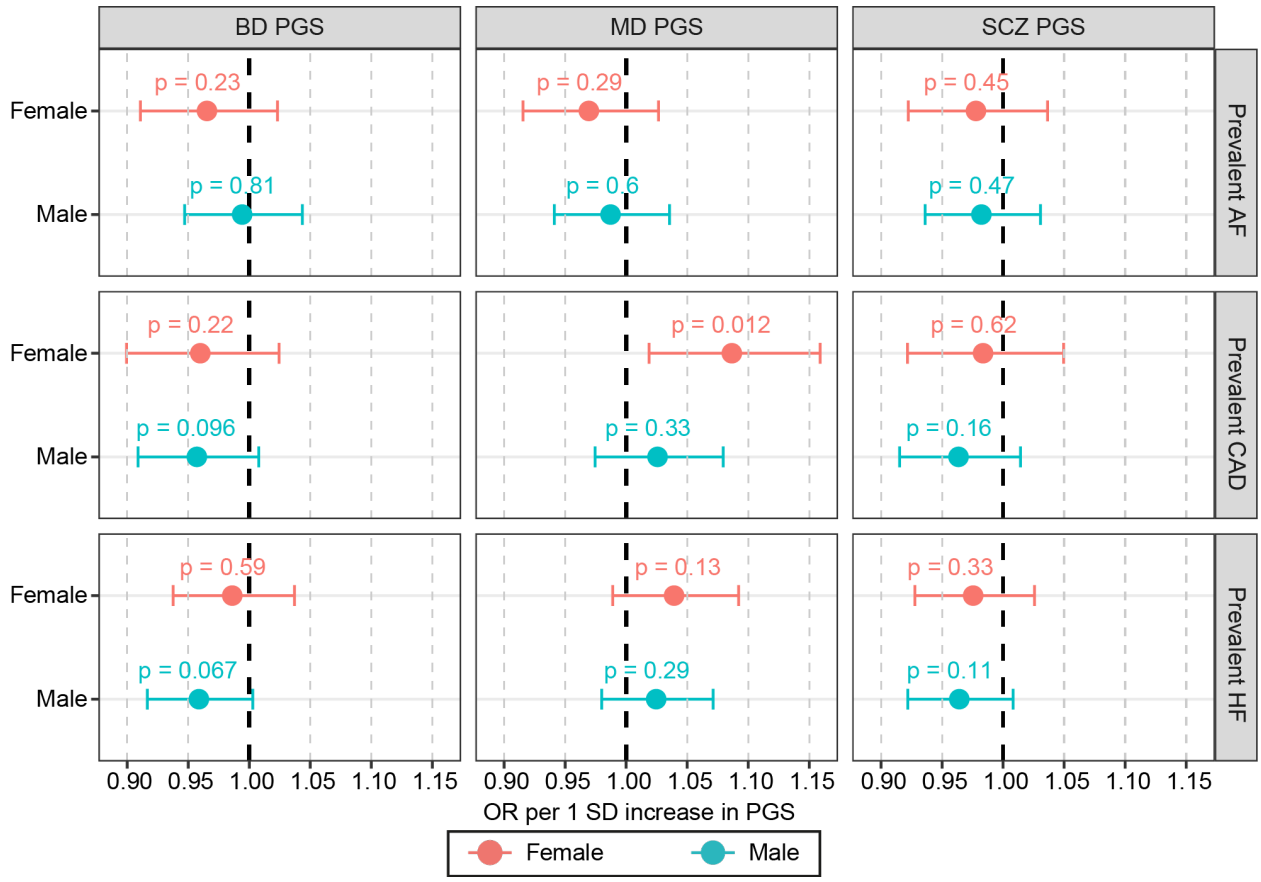
Supplementary Figure 1. Sensitivity analyses on association between PGS_{MD} and incident CVDs, performed on the UK Biobank cohort, where the corresponding CVD PGS was added as a covariate to the model. Associations for the sex-stratified cohorts (red - female; blue - male) were estimated with logistic regression. X-axis shows the OR per SD increase in PGS, with p-values labelled, and error bars indicate 95% confidence intervals. Asterisk (*) indicates a statistically significant difference in the beta values between females and males (two-sided Wald test $p < 0.05$). Dark grey line indicates OR of 1. AF, atrial fibrillation; CAD, coronary artery disease; HF, heart failure; OR, odds ratio; PGS, polygenic score; SD, standard deviation.



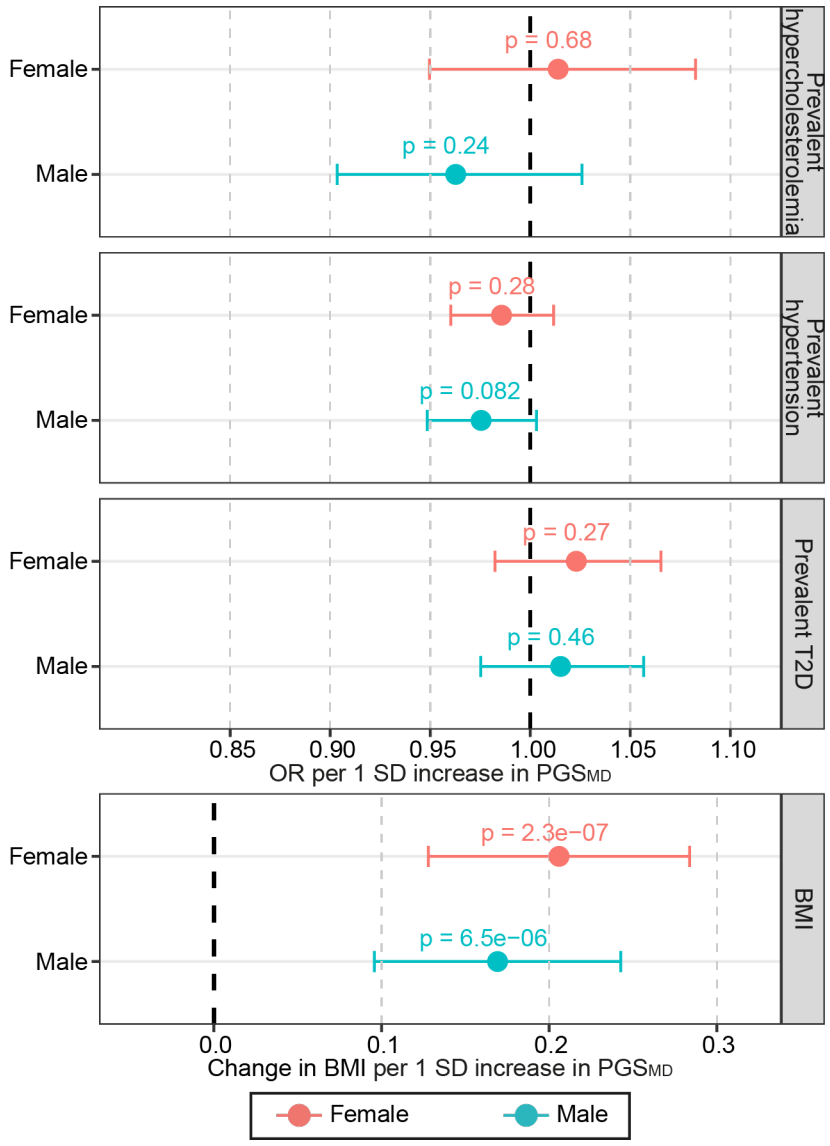
Supplementary Figure 2. Age-specific change in risk of incident CVDs per SD increase in psychiatric disorder PGSs in the UK Biobank cohort. Associations for the sex-stratified cohorts (red - female; blue - male) were estimated with Cox proportional hazards regression models. X-axis shows the HR per SD increase in PGS, with p-values labelled, and error bars indicate 95% confidence intervals. Asterisk (*) indicates a statistically significant difference in the beta values between females and males (two-sided Wald test $p < 0.05$). Dark grey line indicates HR of 1. AF, atrial fibrillation; BD, bipolar disorder; CAD, coronary artery disease; HF, heart failure; HR, hazard ratio; MD, major depression; PGS, polygenic score; SCZ, schizophrenia; SD, standard deviation.



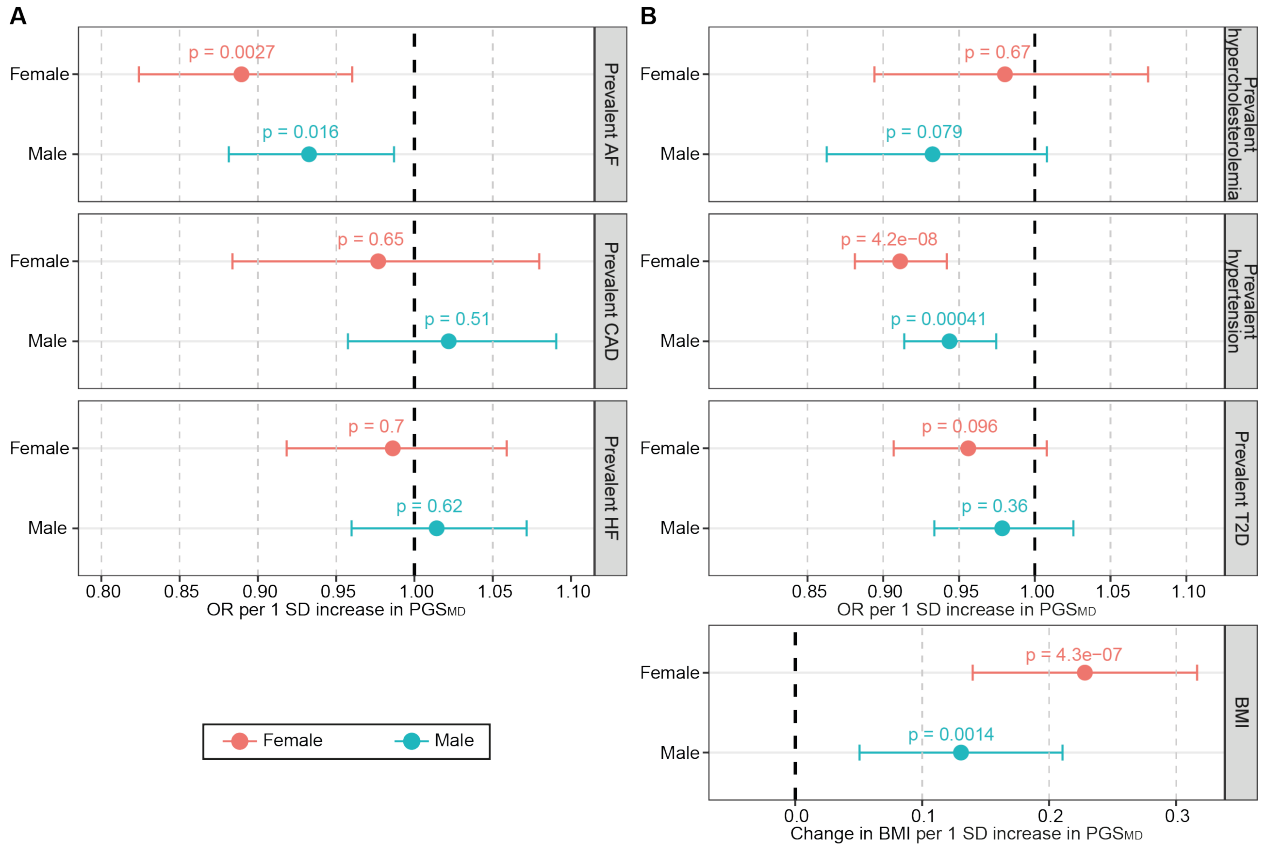
Supplementary Figure 3. Sex-specific associations of PGS_{MD} with CVD-associated risk factors amongst UK Biobank individuals with no diagnosis of psychiatric disorders and psychiatric medications. Associations are shown for the sex-stratified cohorts (red - female; blue - male). Asterisk (*) indicates a statistically significant difference (two-sided Wald test $p < 0.05$) in the risk estimates between females and males. OR was estimated with logistic regression, and change in BMI was estimated with linear regression. X-axis shows the OR or change in BMI per SD increase in PGS, with p-values labelled, and error bars indicate 95% confidence intervals. Dark grey line indicates OR of 1 or change of 0. BMI, body mass index; OR, odds ratio; PGS, polygenic score; SD, standard deviation; T2D, type II diabetes.



Supplementary Figure 4. Association between psychiatric disorder PGSs and prevalent CVDs in the BioVU cohort. Associations for the sex-stratified cohorts (red - female; blue - male) were estimated with logistic regression. X-axis shows the OR per SD increase in PGS, with p-values labelled, and error bars indicate 95% confidence intervals. Dark grey line indicates OR of 1. AF, atrial fibrillation; BD, bipolar disorder; CAD, coronary artery disease; HF, heart failure; MD, major depression; OR, odds ratio; PGS, polygenic score; SCZ, schizophrenia; SD, standard deviation.



Supplementary Figure 5. Association between the genetic risk of MD and CVD risk-factors in the BioVU cohort. Associations are shown for sex-stratified (red - female; blue - male) cohorts. OR was estimated with logistic regression, and change in BMI was estimated with linear regression. X-axis shows the OR or change in BMI per SD increase in PGS, with p-values labelled, and error bars indicate 95% confidence intervals. Dark grey line indicates OR of 1 or change of 0. BMI, body mass index; OR, odds ratio; PGS, polygenic score; SD, standard deviation; T2D, type II diabetes.



Supplementary Figure 6. Association between PGS_{MD} and prevalent CVDs or CVD-associated risk factors amongst individuals of the BioVU cohort who had no psychiatric disorder diagnosis. OR was estimated with logistic regression, and change in BMI was estimated with linear regression. X-axis shows the OR per SD increase in PGS, with p-values labelled, and error bars indicate 95% confidence intervals. Dark grey line indicates OR of 1. AF, atrial fibrillation; BMI, body mass index; CAD, coronary artery disease; HF, heart failure; OR, odds ratio; PGS, polygenic score; SD, standard deviation; T2D, type II diabetes.

1 **Supplementary Tables**

2 Supplementary Table 1. UK Biobank CVD and associated risk factors phenotype ascertainment criteria

Disease	Criteria	Data field	Data field description	Codes
AF	Inclusion	20002	Non-cancer illness code, self-reported	1471, 1483
		20004	Operation code	1524
		41270	ICD10	I48, I480, I481, I482, I483, I484, I489
		41271	ICD9	4273
		41272	Operative procedures - OPCS4	K571, K621, K622, K623, K624
HF	Inclusion	20002	Non-cancer illness code, self-reported	1076, 1079
		41270	ICD10	I110, I130, I132, I255, I420, I425, I428, I429, I500, I501, I509
		41271	ICD9	4254, 4280, 4281, 4289
	Exclusion	20002	Non-cancer illness code, self-reported	1588
		41270	ICD10	I421, I422
CAD	Inclusion	20002	Non-cancer illness code, self-reported	1075
		41270	ICD10	I210, I211, I212, I213, I214, I219, I220, I221, I228, I229, I230, I231, I232, I233, I235, I236, I238, I241, I252
		41271	ICD9	4109, 4119, 4129, 4299
		41272	Operative procedures - OPCS4	K401, K402, K403, K404, K411, K412, K413, K414, K451, K452, K453, K454, K455, K491, K492, K498, K499, K502, K751, K752, K753, K754, K758, K759
Hyper-cholesterolemia	Inclusion for cases	41270	ICD10	E780
		41271	ICD9	27200, 27209
		20002	Non-cancer illness code, self-reported	1473
	Exclusion from controls	6177	Medication for cholesterol, blood pressure or diabetes	1 (Cholesterol lowering medication)
		6153	Medication for cholesterol, blood pressure, diabetes, or take exogenous hormones	1 (Cholesterol lowering medication)
Hypertension	Inclusion for cases	41270	ICD10	I10
		41271	ICD9	4010, 4011, 4019
		20002	Non-cancer illness code, self-reported	1065, 1072
	Exclusion from controls	6177	Medication for cholesterol, blood pressure or diabetes	2 (Blood pressure medication)
		6153	Medication for cholesterol, blood pressure, diabetes, or take exogenous hormones	2 (Blood pressure medication)
T2D	Inclusion for cases	41270	ICD10	E110, E111, E112, E113, E114, E115, E116, E117, E118, E119
		41271	ICD9	25000, 25010
		20002	Non-cancer illness code, self-reported	1223
	Exclusion from controls	41270	ICD10	E100, E101, E102, E103, E104, E105, E106, E107, E108, E109
		41271	ICD9	25001, 25011
		20002	Non-cancer illness code, self-reported	1222
		6177	Medication for cholesterol, blood pressure or diabetes	3 (Insulin)
		6153	Medication for cholesterol, blood pressure, diabetes, or take exogenous hormones	3 (Insulin)

1 Supplementary Table 2. UK Biobank psychiatric disorder phenotype ascertainment criteria

Disease	Criteria	Data field	Data field description	Codes
BD	Inclusion	41270	ICD10	F300, F301, F302, F308, F309, F310, F311, F312, F313, F314, F315, F316, F317, F318, F319
		41271	ICD9	2960, 2961, 2969
SCZ	Inclusion	41270	ICD10	F200, F201, F202, F203, F204, F205, F206, F208, F209, F21, F220, F228, F229, F230, F231, F232, F233, F238, F239, F24, F250, F251, F252, F258, F259, F28
		41271	ICD9	2953, 2959
MD	Inclusion	41270	ICD10	F320, F321, F322, F323, F328, F329, F330, F331, F332, F333, F334, F338, F339, F340, F341, F348, F349, F380, F381, F388, F39
		41271	ICD9	3119, 2962
	Exclusion	BD cases		
		SCZ cases		
		41270	ICD10	F448
		20002	Non-cancer illness code, self-reported	1289
		20003	Treatment/medication code	Antipsychotics: 1141202024, 1141153490, 1141195974, 1140867078, 1140867494, 1141171566, 2038459704, 1140872064, 1140879658, 1140867342, 1140867420, 1140882320, 1140872216, 1140910358, 1141200458, 1141172838, 1140867306, 1140867180, 1140872200, 1140867210, 1140867398, 1140882098, 1140867184, 1140867168, 1140863416, 1140909802, 1140867498, 1140867490, 1140910976, 1140867118, 1140867456, 1140928916, 1140872268, 1140867134, 1140867208, 1140867218, 1140867572, 1140879674, 1140909804, 1140867504, 1140868170, 1140879746, 1141152848, 1141177762, 1140867444, 1140867092, 1141152860, 1140872198, 1140867244, 1140868172, 1140867304, 1140872072, 1140879750, 1140868120, 1140872214, 1141201792, 1140882100, 1141167976
No diagnosis /medication	Exclusion from full cohort	BD cases		
		MD cases		
		SCZ cases		
		20002	Non-cancer illness code, self-reported	1286, 1287, 1288, 1289, 1290, 1291, 1531
		20003	Treatment/medication code	Antipsychotics (see above)
				Antidepressants: 1140867820, 1140867948, 1140879616, 1140867938, 1140867690, 1141190158, 1141151946, 1140921600, 1140879620, 1141201834, 1140867152, 1140909806, 1140879628, 1140867640, 1141200564, 1141151982, 1140916288, 1141180212, 1140867860, 1140867952, 1140879540, 1140867150, 1140909800, 1140867940, 1140879544, 1140879630, 1140867856, 1140867726, 1140867884, 1140867922, 1140910820, 1140879556, 1141152732, 1140867920, 1140882244, 1140867852, 1140867818, 1141174756, 1140867916, 1140867888, 1140867850, 1140867624, 1140867876, 1141151978, 1140882236, 1140867878, 1201, 1140882312, 1140867758, 1140867712, 1140867914, 1140867944, 1140879634, 1140867756, 1140867934, 1140867960, 1140916282, 1141200570, 1141152736
		41270	ICD10	F21, F29, F419, F410, F411, F418, F449, F444, F446, F440, F441, F448, F681, F688, F409, F400, F401, F402, F408, F341, F488, F489, F99, F601, F604, F607, F609
		41271	ICD9	3000, 3001, 3002, 3004, 3005, 3009, 3012, 3015, 3016, 3019

2

Supplementary Table 3. Summary of characteristics of participants in the UK Biobank cohort who had no diagnosis of psychiatric disorders and were not on any antidepressants or antipsychotics

		Female		Male	
Total Number		155517		141304	
Average age at baseline in years (SD)		57.1 (7.9)		57.4 (8.1)	
Average BMI at baseline (SD)		26.7 (5)		27.7 (4.2)	
Current or previous smokers at baseline (%)		62444 (40.2)		71099 (50.3)	
		Case	Control	Case	Control
prevalent AF	N (%)	4243 (2.7)	151274 (97.3)	9246 (6.5)	132058 (93.5)
	Age at baseline (SD)	62.9 (5.7)	56.9 (7.9)	62.6 (6)	57.1 (8.1)
incident AF	N (%)	3096 (2)	151274 (97.3)	6221 (4.4)	132058 (93.5)
	Age at baseline (SD)	63 (5.6)	56.9 (7.9)	62.6 (5.9)	57.1 (8.1)
prevalent CAD	N (%)	3064 (2)	152453 (98)	11907 (8.4)	129397 (91.6)
	Age at baseline (SD)	62.5 (6)	56.9 (7.9)	62.1 (6.1)	57 (8.1)
incident CAD	N (%)	1461 (0.9)	152453 (98)	4550 (3.2)	129397 (91.6)
	Age at baseline (SD)	62.3 (6.1)	56.9 (7.9)	61.5 (6.5)	57 (8.1)
prevalent HF	N (%)	1511 (1)	153923 (99)	3799 (2.7)	137341 (97.2)
	Age at baseline (SD)	62.9 (6)	57 (7.9)	62.7 (6)	57.3 (8.1)
incident HF	N (%)	1218 (0.8)	153923 (99)	2696 (1.9)	137341 (97.2)
	Age at baseline (SD)	63 (5.9)	57 (7.9)	62.8 (6)	57.3 (8.1)
incident hyper-cholesterolemia	N (%)	7766 (5)	131173 (84.3)	12529 (8.9)	101220 (71.6)
	Age at baseline (SD)	61.6 (6.1)	56.2 (7.9)	61.5 (6.4)	55.8 (8.2)
incident hypertension	N (%)	18146 (11.7)	112981 (72.6)	23620 (16.7)	86838 (61.5)
	Age at baseline (SD)	61 (6.6)	55.8 (7.9)	60.9 (6.7)	55.6 (8.2)
incident T2D	N (%)	3479 (2.2)	150032 (96.5)	6355 (4.5)	131251 (92.9)
	Age at baseline (SD)	60.9 (6.7)	56.9 (7.9)	61.2 (6.6)	57.2 (8.1)

Supplementary Table 4. BioVU phenotype ascertainment criteria

Disease	Criteria	Codes
AF	Inclusion	I48, I48.0, I48.1, I48.2, I48.3, I48.4, I48.9
CAD	Inclusion	I21.0, I21.1, I21.2, I21.3, I21.4, I21.9, I22.0, I22.1, I22.8, I22.9, I23.0, I23.1, I23.2, I23.3, I23.5, I23.6, I23.8, I24.1, I25.2
HF	Inclusion	I11.0, I13.0, I13.2, I25.5, I42.0, I42.5, I42.8, I42.9, I50.0, I50.1, I50.9
	Exclusion	I42.1, I42.2
Hypercholesteremia	Inclusion	E78.0
Hypertension	Inclusion	I10
T2D	Inclusion	E11.0, E11.1, E11.2, E11.3, E11.4, E11.5, E11.6, E11.7, E11.8, E11.9
	Exclusion from controls	E10.0, E10.1, E10.2, E10.3, E10.4, E10.5, E10.6, E10.7, E10.8, E10.9
BD	Inclusion	F30.0, F30.1, F30.2, F30.8, F30.9, F31.0, F31.1, F31.2, F31.3, F31.4, F31.5, F31.6, F31.7, F31.8, F31.9
MD	Inclusion	F32.0, F32.1, F32.2, F32.3, F32.4, F32.5, F32.8, F32.9, F33.0, F33.1, F33.2, F33.3, F33.4, F33.8, F33.9, F34.0, F34.1, F34.8, F34.9, F38.0, F38.1, F38.8, F39
SCZ	Inclusion	F20.0, F20.1, F20.2, F20.3, F20.4, F20.5, F20.6, F20.8, F20.9, F21, F22.0, F22.8, F22.9, F23.0, F23.1, F23.2, F23.3, F23.8, F23.9, F24, F25.0, F25.1, F25.2, F25.8, F25.9, F28
Neurotic disorders	Inclusion	F21, F34.1, F40.0, F40.1, F40.2, F40.8, F40.9, F41.0, F41.1, F41.8, F41.9, F44.0, F44.1, F44.4, F44.6, F44.8, F44.9, F48.8, F48.9, F60.1, F60.4, F60.7, F60.9, F68.1, F68.8, F99
Psychosis	Inclusion	F29

Supplementary Table 5. Summary of phenotypes in the whole BioVU cohort

		Female		Male	
Total Number		35160		27481	
Median age in years (SD)		47.91 (21.09)		50.70 (22.37)	
Median BMI (SD)		27.87 (7.63)		27.51 (6.66)	
Tobacco Use Disorder (%)		3460 (9.84)		3971 (14.45)	
		Case	Control	Case	Control
BD	N (%)	659 (1.87)	34501 (98.13)	404 (1.47)	27077 (98.53)
	Median age (SD)	43.97 (17.29)	47.98 (21.15)	44.56 (18.37)	50.79 (22.41)
MD	N (%)	5565 (15.83)	29595(84.17)	2932 (10.67)	24549 (89.33)
	Median age (SD)	50.89 (18.33)	47.35 (21.52)	53.64 (18.54)	50.35 (22.77)
SCZ	N (%)	115 (0.33)	35045 (99.67)	103 (0.37)	27378 (99.63)
	Median age (SD)	45.14 (19.34)	47.92 (21.09)	41.75 (15.59)	50.73 (22.39)
hyper-cholesterolemia	N (%)	950 (2.7)	34210 (97.3)	1050 (3.82)	26431 (96.18)
	Median age (SD)	63.64 (12.35)	47.47 (21.11)	64.67 (11.95)	50.14 (22.5)
hypertension	N (%)	9290 (26.42)	25870 (73.58)	8945 (32.55)	18536 (67.45)
	Median age (SD)	59.78 (14.82)	43.64 (21.37)	60.87 (14.64)	45.79 (23.76)
T2D	N (%)	2877 (8.18)	32283 (91.81)	3118 (11.35)	24363 (88.65)
	Median age (SD)	59.46 (13.83)	46.88 (21.31)	62.71 (11.91)	49.16 (22.92)
AF	N (%)	1329 (3.78)	33831 (96.22)	2043 (7.43)	25438 (92.57)
	Median age (SD)	69.65 (11.40)	47.05 (20.92)	67.80 (11.54)	49.32 (22.46)
CAD	N (%)	1021 (2.9)	34139 (97.1)	1789 (6.51)	25692 (93.49)
	Median age (SD)	64.16 (13.30)	47.42 (21.08)	65.68 (11.17)	49.65 (22.58)
HF*	N (%)	1763 (5.01)	33293 (94.99)	2368 (8.62)	24986 (91.38)
	Median age (SD)	63.16 (16.45)	47.08 (21.00)	64.33 (14.78)	49.39 (22.55)

Supplementary Table 6. Summary of phenotypes in the BioVU cohort with no diagnosis of psychiatric disorders

		Female		Male	
Total Number		26807		22917	
Median age in years (SD)		47.33 (21.84)		50.47 (22.97)	
Median BMI (SD)		27.57 (7.6)		27.36 (6.66)	
Tobacco Use Disorder (%)		1930 (7.2)		2733 (11.93)	
		Case	Control	Case	Control
AF	N (%)	706 (2.63)	26101 (97.37)	1414 (6.17)	21503 (93.83)
	Median age (SD)	70.2 (11.21)	46.62 (21.64)	68.26 (11.62)	49.30 (22.97)
CAD	N (%)	396 (1.48)	26411 (95.52)	1063 (4.64)	21854 (95.36)
	Median age (SD)	66.15 (12.18)	46.96 (21.75)	67.21 (10.14)	49.66 (23.03)
HF	N (%)	801 (2.99)	25953 (97.01)	1471 (6.42)	21364 (93.58)
	Median age (SD)	64.16 (17.01)	46.72 (21.69)	65.44 (15.04)	49.43 (23.00)
hyper-cholesterolemia	N (%)	466 (1.74)	26341 (98.26)	688 (3.00)	22229 (97.00)
	Median age (SD)	64.23 (12.38)	46.95 (21.78)	65.51 (11.9)	50.01 (22.99)
hypertension	N (%)	4753 (17.73)	22054 (82.27)	5901 (25.75)	17016 (74.25)
	Median age (SD)	60.79 (15.10)	44.33 (21.87)	61.88 (14.68)	46.52 (23.88)
T2D	N (%)	1534 (5.72)	25158 (94.28)	2129 (9.29)	20708 (90.71)
	Median age (SD)	58.59 (16.5)	46.62 (21.85)	62.35 (13.28)	49.32 (23.32)

Supplementary Table 7. Sources of GWAS summary statistics used in PGS SNP weight estimation for UK Biobank and BioVU cohorts

Prediction cohort	Trait	Sample size	Ancestry	Reference
UK Biobank	AF	17931 cases and 115142 controls	Predominantly European	Christophersen et al. [44]
	BD	20352 cases and 31358 controls	European	Stahl et al. [45]
	CAD	60801 cases and 123504 controls	Predominantly European	Nikpay et al. [46]
	HF	40805 cases and 542362 controls	European	The full dataset (including UK Biobank individuals) is available in Shah et al. [47]). The dataset used in this study (with no UK Biobank individuals) is unpublished.
	MD	121198 cases and 329421 controls	European	The full dataset (including UK Biobank individuals) is available in Wray et al. [48]). The dataset used in this study (with no UK Biobank individuals) is unpublished.
	SCZ	40675 cases and 64643 controls	Predominantly European	Pardiñas et al. [49]
BioVU	BD	41917 cases and 371549 controls	European	Mullins et al. [50]
	MD	170756 cases and 329443 controls (excluding the 23andMe cohort)	European	Howard et al. [51]
	SCZ	53386 cases and 77258 controls	European	Trubetskoy et al. [52]

Supplementary Table 8. Prediction accuracy of PGSs generated for the UK Biobank cohort

Prediction cohort	PGS	Number of SNPs	N_{cases}	$N_{controls}$	OR per SD increase of PGS (95% CI)	p-value	AUC	Nagelkerke R^2
UK Biobank	AF	1002671	16217	328984	1.58 (1.55 - 1.6)	< 2E-16	0.77	0.030
	BD	1008170	1027	296821	2.04 (1.91 - 2.17)	< 2E-16	0.73	0.037
	CAD	1007431	18271	326930	1.63 (1.6 - 1.66)	< 2E-16	0.80	0.035
	HF	1006789	6942	337967	1.19 (1.16 – 1.22)	< 2E-16	0.77	0.0033
	MD	1007847	11441	296821	1.30 (1.27 - 1.32)	< 2E-16	0.65	0.0088
	SCZ	987863	823	296821	2.00 (1.86 - 2.15)	< 2E-16	0.72	0.034
BioVU	BD	778307	1063	12917	1.36 (1.28 - 1.45)	< 2E-16	0.73	0.0098
	MD	770389	8497	12917	1.14 (1.11 - 1.16)	< 2E-16	0.67	0.0033
	SCZ	779202	218	12917	1.55 (1.34 - 1.79)	6.5E-09	0.71	0.013

Supplementary Table 9. CVD incidence and average recruitment age of females who were baseline pre- and post-menopause in the UK Biobank (*** indicates a chi square p < 2.2E-16)

		Baseline pre-menopause		Baseline post-menopause		
		Case	Control	Case	Control	
AF	Count (%)	83 (0.3)	28287 (99.7)	3055 (2.7)	110208 (97.3)	***
	Average age at baseline in years (SD)	45.6 (2.7)	44.9 (2.6)	63.7 (4.5)	60.7 (5.4)	NA
CAD	Count (%)	74 (0.3)	28274 (99.7)	1530 (1.4)	111224 (98.6)	***
	Average age at baseline in years (SD)	45.7 (2.3)	44.9 (2.6)	63.3 (4.9)	60.7 (5.4)	NA
HF	Count (%)	62 (0.2)	28331 (99.8)	1346 (1.2)	112634 (98.8)	***
	Average age at baseline in years (SD)	45.4 (2.6)	44.9 (2.6)	63.6 (5)	60.7 (5.4)	NA