Circulating cardiovascular proteomic associations with genetics and disease

Kathryn A. McGurk, PhD[1,2,3,*], Lara Curran, MBBS, BSc[1,2,4], Arunashis Sau, MRCP, PhD[1,5], Fu Siong Ng, FRCP, PhD[1,5,6], Brian Halliday, FRCP, PhD[1,4], James S. Ware, PhD, MRCP[1,2,3,4,5], Declan P. O'Regan, FRCR, PhD[2]

[1]National Heart and Lung Institute, Imperial College London, London, UK.
[2]MRC Laboratory of Medical Sciences, Imperial College London, London, UK.
[3]Program in Medical and Population Genetics, The Broad Institute of MIT and Harvard, Cambridge, US.
[4]Royal Brompton & Harefield Hospitals, Guy's and St. Thomas' NHS Foundation Trust, London, UK.
[5]Department of Cardiology, Imperial College Healthcare NHS Trust, London, UK.
[6]Chelsea and Westminster Hospital NHS Foundation Trust, London, UK.
*Corresponding author: Dr. Kathryn A. McGurk, k.mcgurk@imperial.ac.uk, LMS building, Hammersmith Campus, Imperial College London, London, UK.

Keywords: proteomics, genetics, heart failure
Short title: Circulating cardiovascular proteomics
Word count: 5,379

Abstract

Background

The analysis of the circulating proteome can identify translational modifiers and biomarkers of disease expressivity and severity at a given time point. Here we explore the relationships between protein measures implicated in cardiovascular disease and whether they mediate causal relationships between cardiovascular risk factors and disease development.

Methods

To understand the relationships between circulating biomarkers and genetic variants, medications, anthropometric traits, lifestyle factors, imaging-derived measures, and diagnoses of cardiovascular disease, we analysed measures of nine plasma proteins with *a priori* roles in genetic and structural cardiovascular disease or treatment pathways (ACE2, ACTA2, ACTN4, BAG3, BNP, CDKN1A, NOTCH1, NT-proBNP, and TNNI3) from the Pharma Proteomics Project of the UK Biobank cohort (over 45,000 participants sampled at recruitment).

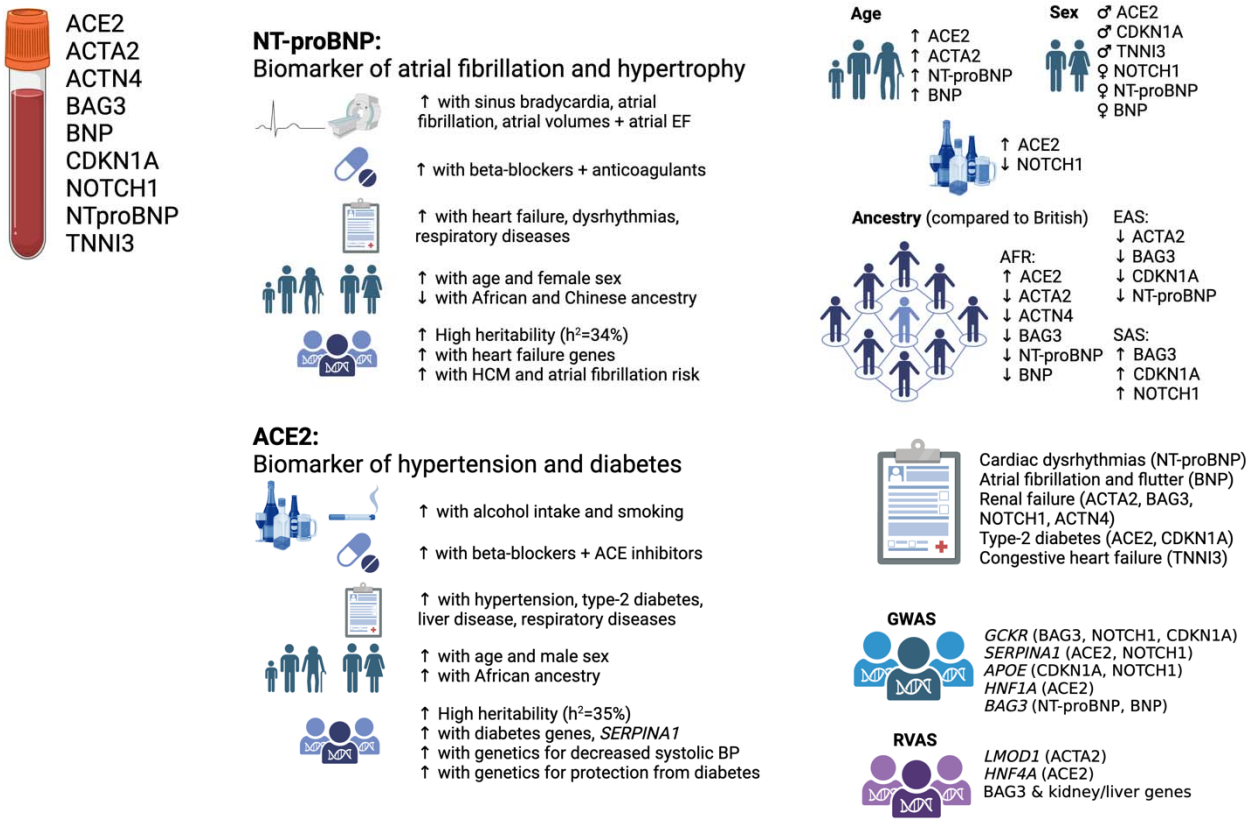
Results

We identified significant variability in circulating proteins with age, sex, ancestry, alcohol intake, smoking, and medication intake. Phenome-wide association studies highlighted the range of cardiovascular clinical features with relationships to protein levels. Genome-wide genetic association studies identified variants near *GCKR*, *APOE*, and *SERPINA1*, that modified multiple circulating protein levels (BAG3, CDKN1A, and/or NOTCH1). NT-proBNP and BNP levels associated with variants in *BAG3*. ACE2 levels were increased with a diagnosis of hypertension or diabetes and were influenced by

variants in genes associated with diabetes (*HNF1A, HNF4A*). Two-sample Mendelian randomisation identified ACE2 as protective for systolic blood pressure and Type-2 diabetes.

Conclusions

From a panel of circulating proteins, the results from this observational study provide evidence that ACE2 is causally associated with hypertension and diabetes. This suggests that ACE2 stimulation may provide additional protection from these cardiovascular diseases. This study provides an improved understanding of the circulating pathways depicting cardiovascular disease dynamics.



Graphical abstract

Introduction

Circulating proteomics provides information on the landscape of biological function, metabolism, disease, and homeostasis. While genetic testing can provide a once-off invariable assessment from birth, proteome analyses can identify translational modifiers and biomarkers of disease expressivity and severity at a particular sampling time point. In addition to large-scale proteome-wide discovery and risk score analyses, in-depth assessments of selected proteins based on *a priori* implication in disease are required to fully identify and interpret the relationships between protein measures of interest and cardiovascular disease, and whether they mediate causal relationships between cardiovascular risk factors and disease development.

To understand these relationships, we undertook an in-depth assessment of the circulating levels of nine plasma proteins implicated through genetic studies for roles in structural cardiovascular disease or treatment pathways (aortopathies, cardiomyopathies, congenital heart disease, and heart failure[1]): ACE2 (*ACE2*; angiotensin-converting enzyme 2), ACTA2 (*ACTA2*; actin alpha 2, smooth muscle), ACTN4 (*ACTN4;* Actinin Alpha 4), BAG3 (*BAG3*; BAG cochaperone 3), BNP (*NPPB;* brain natriuretic peptide or B-type natriuretic peptide), CDKN1A (*CDKN1A;* cyclin-dependent kinase inhibitor 1A), NOTCH1 (*NOTCH1; Neurogenic locus notch homolog protein 1*), NT-proBNP (*NPPB*; N-terminal prohormone of brain natriuretic peptide), and TNNI3 (*TNNI3*; Troponin I).

ACE2, with functionally opposing roles of ACE1 (which is unmeasured here, is the target of ACE inhibitors, and is involved in the biosynthesis of the angiotensin II vasoconstrictor), is a part of the renin-angiotensin-aldosterone system that regulates

blood pressure by catalysing the hydrolysis of vasoconstrictor angiotensin II to other vasodilator angiotensins. ACTA2 is an actin protein involved in the contraction of smooth muscle and genetic variants in *ACTA2* cause autosomal dominant familial thoracic aortic aneurysm and aortic dissection[2]. ACTN4 is an actin-binding protein that anchors the myofibrillar actin filaments in muscle cells. BAG3 is involved in chaperone-assisted selective autophagy of damaged cytoskeletal components and variants in the *BAG3* gene cause autosomal dominant dilated cardiomyopathy and myofibrillar myopathy[1,3]. The genetic locus of CDKN1A has been recently identified in case-control GWAS analyses of cardiomyopathies[4,5] and it is a regulator of cardiomyocyte cell cycle arrest[6]. NOTCH1 controls cell fate decisions and variants at the locus have been previously associated with congenital heart disease[7] and trabeculation[8]. PreproBNP and proBNP (both unmeasured here) are cleaved by a convertase to create BNP and NT-proBNP (inactive with a longer half-life than BNP); both are biomarkers and upregulated with heart failure and myocardial stretching. PreproBNP is encoded by *NPPB* which has unique expression in the heart, highest in the atrial appendage (GTEx)[9]. BNP has roles in natriuresis, diuresis and vasodilatation. TNNI3 is a mediator of relaxation in the sarcomeric thin filament of cardiac striated muscle and is exclusively expressed in the heart. Variants in *TNNI3* cause autosomal dominant hypertrophic cardiomyopathy[1,10].

Exploration of the relationships between circulating proteins and upstream factors that may influence their levels (e.g., anthropometric traits and protein quantitative trait loci identifiable from genome-wide common and rare genetic factors) as well as downstream clinical endpoints they may predict or prevent (cardiovascular diagnoses, clinical features, and ECG and MRI-derived traits), would aid our understanding of these

biomarkers of cardiovascular disease. Here, we use the results of the Pharma Proteomics Project of the UK Biobank cohort: Olink proteomic data from plasma samples collected at recruitment. We analysed the proteomics from over 45,000 participants to identify relationships with the full spectrum of available UK Biobank data and tested for causal relationships (**Graphical abstract**).

Methods

Study overview

The UK Biobank (UKB) cohort study recruited 500,000 participants aged 40 to 69 years old from across the United Kingdom between 2006 and 2010 (National Research Ethics Service, 11/NW/0382)[11]. This study was conducted under the terms of access of projects 47602 and 40616. All participants provided written informed consent.

Genotyping array data and exome sequencing data were available for over 450,000 participants. Sub-studies analysed baseline plasma proteomics in over 45,000 participants[12] and recalled participants for electrocardiograms (ECG)[13] and cardiac magnetic resonance (CMR) imaging[14,15]. Additional phenotypic and outcome data included hospital episode statistics, self-reported questionnaire data, alcohol intake[16], and smoking status. Protein levels were assessed for association with genetic, phenotypic, and clinical outcome data (**Figure 1**). Two-sample Mendelian randomization was used to determine causality and risk prediction of Cox proportional hazards regression models were created to understand whether the biomarkers predicted incident diagnoses.

Proteomics

Plasma from the initial UKB assessment visit (2006-2010) was collected. Details of participant randomisation, sample handling, Olink proteomics assay through the antibody-based Olink Explore 3072 PEA, data processing, and quality control are as detailed previously[12]. Briefly, proteomics was undertaken on samples of a randomly selected subset of 46,595 UKB participants at the baseline visit, of which 46,011

participants had complete measures of the nine proteins of interest with *a priori* implication in structural cardiovascular disease or treatment pathway through genetic studies (aortopathies, cardiomyopathies, congenital heart disease, and heart failure; ACE2, ACTA2, ACTN4, BAG3, BNP, CDKN1A, NOTCH1, NT-proBNP, and TNNI3). Covariates were assessed (age at recruitment, UKB recruitment centre, genetically determined sex, and proteomics batch). The protein levels are provided in normalized protein eXpression (NPX); Olink's arbitrary unit in $\log_2$ scale.

Genetic analyses

Genotyping array data and exome sequencing data were available for over 450,000 participants. Genotype calling and exome sequencing were performed and imputed as described previously[11,17]. For genome-wide association studies (GWAS), the imputed UKB genotyping data was used, where a minor allele frequency of >0.001 in autosomes was included. Individuals with more than 5% missing genotypes and SNPs with more than 5% missingness were excluded. Participant sex discrepancies, heterozygosity, and relatedness were handled by keeping only genetically European individuals and participants included in the UKB principal components analysis[11]. SNPs deviating from Hardy-Weinberg equilibrium ($1x10^{-8}$) and those with an imputation INFO score of <0.4 were excluded. Individuals with proteomics data were extracted and GWAS was undertaken using GCTA software (version 64)[18]. A sparse genetic relationship matrix (GRM) was created and FastGWA was undertaken with a mixed linear model, adjusting for the genotyping array batch. Genes of independent loci were prioritized through LocusZoom and eQTLs from GTEx (v8). Phenotype associations through GWAS catalog and PheWEB, were assessed. Heritability was estimated by creating a genetic

relationship matrix in GCTA and using a restricted maximum likelihood analysis (REML) to estimate the variance explained by the SNPs that were used to estimate the GRM. Mendelian randomization was undertaken using GWAS summary statistics from published literature[4,5,19,20] with NT-proBNP, BNP, and ACE2, GWAS results using the R package TwoSampleMR. Exposure variants were included if GWAS significant ($P<5x10^{-8}$). Tests of pleiotropy, Steiger directionality, and heterogeneity were assessed.

Rare variant association studies (RVAS) were undertaken using Regenie software on the DNA Nexus Research Analysis Platform[21]. The genotyping data for step 1 of Regenie included SNPs in autosomes with a minor allele frequency <0.01, missingness of >0.01, a minor allele count of <20, deviations from Hardy-Weinberg equilibrium ($5x10^{-15}$), and individuals with greater than 10% missingness, were excluded from the analysis. Interchromosome, SNPs in linkage disequilibrium (indep-pairwise 1000 100 0.9), and areas of low complexity, were excluded for step 1. Exome sequencing data for step 2 was quality controlled for variants in the autosomes with missingness less than 10%, variants where less than 90% of all genotypes for that variant had a read depth less than 10, deviations from Hardy-Weinberg equilibrium ($1x10^{-15}$), and individuals with more than 10% missingness. Step 2 of Regenie was run over different allele frequencies (singletons, 0.01, 0.001) for 6 overlapping, protein-altering variant, custom masks (LoF only; missense only (flagged by >1 of 5 deleterious software); missense only (all); missense only (flagged by all of 5 deleterious software); protein-altering variants (LoF and missense flagged by >1 of 5 deleterious software); protein-altering variants (all Lof and missense)), where the minimum minor allele count was at least 3. Bonferroni significance for 18,117 included genes was $P<2.76x10^{-6}$.

Cardiomyopathy-associated rare variants were identified as previously published[5,22,23] for HCM and DCM. Individuals were classified as genotype negative (SARC-NEG) if they had no rare protein-altering genetic variation (minor allele frequency <0.001 in the UKB and the Genome Aggregation Database) in any genes that may cause or mimic HCM or DCM. These genes represented an inclusive list of genes with definitive or strong evidence of an association with cardiomyopathy, moderate evidence, and genes associated with syndromic phenotypes[1,3,10]. This SARC-NEG group was compared with individuals with disease-associated rare variants in genes with strong or definitive evidence for HCM (*MYBPC3, MYH7, MYL2, MYL3, TNNI3, TNNT2, TPM1,* and *ACTC1*) and DCM (*BAG3, DES, DSP, FLNC, LMNA, MYH7, PLN, RBM20, SCN5A, TNNC1, TNNT2,* and *TTN*). Analysis was restricted to robustly disease-associated variant classes for each gene[3,10] and to variants sufficiently rare to cause penetrant disease (filtering allele frequency <0.00004 for HCM and 0.000084 for DCM[24]). Variants were classified as pathogenic/likely pathogenic (SARC-P/LP) if reported as P/LP for cardiomyopathy in ClinVar and confirmed by manual review.

Cardiac MRI and ECG analyses

A sub-study recalled participants for imaging, including CMR[25], and ECG. For CMR, volumetric traits were measured using quality-controlled deep learning algorithms[14]. Deep neural networks were used for short-axis cine segmentation via a fully convolutional network to label pixels containing myocardium. The performance of image annotation using this algorithm is equivalent to a consensus of expert human readers and achieves subpixel accuracy for cardiac segmentation[14]. 5,324 participants with proteomics had imaging data available.

The ECGs were performed according to a defined protocol and analysed using proprietary software (GE CardioSoft, Boston, MA). Data from the first imaging visit (instance 2, n=42,386) was labeled using a previously trained convolutional neural network designed to identify six diagnoses from the ECG[13]: sinus bradycardia, sinus tachycardia, left bundle branch block, right bundle branch block, 1st degree AV block, and atrial fibrillation. Automated diagnoses had F1 scores above 80% and specificity over 99%[13]. The ECGs were preprocessed with a bandpass filter 0.5 to 100hz, a notch filter at 60hz, and re-sampling to 400hz. Zero padding resulted in a signal with 4,096 samples for each lead for a 10s recording, which was used as input to the neural network model. The binary outputs (presence or absence of each diagnosis) were used for subsequent analyses. 4,831 participants with proteomics had ECG data available.

For analyses comparing plasma proteomics sampled at recruitment to records from the subsequent imaging appointment, the analyses were duplicated to include an adjustment for the difference in time.

Statistical analyses

The analyses were undertaken using R (v4.1.2) and the UKB research analysis platform. The results are expressed as mean and standard deviation (SD). The Student's t-test was used to assess differences in means for quantitative traits and Fisher's exact test for counts. Pearson's correlation coefficient described relationships. Effect sizes are presented as standardised beta coefficients. The proteomic measures were adjusted using multiple linear regression for age at recruitment, age^2, UKB recruitment centre, genetically determined sex, age x sex interaction, proteomics batch, and genetically determined European ancestry, and the resulting standardized residuals (mean=0,

SD=1) were normalized by an inverse rank normalization. Cardiovascular-associated medication intake was self-reported and curated for participants who reported taking medications of interest (**Table S6**). Phenome-wide association studies (PheWAS) were undertaken using the PheWAS R package with clinical outcomes and coded phenotypes converted to 1,840 categorical PheCodes. P-values were deemed significant with Bonferroni adjustment for the number of PheCodes measured.

Cox proportional hazards regression models were assessed with the full cohort of participants for NT-proBNP, BNP, and ACE2, and created using the first reported UKB data (summarising the first date of a report from all UKB data (Hospital episode statistics, primary care, self-reported, etc.)) for heart failure, cardiomyopathy, atrial fibrillation, hypertension, diabetes, and myocardial infarction, as identified through PheWAS analysis, using the survival and survminer R packages by age to death, diagnosis, or last date of follow up report. Participants diagnosed before recruitment were excluded. Participants who died without a diagnosis were also excluded as a sensitivity analysis **(Figure S15**).

Results

Circulating biomarkers of age, sex, alcohol intake, smoking status, and ancestry

46,011 participants had measures of the proteins assessed (participant characteristics: **Table S8**). The correlation between the levels of NT-proBNP and BNP was R=0.67 (**Figure S9**). Levels of ACTA2 had the strongest relationships with the other circulating protein levels (e.g., ACTN4 (R=0.30), NT-proBNP (R=0.30), BNP (R=0.21), BAG3 (R=0.21); **Figure S9**).

A positive relationship was identified with age at recruitment for measures of ACTA2 (R=0.42), NT-proBNP (R=0.34), BNP (R=0.23), and ACE2 (R=0.16; **Figure S1**). ACE2, CDKN1A, and TNNI3, were increased in male compared to female participants (β=0.47, $P=1.0x10^{-16}$; β=0.20, $P=6.61x10^{-68}$; β=0.21, $P=3.97x10^{-127}$; respectively; **Figure S2**). NOTCH1, NT-proBNP, and BNP were increased in female participants (β=0.06, $P=8.61x10^{-258}$; β=0.42, $P=1.62x10^{-278}$; β=0.25, $P=3.50x10^{-69}$; respectively; **Figure S2**).

A positive relationship was observed between alcohol intake and smoking status (smoker at recruitment compared to never smoked), and measures of ACE2 (R=0.14 and β=0.22, $P=1.70x10^{-44}$; respectively) and NOTCH1 (R=-0.15 and β=-0.16, $P=6.11x10^{-23}$). BAG3, CDKN1A, and TNNI3 also had relationships with smoking status (β=-0.08, $P=7.94x10^{-8}$; β=0.10, $P= 2.38x10^{-10}$; β=-0.05, P=0.0008; respectively).

Proteomic variability was observed with ancestry. Participants of self-reported African or Caribbean ancestry (n=584 and n=434, respectively) had increased average ACE2 and decreased ACTA2, ACTN4, BAG3, NT-proBNP, and BNP, compared to British ancestry which dominates the UKB cohort (n=40,228, 87% British; **Figure 2**). Participants of

Chinese ancestry (n=131) had decreased average ACTA2, BAG3, CDKN1A, and NT-proBNP, and participants of Indian ancestry (n=495) had increased average BAG3, CDKN1A, and NOTCH1, compared to British ancestry (**Figure 2**).

Medication use

Medication use was reported at recruitment when the blood was sampled for proteomics (n=46,011 participants) and 8.4 years later (range 3.8-12.7 years) at the imaging appointment (n=5,324). 39%-75% of the individuals with reported cardiovascular medications at recruitment also reported the medication at the imaging appointment (39% on antiplatelets, 57% on beta blockers, 60% on angiotensin receptor blockers (ARBs), 65% on ACE inhibitors, 69% on calcium channel blockers, and 75% on anticoagulants).

Most of the protein levels (measured from blood samples collected at recruitment) had significant associations with medication use reported at recruitment (**Table S6**). To understand whether the protein levels may herald a disease that will require future treatment, protein measures at recruitment were assessed for an association with medication reported at the imaging visit. Participants with increased ACE2, NT-proBNP, and BNP levels at recruitment positively associated with beta-blocker use at the imaging visit 8 years later. ACE2 was also associated with ACE inhibitor (β=0.27, P=7.90x10^{-11}) and ARB use (β=0.22, P=8.44x10^{-4}) and participants with increased NT-proBNP or BNP at recruitment were more likely to report anticoagulant use (β=1.02, P=5.44x10^{-6}; β=0.91, P=6.07x10^{-6}; respectively) at the imaging visit (**Table S6**). The results were similar when the proteins were adjusted for the time between recruitment and the imaging visit (**Table S6**).

To further stratify these associations and understand whether the increase in plasma proteins at recruitment was due to medication use at recruitment or predictive of future medication use, we assessed the change in medication use: whether the protein levels at recruitment were significantly altered with medication i) intake reported only at recruitment, ii) uptake by the imaging visit, or iii) longer-term use and reported during both visits, compared to participants without reported medication (**Figure 3**). ACE2 levels predicted the future uptake of beta-blockers and ACE inhibitors, suggesting that ACE2 levels herald prescription and associated diseases. NT-proBNP and BNP levels predicted long-term use and future uptake of anticoagulants (**Figure 3**). The increase in NT-proBNP and BNP levels observed for beta-blockers may be in part due to the use of the respective medications and/or overt disease at recruitment, as they predicted both current and future use. Adjustment for time since imaging had little effect on the results (**Figure S13**).

Association with clinical features and diagnoses

Through phenome-wide association studies (PheWAS), the most significant associations with each protein were with cardiac dysrhythmias (NT-proBNP), atrial fibrillation and flutter (BNP), chronic renal failure (ACTA2, BAG3, NOTCH1, ACTN4), type-2 diabetes (ACE2, CDKN1A), and congestive heart failure (TNNI3; **Figure 4, Table S1, Figures S16-S30**). Renal diseases associated with most of the protein measures analysed (**Figure S25, Table S1**). ACE2, ACTA2, NT-proBNP, and BNP levels associated with a range of respiratory conditions (**Figure S17, Table S1**). ACE2 levels associated with type-2 diabetes, epilepsy, tobacco use disorder, alcohol-related disorders, and liver diseases (**Figures S16-S30, Table S1**).

A curated analysis was used to assess associations between the nine circulating proteins and specific diagnoses of cardiomyopathies, muscular dystrophy, heart failure, scoliosis, respiratory failure, coronary disease, cardiac arrhythmia (including atrial fibrillation and flutter), stroke, hypertension, valve disease, hypercholesterolemia, and diabetes (**Figure S14, Table S2**). ACE2, ACTA2, TNNI3, NT-proBNP, and BNP levels, had the most associations with curated traits. Hypertension and heart failure correlated with all proteins, except NOTCH1 and CDKN1A, respectively.

Cardiac ECG-AI diagnoses and MRI parameters

Participants with increased NT-proBNP at recruitment were diagnosed more frequently with sinus bradycardia (β=0.25, P=2.82x10^{-7}) and atrial fibrillation (β=0.92, P=7.24x10^{-9}) on future ECG-AI analysis. The associations remained significant when individuals with reported beta-blocker use were removed from the analysis (sinus bradycardia, β=0.25, P=9.44x10^{-7}; atrial fibrillation, β=0.83, P=0.0001). The association with atrial fibrillation was also observed for BNP (β=0.81, P=2.27x10^{-6}). Adjustment for the time between recruitment and imaging had little effect (NT-proBNP and sinus bradycardia, β=0.27, P=2.64x10^{-7}; NT-proBNP and atrial fibrillation, β=1.17, P=2.29x10^{-10}; BNP and atrial fibrillation, β=0.96, P=1.15x10^{-7}).

Analyses of CMR-derived traits identified relationships between NT-proBNP and BNP and increased left atrial volume (R=0.14-0.17) and decreased atrial ejection fraction (R=-0.14--0.19). NT-proBNP also correlated with decreased ventricular wall thickness (R=-0.21) and volumes (R=-0.11--0.18) and increased ventricular ejection fraction (R=0.12). The relationships between NT-proBNP and BNP with atrial volumes (R=0.11-

0.21) and with atrial ejection fractions (R=-0.11--0.17) were not affected by adjustment for the time between the recruitment and imaging visit (**Table S7**).

Carriers of HCM and DCM pathogenic variants

Participants carrying a pathogenic/likely pathogenic variant in a HCM-associated gene, had increased levels of ACE2 and NT-proBNP (but not BNP) at recruitment (β=0.47, P=0.0004; β=0.62, P=0.0007; respectively). Participants carrying a pathogenic/likely pathogenic variant in a DCM-associated gene also had increased NT-proBNP at recruitment (β=0.32, P=0.0002). These associations were likely due to overt cardiomyopathy, as the signals became non-significant with the removal of individuals with diagnosed cardiomyopathy.

Genetic association studies

SNP-based heritability of the protein levels, the amount of variation in the protein levels estimated to be due to genetic factors, identified ACE2 levels with the highest estimated heritability (34.5%), followed by NT-proBNP (33.5%), BAG3 (25.7%), NOTCH1 (22.3%), BNP (18.9%), ACTA2 (17.0%), and CDKN1A (14.2%). ACTN4 and TNNI3 levels had very low and non-significant heritability estimates; either more influenced by non-genetic factors or measurement error.

To identify genetic modifiers of the circulating protein levels, we undertook a genome-wide association study (GWAS; **Figure 5, Table S3**) and a rare variant association study (RVAS) for each protein measure. The GWAS and/or RVAS of the levels of ACTA2, BAG3, CDKN1A, NOTCH1, NT-proBNP, and BNP, identified the expected

gene-protein pair (cis-expression quantitative trait loci were identified at the genetic loci of *ACTA2, BAG3, CDKN1A, NOTCH1,* and *NPPB*).

Recurrent modifiers included *GCKR* for BAG3, CDKN1A, and NOTCH1 levels, *APOE* for CDKN1A and NOTCH1 levels, and *SERPINA1* for ACE2 and NOTCH1 levels (**Table S3**). Glucokinase regulator (*GCKR*) is a regulatory protein that inactivates glucokinase in liver and pancreatic islet cells and has been previously associated with hyperlipidemia and diabetes. Apolipoprotein E (*APOE*) is an apolipoprotein involved in lipoprotein metabolism. Serpin family A member 1 (*SERPINA1*) is a serine protease inhibitor associated with alpha 1-antitrypsin deficiency. The gene is linked to chronic obstructive pulmonary disease, emphysema, and chronic liver disease.

Rare variant association studies (RVAS; **Table S4**) found that ACTA2 levels associated with variants in *LMOD1*. Leiomodin 1 (*LMOD1*) has been implicated in smooth muscle dysfunction and thoracic aortic aneurysm and dissection previously and is predicted to interact with ACTA2[26].

ACE2 and its potential role in hypertension and diabetes

Increasing ACE2 levels at recruitment was correlated with systolic blood pressure (R=0.21; **Figure S14**)) and was predictive of an incident hypertension diagnosis (**Figure 7, Figure S15**). Two-sample Mendelian randomisation with a genetic instrument for systolic blood pressure[19] showed evidence for decreased systolic blood pressure to increase circulating ACE2 (**Figure 6, Figure S3-S4, Table S5**). This suggests a protective role for increased ACE2 levels in blood pressure control.

We showed that ACE2 was significantly increased on average in participants of African or Caribbean ancestry and some guidelines suggest ACE inhibitors (targeting the ACE1 protein) are less effective in “Black” individuals (see the **Discussion**). A high level of ACE2 at recruitment in participants of African or Caribbean ancestry was not significantly predictive of a hypertension diagnosis, however, a limitation of this experiment is the small sample size (n=682).

ACE2 levels were associated with the loci of diabetes-associated genes *HNF1A* and *HNF4A* at GWAS and RVAS (**Table S3-S4**). Increasing ACE2 levels at recruitment were predictive of an incident type-2 diabetes diagnosis (**Figure 7, Figure S15**). Two-sample Mendelian randomisation with the results of a GWAS of type-2 diabetes[27], showed evidence for an inverse relationship: ACE2 genetic instrument decreased the risk of type-2 diabetes while type-2 diabetes increases ACE2 levels (**Figure 6, Figure S7**). The causal protective relationship of ACE2 as an exposure for type-2 diabetes can be exemplified through a missense variant (rs1169288) in the gene hepatocyte nuclear factor-1 alpha (*HNF1A*). The variant is associated with an increased risk of type-2 diabetes (GWAS β=0.04) and decreased levels of circulating ACE2 (GWAS β=-0.16).

NT-proBNP and BNP and their role in heart failure

The GWAS of NT-proBNP identified cardiomyopathy- or heart failure-associated genetic loci (e.g., *HSPB7/CLCNKA*, *CDKN1A*, *TTN, FLNC,* and *BAG3*). BNP also associated with variants in the loci of *BAG3* at GWAS. Through RVAS, NT-proBNP and BNP associated with variants in *NPPB* and the region surrounding the natriuretic peptide genomic locus (*NPPA/NPPB*). This suggests that the natriuretic peptide locus is the

major effect locus of NT-proBNP and BNP levels as the loci of the suggested furin and corin convertase enzymes involved in their biosynthesis were not identified.

NT-proBNP and BNP levels were increased in individuals with overt cardiomyopathy (**Figure S14**) and were predictive of incident diagnoses (heart failure, cardiomyopathy, and atrial fibrillation (**Figure 7, Figure S8, Figure S12, Figure S15**)). To understand whether these increased levels are disease-causing or protective in cardiomyopathies, we undertook two-sample Mendelian randomisation with the results of GWAS of HCM[28] and DCM[5]. This analysis suggested that left ventricular hypertrophy increases NT-proBNP and BNP circulating levels (**Figures S5-S6, Table S5**). An example of this effect is the lead SNP (downstream variant rs198379) of the *NPPB* locus associated with increased NT-proBNP and BNP circulating levels, here and previously[29,30]. The variant increases *NPPB* atrial appendage expression on GTEx[9] and increases the risk of heart failure and HCM through case-control GWAS[20,31] (**Figure S5**).

The opposite was observed for the DCM Mendelian randomisation; common genetic variants associated with DCM risk[5] predisposed to decreased NT-proBNP levels (**Figure S5, Table S5**). Using the same *NPPB* SNP as described as an example, the variant increased levels of NT-proBNP and BNP, and decreased the risk of DCM. Opposing genetic relationships between HCM and DCM common variants have been described elsewhere[6,28] and indicate that genetic loci underlying the variability of left ventricular function in the general population may be differentially involved in susceptibility to HCM and DCM.

Atrial fibrillation[32] increased NT-proBNP circulating levels observationally and through Mendelian randomisation (**Figure S10, Table S5**). However, the observed association

between NT-proBNP and myocardial infarction[33] at PheWAS (**Figure 4**) and with incident myocardial infarction risk (**Figure 7, Figure S15**) did not have evidence of causality (**Table S5, Figure S11**).

NT-proBNP had a positive, relationship with a heart failure diagnosis alongside older age, a diagnosis of cardiomyopathy or atrial fibrillation, and carriers of pathogenic cardiomyopathy-associated variants (**Figure 7, Figure S12, Figure S15**). We showed that NT-proBNP and BNP were significantly decreased on average in participants of African or Caribbean ancestry. Increasing NT-proBNP levels at recruitment in an analysis of participants of African or Caribbean ancestry only remained significantly predictive of heart failure (BNP was not significant).

Discussion

We identified relationships between circulating proteins with cardiovascular disease and those that mediate causal relationships between cardiovascular risk factors and disease development. The results presented identified age, ancestry, and sex-specific protein biomarkers of cardiovascular disease risk. Genetic studies identified variants in *GCKR*, *APOE*, and *SERPINA1*, as modifiers for more than one cardiovascular-associated protein.

We show that variants in *BAG3* influence plasma NT-proBNP and BNP levels. We have previously described a particular common missense variant within BAG3 (C151R; rs2234962) that demonstrates BAG3's potential cardioprotective function in GWAS of DCM, HF, and ejection fraction, alongside risk for HCM[6]. The variant is associated with proteins maintaining myofibrillar integrity and causes improved response to proteotoxic stress. BAG3's role in the protection or risk for opposing cardiomyopathies and carcinomas holds promise for devising therapeutic interventions, diagnostics, and tailored treatments.

Six of the nine proteins assessed have been previously identified (Table S19 at reference[34]) as the strongest predictors in large-scale proteomic risk scores of atrial fibrillation (NT-proBNP, BNP), cardiomyopathy (NT-proBNP, TNNI3, BNP), heart failure (NT-proBNP), hypertension (NT-proBNP, ACE2, ACTA2), nonrheumatic mitral valve disorders (NT-proBNP, BNP), pulmonary hypertension (NT-proBNP, ACE2, BNP), stable angina (NT-proBNP), hyperplasia of prostate (NT-proBNP, ACTA2, BNP), kidney disease (NT-proBNP, BNP), infections (NT-proBNP, ACTN4, BNP), and pleural effusion (NT-proBNP).

ACE2 and COVID-19, diabetes, and hypertension

The observation of the upregulation of ACE2 with smoking[35,36] and alcohol[37–39] has been identified in extensive studies of ACE2 in COVID-19 (where the COVID-19 virus was found to engage ACE2 for cellular entry). There is minimal information on the impact of this on cardiovascular disease risk. Through two-sample Mendelian randomisation, we showed that genetic instruments for decreased systolic blood pressure inversely correlated with circulating ACE2 levels. The relationship of ACE2 circulating levels with cardiovascular risk mediators and genetic factors here suggests that it may be an important target for therapeutic intervention. The increased ACE2 levels observed here with ACE inhibitor treatment (targeting the ACE1 protein) suggests that teasing apart the role of the increase in vasodilator ACE2 effects alongside the decrease in vasoconstriction ACE1 effects would aid our understanding of the success of ACE inhibitors, whether their impact is modified by ACE2, and whether the upregulation or stimulation of ACE2-induced vasodilation would be beneficial in clinical practice.

However, ACE inhibitors are thought to have less blood pressure-lowering effects and increased risk of angioedema in “Black” individuals with hypertension and international guidelines preferentially recommend diuretics and calcium channel blockers over ACE inhibitor treatment in these individuals. Concerns have been expressed with the generalisability of the guidelines and the lack of mechanistic understanding[40]. Further exploration is required into the increased average circulating ACE2 protein levels observed here in individuals of self-reported African or Caribbean ancestry.

Angiotensin receptor blockers prevent the action of angiotensin II for high blood pressure regulation, preventing heart failure, and preventing kidney failure in people with diabetes. *ACE2* increased expression in the endocrine pancreas in diabetes is hypothesized to act in a compensatory manner[41] and here we provide evidence of genetic associations that strengthen a potential role for *ACE2* expression and ACE2 protection against diabetes. Conversely, as a biomarker, ACE2 may have potential as a measure of increased cardiovascular disease risk.

NT-proBNP: a predictive biomarker of hypertrophic cardiomyopathy

NT-proBNP is a prohormone with an N-terminal that is cleaved to release brain or b-type natriuretic peptide 32 (BNP). BNP is released by the heart upon myocardial wall stretch; it reduces fluid and sodium retention and causes mild vasodilation, regulating blood pressure. It has been implicated in hypertrophy, fibrosis, angiogenesis, and cardiomyocyte proliferation and viability. BNP and NT-proBNP are circulating biomarkers of heart failure and hypertrophy due to the reactivation of *NPPA* and *NPPB*[42]. Here, NT-proBNP circulating levels are associated with variants in the loci of these atrial-expressed genes.

We identified an opposing relationship between HCM and DCM and the BNPs; while the levels of BNPs are increased with the progression of both cardiomyopathies, causality via Mendelian randomisation suggested that BNPs are part of HCM pathology and progression while the observed increase in DCM is an adaptive response to contractile dysfunction and cardiomyocyte stretch. The roles of BNPs in natriuresis and promoting hypertrophy may protect against DCM-associated systolic dysfunction. In models, endogenous natriuretic peptides have been shown to protect the heart in a mouse

model of DCM and sudden death, *NPPB* knockout rats at 3 months showed hypertrophy without alteration to the ventricles or function, which transitioned at 6 months into DCM, and diabetic cardiomyopathy mice models treated with exogenous BNP prevented the development of DCM, while knockdown of endogenous BNP accelerated DCM[43–45].

ARBs have been trialled as a treatment for early-stage HCM[46] based on data that suggest they abrogated the development of hypertrophy and fibrosis. Initial results of the VANISH trial were encouraging. Other studies are investigating the combination of sacubitril-valsartan[47] (e.g., NCT04164732). Sacubitril is a neprilysin inhibitor that inhibits the degradation of BNPs. Our results question whether this combination may have negative consequences in patients with hypertrophy, given the possible effects of increased BNP in this context. Our results are in keeping with existing data that support a positive effect in patients with heart failure with reduced ejection fraction[48], including those with dilated cardiomyopathy. Further study is required to address this.

We show better predictive capacity for measures of circulating NT-proBNP than BNP with cardiovascular disease. While the measures are highly correlated (R=0.67), NT-proBNP is measured at a higher concentration, has a higher prognostic value[49], has been shown to have sustained elevation for 12 weeks[50], the predictive capacity of the ratio of NT-proBNP:BNP has been explored previously[51], we show that NT-proBNP was more predictive of heart failure, other cardiovascular diseases, and risk factors, and is more influenced by genetic factors (heritability) than BNP, suggesting important differences in measurable levels. NT-proBNP is inactive and has a longer half-life than BNP which likely explains the improved prediction capacity and increased heritability of plasma measures of NT-proBNP identified here.

The relationships identified with sex, ancestry, and specific cardiovascular diseases, may have implications for NT-proBNP's predictive capacity in clinical settings. It could be suggested that the association of NT-proBNP levels with female sex may be due to lower BMI, but adjustment for BMI did not alter this finding. NT-proBNP >125 pg/mL is common in females without classical cardiovascular risk factors as well as older people[52]. The alteration in average NT-proBNP levels with ancestry is thought to convey an altered risk for hypertension[53,54], but as average levels are lower in individuals of Chinese and African ancestry, NT-proBNP may not have similar predictive capacity across ancestries. Analyses of proteomics in more diverse ancestries would aid this assessment.

Altered NT-proBNP has been previously noted in patients with atrial fibrillation[55], which may influence heart failure or cardiac event risk prediction capabilities in atrial fibrillation patients. Enlargement of the left atrium has been shown to increase NT-proBNP and BNP in individuals diagnosed with atrial fibrillation[56]. Enlarged atria are associated with both existing and incident atrial fibrillation and atrial fibrillation and heart failure have shared pathogenesis. The association of NT-proBNP (and BNP) with future uptake of anticoagulant medication is likely due to a diagnosis of atrial fibrillation or flutter.

The increase of NT-proBNP circulating levels in participants with a pathogenic/likely pathogenic cardiomyopathy-associated variant may be indicative of individuals at particular risk of cardiomyopathy, incident heart failure, or atrial fibrillation. The association of NT-proBNP circulating levels with sinus bradycardia has been identified previously and thought to be through increased stroke volume and wall tension[57], and

we show evidence here that it is unlikely to be due to beta-blockers prescribed for hypertension, as previously suggested[58].

Limitations

There are several limitations to this study. The UKB has biases (survivorship, dominated by European ancestry, etc.). Assessing single time point data limits assessments of causality and directionality. While informative, assessments of medication use require longitudinal follow-up, and the medication associations described may be due to the diagnoses they were prescribed for. The proteomics was measured in samples collected at baseline recruitment while the imaging appointment was on average 8 years later and sensitivity analyses adjusting for this have been untaken. Participants with both proteomics and imaging is currently limited.

Conclusions

We describe the relationships between nine plasma proteins with roles in genetic or structural cardiovascular disease or treatment pathways, that may mediate relationships between cardiovascular risk factors and disease development. We discuss the potential for additional avenues of therapeutic intervention with studies of ACE2 in hypertension and diabetes, and BAG3 in cardiomyopathies, and the need to understand the relationships with NT-proBNP for diagnostic purposes in stratified groups of patients. This study provides an improved understanding of the circulating pathways depicting cardiovascular disease dynamics and the influencing modifiers and risk factors.

Sources of funding

K.A.M. is supported by the British Heart Foundation [BHF; FS/IPBSRF/22/27059, RE/18/4/34215]. L.C. is supported by the BHF [RE/18/4/34215]. A.S. is supported by the BHF [FS/CRTF/21/24183] and a National Institute for Health Research (NIHR) Clinical Lectureship. F.S.N. is supported by the BHF [RG/F/22/110078, RE/18/4/34215, RE/24/130023]. B.H. is funded by the BHF [FS/ICRF/21/26019] and Rosetrees Trust. J.S.W. is supported by the Sir Jules Thorn Charitable Trust [21JTA], the Medical Research Council (UK) [MC_UP_1605/13], and the BHF [RG/19/6/34387, RE/18/4/34215]. D.P.O is supported by the Medical Research Council [MC_UP_1605/13], and the BHF [RG/19/6/34387, RE/18/4/34215, CH/P/23/80008]. All authors were supported by the National Institute for Health Research (NIHR) Imperial College Biomedical Research Centre. The views expressed in this work are those of the authors and not necessarily those of the funders. For open access, the authors have applied a CC BY public copyright license to any Author Accepted Manuscript version arising from this submission. The graphical abstract was created with Biorender.com.

Disclosures

K.A.M. has consulted for Checkpoint Capital LP. D.P.O'R has consulted for Bayer AG and Bristol Myers-Squibb. J.S.W. has consulted for MyoKardia, Inc., Pfizer, Foresite Labs, Health Lumen, and Tenaya Therapeutics, and has received research support from Bristol Myers-Squibb. None of these activities are directly related to the work presented here. All other authors have nothing to disclose.

Data availability

All data included in this study is available through the UK Biobank (https://biobank.ndph.ox.ac.uk/showcase/) and GWAS results will be made available through the GWAS catalog (https://www.ebi.ac.uk/gwas/).

CrediT statement

Conceptualization and guarantor: K.A.M.; Formal analysis: K.A.M.; Resources: L.C., A.S., F.S.N., J.S.W., D.P.O.; Writing – original draft: K.A.M.; Writing – review & editing: (all authors).

Figures

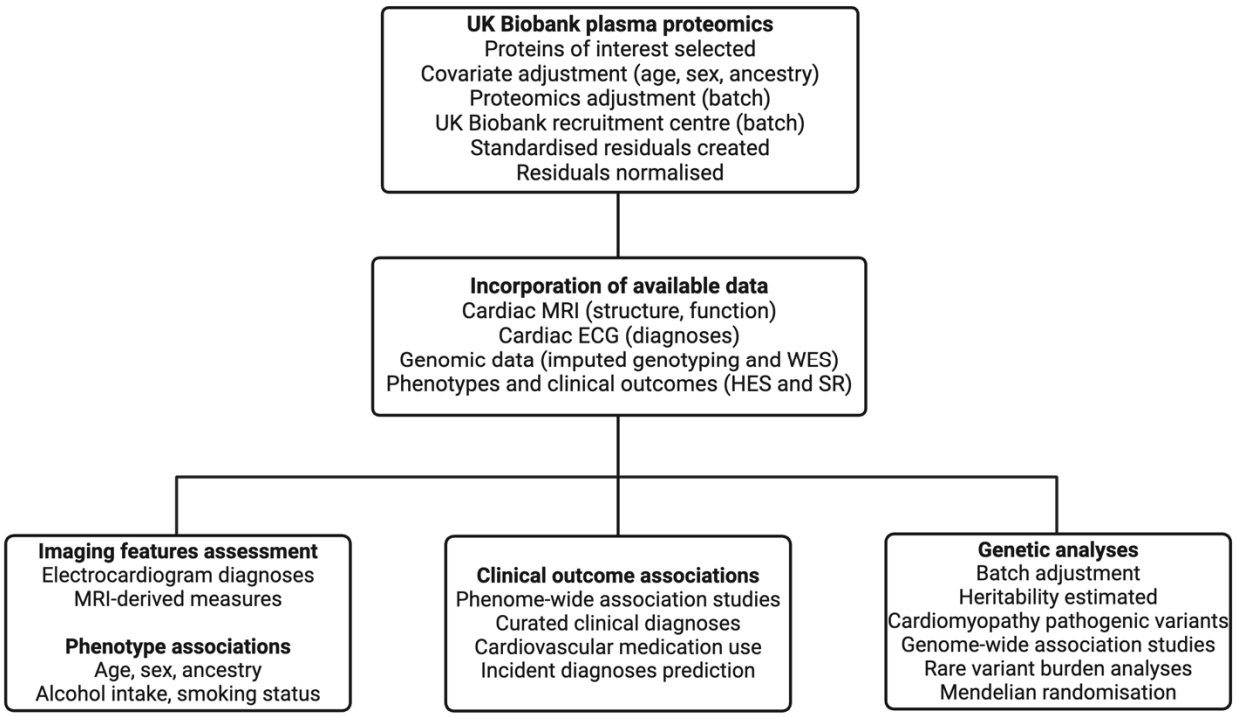


Figure 1 Study flow chart and protein summary. A summary of the main analysis steps and data available for the analysis of nine plasma proteins and the genetic and outcome associations. MRI, magnetic resonance imaging; CMR, cardiac MRI; WES, whole exome sequencing; HES, hospital episode statistics; SR, self-reported.

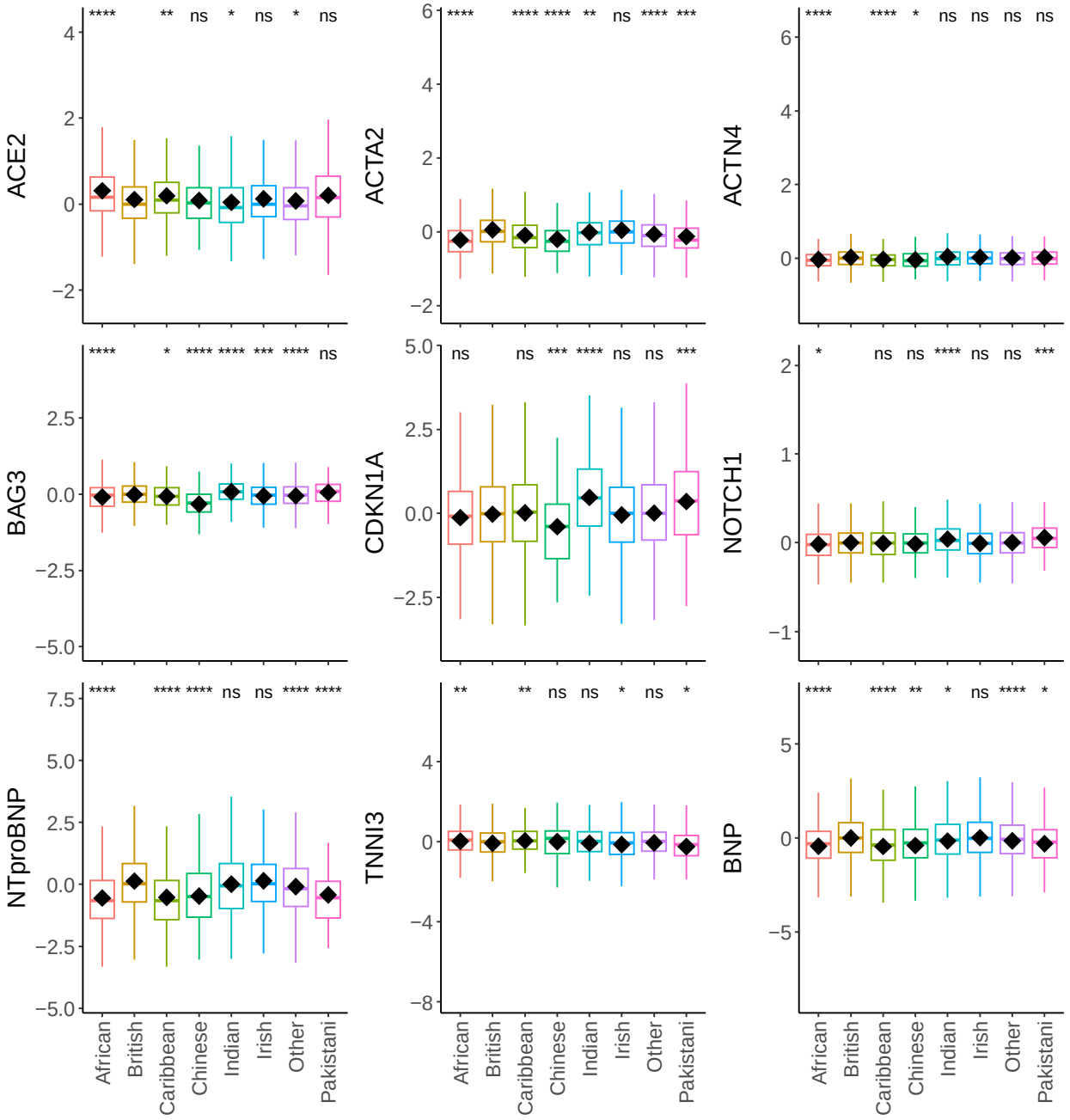


Figure 2 The relationships with proteins and ancestry. The plots depict the significant differences in protein levels across self-reported ancestries. Participants of self-reported African or Caribbean ancestry had increased average ACE2 and decreased ACTA2, ACTN4, BAG3, NT-proBNP, and BNP, compared to British ancestry. Participants with Chinese ancestry had decreased average ACTA2, BAG3, CDKN1A,

and NT-proBNP, and participants with Indian ancestry had increased average BAG3, CDKN1A, and NOTCH1, compared to British ancestry. The significance of differences in means as derived by Student's t-test are denoted as stars compared to British ancestry. The y-axis units are Olink's arbitrary unit in $\log_2$ scale. The sample sizes were as follows (African n=584, British n=40228, Caribbean n=434, Chinese n=131, Indian n=495, Irish n=1200, Other n=2736, Pakistani n=143).

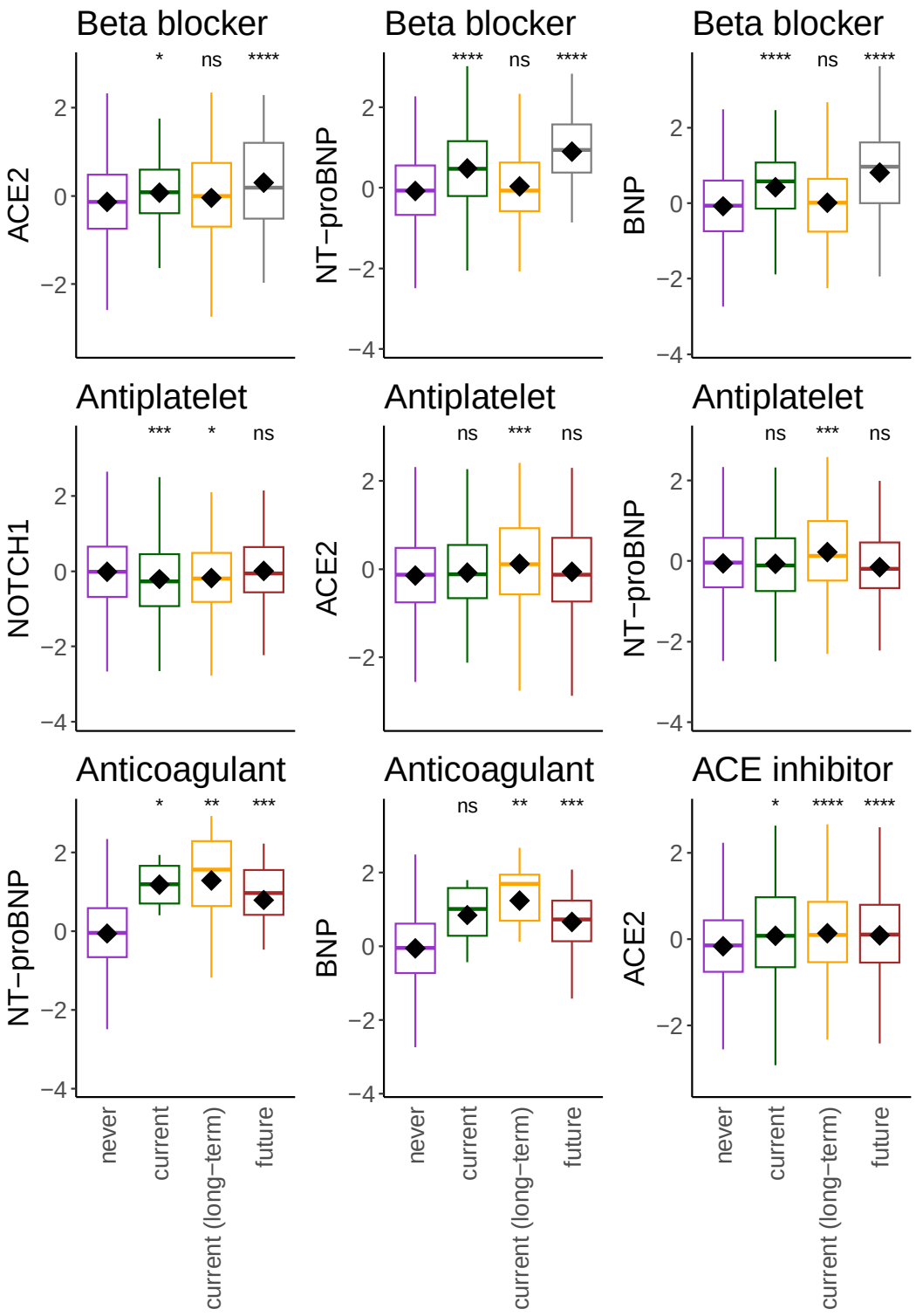


Figure 3 The relationships with proteins and medication intake. The plots depict the proteins measured at recruitment that were significantly increased with medication

intake reported only at recruitment (current), or the imaging visit on average 8 years later (future), or reported during both visits (current (long-term)). The significance of differences in means as derived by Student's t-test are denoted as stars compared to no report of the medication (never). The y-axis units are standardized residuals after adjustment for covariates. The data only includes those with proteomics who attended the imaging visit (n=5,324).

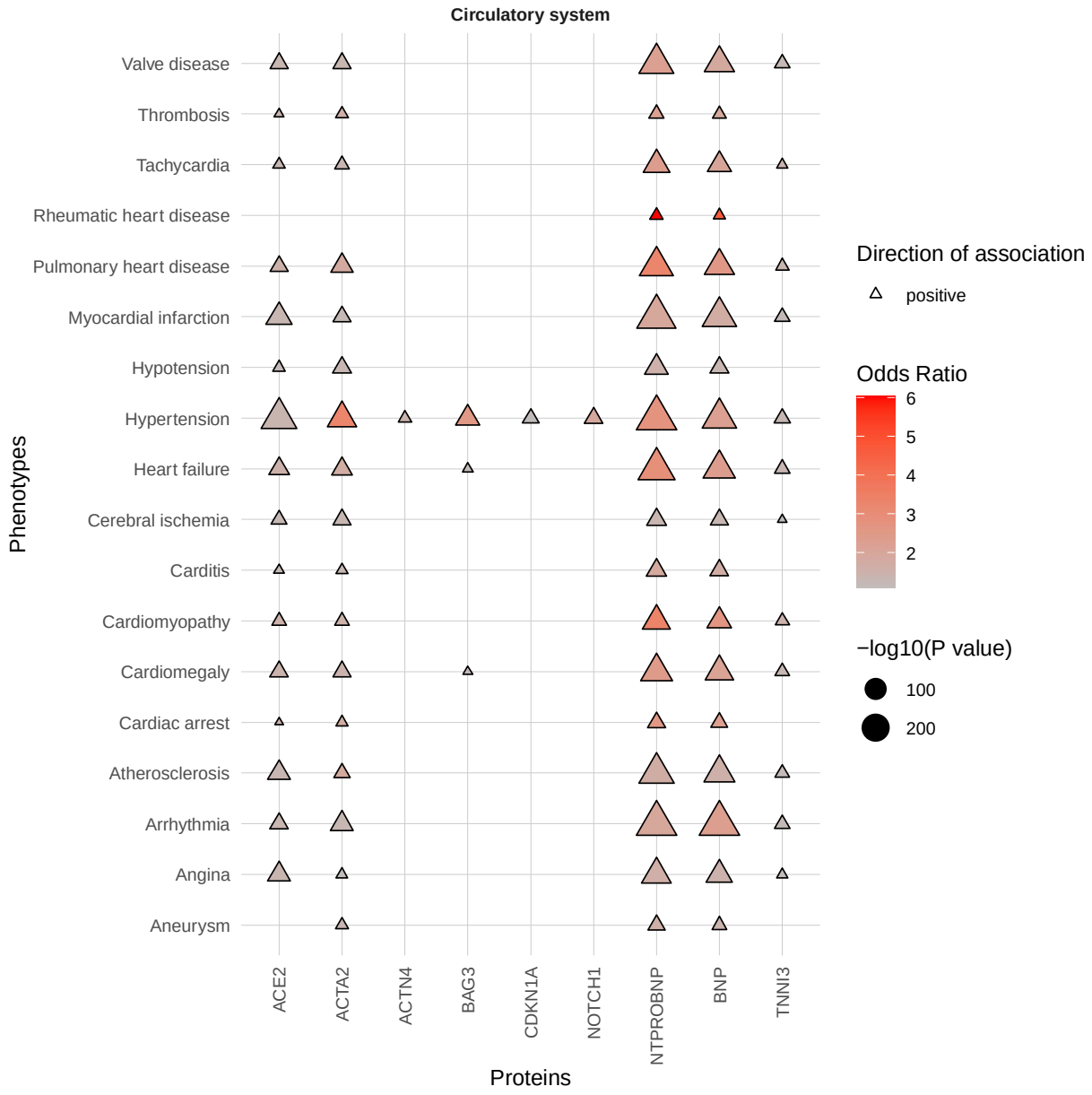


Figure 4 Phenome-wide association study results of the plasma protein levels with selected circulatory disorders.

Phenotypes as phecodes are described on the y-axis and the protein traits on the x-axis. Each point denotes a significant PheWAS association with a Bonferroni correction for the number of analyzed phecodes. The shape and colour denote the direction of effect

and odds ratio. Only the most significant associations with selected, non-redundant phenotypes of the circulatory disorder category are presented for clarity. See **Table S1** for the full PheWAS results.

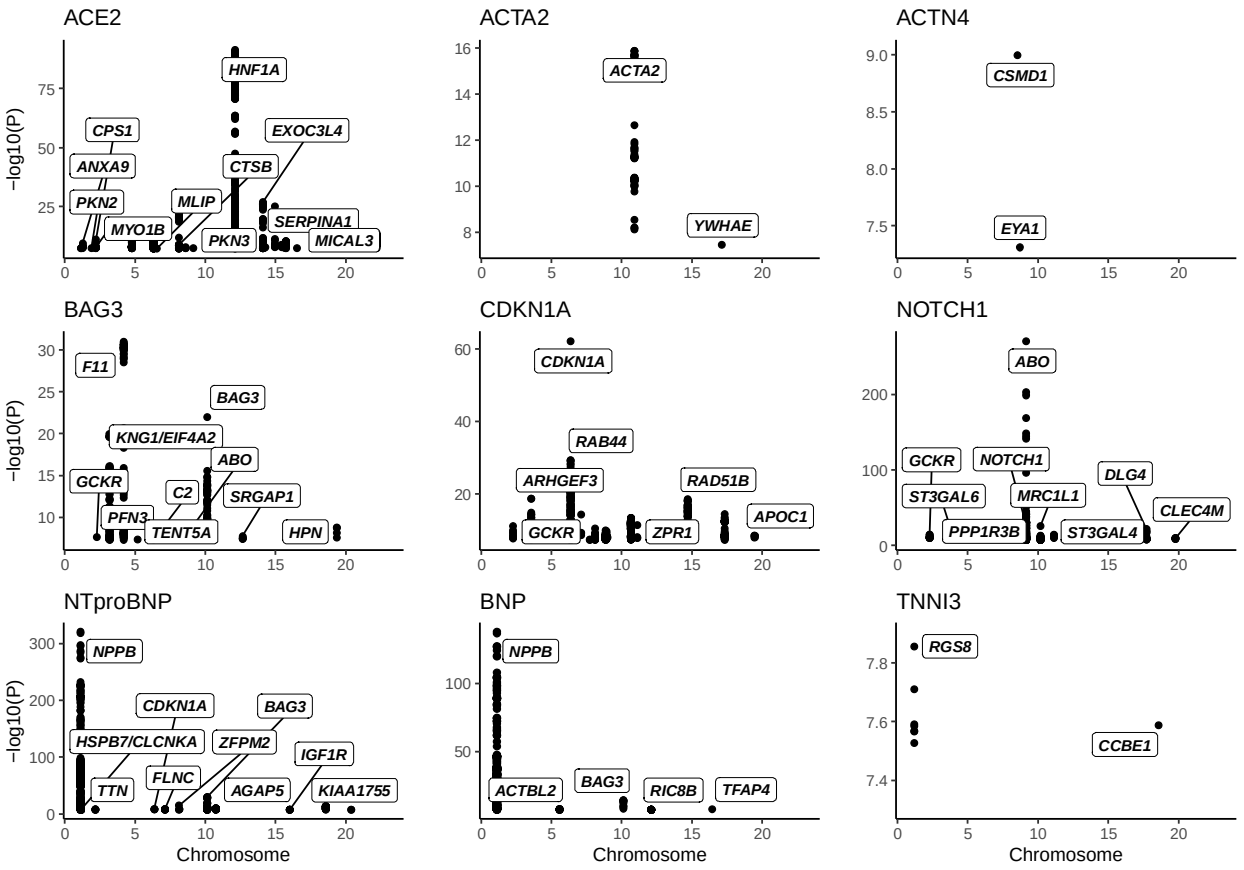


Figure 5 Significant genome-wide association study results. The Manhattan plots present the GWAS significant SNPs for the nine protein levels. The prioritized gene is noted for the significant loci identified. A subset of gene labels for ACE2, CDKN1A, and NT-proBNP, have been selected to allow for presentation. The y-axis is cut at a minimum of 5e-08 (7.3). Please see **Table S3** for the full GWAS results.

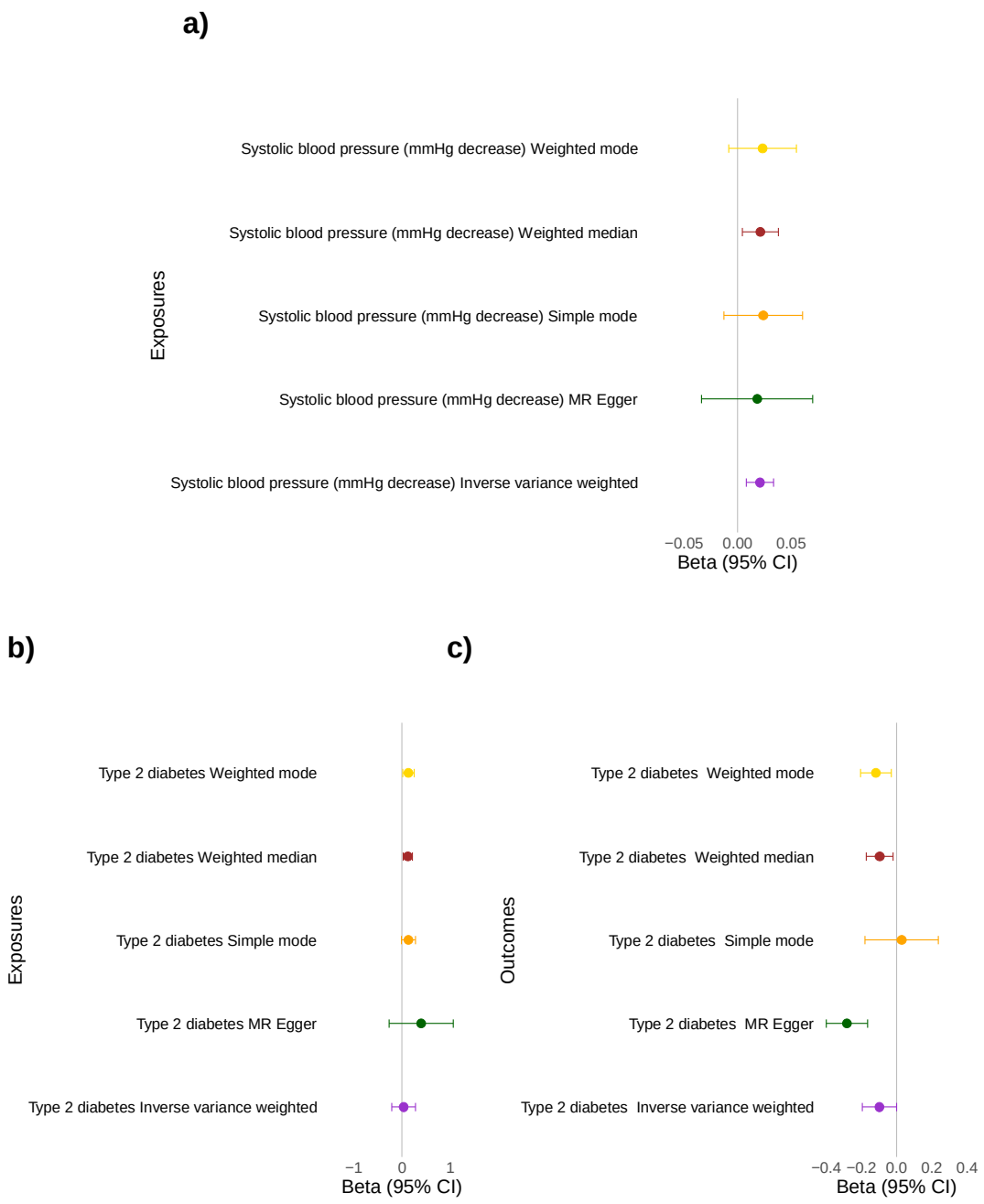
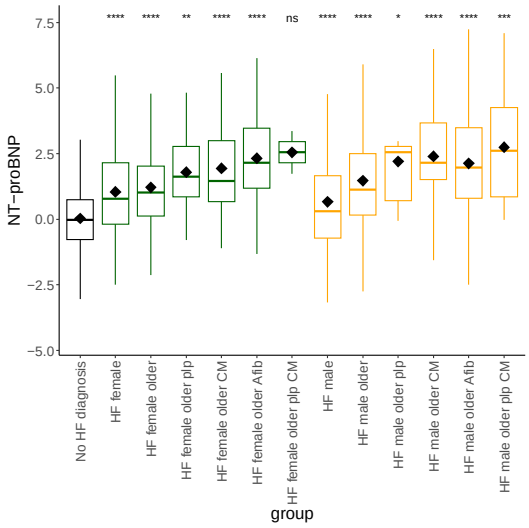


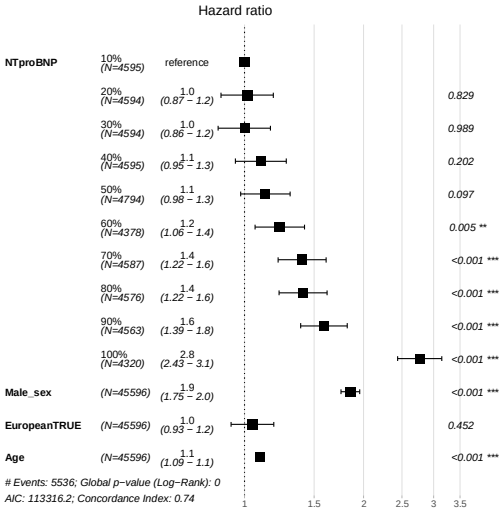
Figure 6 Evidence of a causal relationship between systolic blood pressure, type-2 diabetes, and ACE2. Increased circulating ACE2 can decrease blood pressure through the creation of vasodilators and is causally associated with type-2 diabetes. A genetic predisposition for decreased systolic blood pressure is associated with increased ACE2. **a)** Mendelian randomization genetic determination model of systolic

blood pressure (mmHg decrease) genetic instruments as exposures for ACE2 outcome. Two-sample Mendelian randomization was undertaken with ACE2 (using the GWAS results) and decreased systolic blood pressure (from GWAS summary statistics of published data). **b)** Mendelian randomization genetic determination model of type-2 diabetes genetic instruments as exposures for ACE2 outcome. Two-sample Mendelian randomization was undertaken with ACE2 (using the GWAS results) and type-2 diabetes (from GWAS summary statistics of published data). **c)** Mendelian randomization genetic determination model of ACE2 genetic instrument as an exposure for type-2 diabetes. See **Table S5** and **Figures S3-S4** for further details.

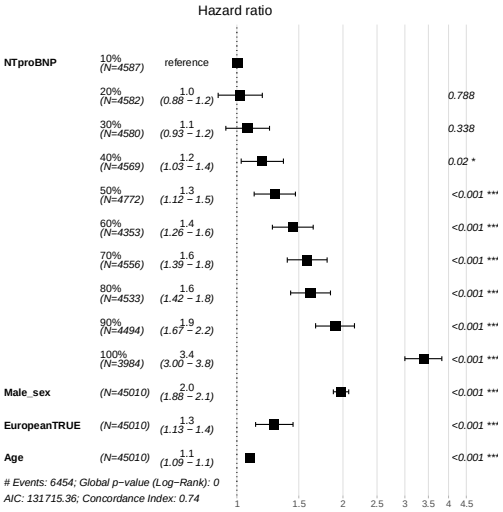
a) NT-proBNP and HF



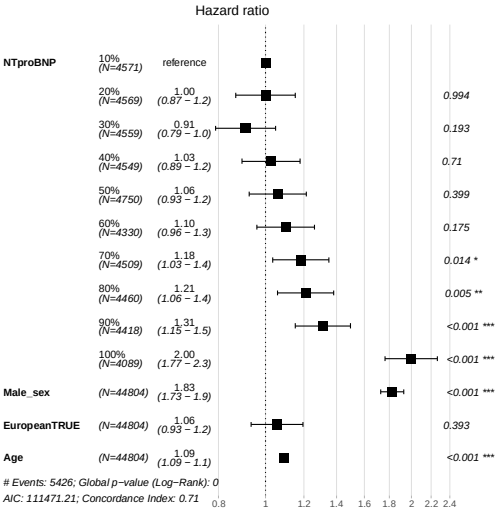
b) NT-proBNP and HF



c) NT-proBNP and Afib



d) NT-proBNP and MI



e) ACE2 and hypertension **f) ACE2 and type-2 diabetes**

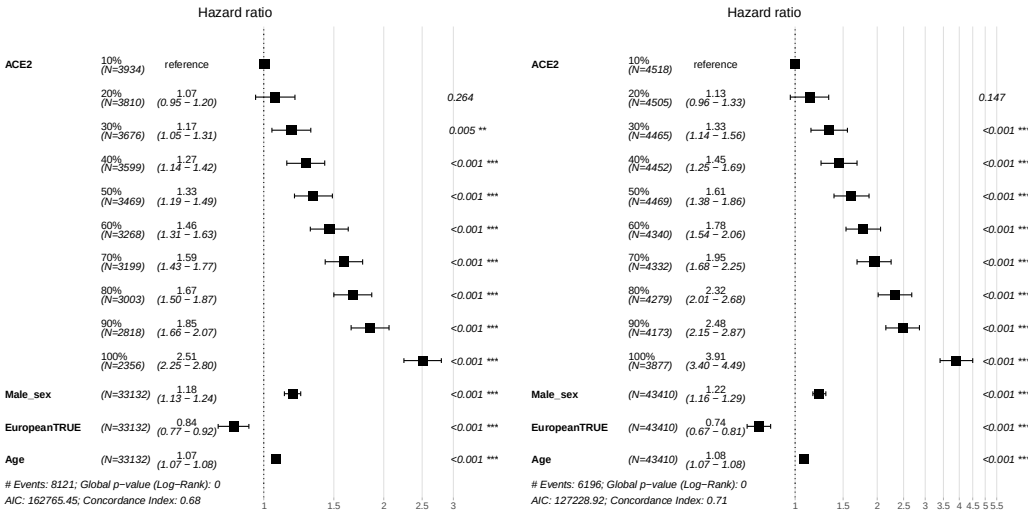


Figure 7 Increasing levels of NT-proBNP are observed in participants diagnosed with incident heart failure, atrial fibrillation, and myocardial infarction. Increasing levels of ACE2 are observed in participants diagnosed with incident hypertension and type-2 diabetes. a) The figure shows the sequential increase in mean NT-proBNP with overt diseases and other modifiers influencing the protein's levels. This includes a heart failure (HF) diagnosis alongside sex, age at recruitment, a diagnosis of cardiomyopathy or atrial fibrillation, and carriers of pathogenic cardiomyopathy-associated variants. The Student's t-test compared "No HF diagnosis" as the reference group. The groups contained the following sample sizes, respectively: (No HF diagnosis) 44050, (female groups:) 103, 286, 12, 33, 234, 2, (male groups:) 202, 496, 7, 42, 470, 14. NT-proBNP units are Olink's arbitrary unit in $\log_2$ scale. HF, heart failure; plp, P/LP variant carrier; Afib, atrial fibrillation; CM, cardiomyopathy diagnosis. The forest plots (**b)-f)**) of Cox proportional hazards regression models were created assessing death or diagnosis from recruitment with those diagnosed before recruitment

excluded. Sex (increasing risk is male), European ancestry (increasing risk is European), and age at recruitment (incremental risk per year lived), were added to this multivariable analysis for comparison. Forest plots are presented for deciles of NT-proBNP levels with incident **b)** heart failure (HF), **c)** atrial fibrillation (Afib), **d)** myocardial infarction (MI), from recruitment, and ACE2 levels by decile with incident **e)** hypertension and **f)** type-2 diabetes, from recruitment.

References

1. Josephs KS, Roberts AM, Theotokis P, Walsh R, Ostrowski PJ, Edwards M, Fleming A, Thaxton C, Roberts JD, Care M, Zareba W, Adler A, Sturm AC, Tadros R, Novelli V, Owens E, Bronicki L, Jarinova O, Callewaert B, Peters S, Lumbers T, Jordan E, Asatryan B, Krishnan N, Hershberger RE, Chahal CAA, Landstrom AP, James C, McNally EM, Judge DP, Tintelen P van, Wilde A, Gollob M, Ingles J, Ware JS. Beyond gene-disease validity: capturing structured data on inheritance, allelic-requirement, disease-relevant variant classes, and disease mechanism for inherited cardiac conditions. *Genome Med*. 2023;15.

2. Rehm HL, Berg JS, Brooks LD, Bustamante CD, Evans JP, Landrum MJ, Ledbetter DH, Maglott DR, Martin CL, Nussbaum RL, Plon SE, Ramos EM, Sherry ST, Watson MS. ClinGen — The Clinical Genome Resource. *New England Journal of Medicine*. 2015;372.

3. Jordan E, Peterson L, Ai T, Asatryan B, Bronicki L, Brown E, Celeghin R, Edwards M, Fan J, Ingles J, James CA, Jarinova O, Johnson R, Judge DP, Lahrouchi N, Lekanne Deprez RH, Lumbers RT, Mazzarotto F, Medeiros Domingo A, Miller RL, Morales A, Murray B, Peters S, Pilichou K, Protonotarios A, Semsarian C, Shah P, Syrris P, Thaxton C, van Tintelen JP, Walsh R, Wang J, Ware J, Hershberger RE. Evidence-Based Assessment of Genes in Dilated Cardiomyopathy. *Circulation*. 2021;144:7–19.

4. Tadros R, Zheng SL, Grace C, Jordà P, Francis C, Jurgens SJ, Thomson KL, Harper AR, Ormondroyd E, West DM, Xu X, Theotokis PI, Buchan RJ, McGurk KA, Mazzarotto F, Boschi B, Pelo E, Lee M, Noseda M, Varnava A, Vermeer AMC, Walsh R, Amin AS, Slegtenhorst MA van, Roslin N, Strug LJ, Salvi E, Lanzani C, Marvao A de, Collaborators HI, Roberts JD, Tremblay-Gravel M, Giraldeau G, Cadrin-Tourigny J, L'Allier PL, Garceau P, Talajic M, Pinto YM, Rakowski H, Pantazis A, Baksi J, Halliday BP, Prasad SK, Barton PJR, O'Regan DP, Cook SA, Boer RA de, Christiaans I, Michels M, Kramer CM, Ho CY, Neubauer S, Investigators H, Matthews PM, Wilde AA, Tardif J-C, Olivotto I, Adler A, Goel A, Ware JS, Bezzina CR, Watkins H. Large scale genome-wide association analyses identify novel genetic loci and mechanisms in hypertrophic cardiomyopathy. *medRxiv*. 2023;2023.01.28.23285147.

5. Zheng SL, Henry A, Cannie D, Lee M, Miller D, McGurk KA, Bond I, Xu X, Issa H, Francis C, Marvao A De, Theotokis PI, Buchan RJ, Speed D, Abner E, Adams L, Aragam KG, Ärnlöv J, Raja AA, Backman JD, Baksi J, Barton PJR, Biddinger KJ, Boersma E, Brandimarto J, Brunak S, Brundgaard H, Carey DJ, Charron P, Cook JP, Cook SA, Denaxas S, Deleuze J-F, Doney AS, Elliott P, Erikstrup C, Esko T, Farber-Eger EH, Finan C, Garnier S, Ghouse J, Giedraitis V, Guðbjartsson DF, Haggerty CM, Halliday BP, Helgadottir A, Hemingway H, Hillege H, Kardys I, Lind L, Lindgren CM,

Lowery BD, Manisty C, Margulies KB, Moon JC, Mordi IR, Morley MP, Morris AD, Morris AP, Morton L, Noursadeghi M, Ostrowski SR, Owens AT, Palmer CNA, Pantazis A, Pedersen OB V, Prasad SK, Shekhar A, Smelser DT, Srinivasan S, Stefansson K, Sveinbjörnsson G, Syrris P, Tammesoo M-L, Tayal U, Teder-Laving M, Thorgeirsson G, Thorsteinsdottir U, Tragante V, Trégouët D-A, Treibel TA, Ullum H, Valdes AM, Setten J van, Vugt M van, Veluchamy A, Verschuuren WMM, Villard E, Yang Y, COVIDsortium, Consortium DG, Consortium GER, Consortium H, Asselbergs FW, Cappola TP, Dube M-P, Dunn ME, Ellinor PT, et al. Genome-wide association analysis reveals insights into the molecular etiology underlying dilated cardiomyopathy. *medRxiv*. 2023;2023.09.28.23295408.

6. Mukhopadhyay S, Dixit P, Khanom N, Sanghera G, McGurk KA. The Genetic Factors Influencing Cardiomyopathies and Heart Failure across the Allele Frequency Spectrum. *J Cardiovasc Transl Res*. 2024;

7. McGurk KA, Qiao M, Zheng SL, Sau A, Henry A, Ribeiro ALP, Ribeiro AH, Ng FS, Lumbers RT, Bai W, Ware JS, O'Regan DP. Genetic and phenotypic architecture of human myocardial trabeculation. *medRxiv*. 2024;2024.03.26.24304726.

8. Page DJ, Miossec MJ, Williams SG, Monaghan RM, Fotiou E, Cordell HJ, Sutcliffe L, Topf A, Bourgey M, Bourque G, Eveleigh R, Dunwoodie SL, Winlaw DS, Bhattacharya S, Breckpot J, Devriendt K, Gewillig M, Brook JD, Setchfield KJ, Bu'Lock FA, O'Sullivan J, Stuart G, Bezzina CR, Mulder BJM, Postma A V., Bentham JR, Baron M, Bhaskar SS, Black GC, Newman WG, Hentges KE, Lathrop GM, Santibanez-Koref M, Keavney BD. Whole Exome Sequencing Reveals the Major Genetic Contributors to Nonsyndromic Tetralogy of Fallot. *Circ Res*. 2019;124:553–563.

9. Aguet F, Barbeira A, Bonazzola R, Brown A, Castel S, Jo B, Kasela S, Kim-Hellmuth S, Liang Y, Oliva M, Parsana P, Flynn E, Fresard L, Gaamzon E, Hamel A, He Y, Hormozdiari F, Mohammadi P, Muñoz-Aguirre M, Park Y, Saha A, Segrć A, Strober B, Wen X, Wucher V, Das S, Garrido-Martín D, Gay N, Handsaker R, Hoffman P, Kashin S, Kwong A, Li X, MacArthur D, Rouhana J, Stephens M, Todres E, Viñuela A, Wang G, Zou Y, Brown C, Cox N, Dermitzakis E, Engelhardt B, Getz G, Guigo R, Montgomery S, Stranger B, Im HK, Battle A, Ardlie K, Lappalainen T. The GTEx Consortium atlas of genetic regulatory effects across human tissues. *Science (1979)*. 2020;369:1318–1330.

10. Ingles J, Goldstein J, Thaxton C, Caleshu C, Corty EW, Crowley SB, Dougherty K, Harrison SM, McGlaughon J, Milko L V., Morales A, Seifert BA, Strande N, Thomson K, Peter Van Tintelen J, Wallace K, Walsh R, Wells Q, Whiffin N, Witkowski L, Semsarian C, Ware JS, Hershberger RE, Funke B. Evaluating the Clinical Validity of Hypertrophic Cardiomyopathy Genes. *Circ Genom Precis Med*. 2019;12:57–64.

11. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, Motyer A, Vukcevic D, Delaneau O, O'Connell J, Cortes A, Welsh S, Young A, Effingham M, McVean G, Leslie S, Allen N, Donnelly P, Marchini J. The UK Biobank resource with deep phenotyping and genomic data. *Nature*. 2018;562:203–209.

12. Sun BB, Chiou J, Traylor M, Benner C, Hsu YH, Richardson TG, Surendran P, Mahajan A, Robins C, Vasquez-Grinnell SG, Hou L, Kvikstad EM, Burren OS, Davitte J, Ferber KL, Gillies CE, Hedman ÅK, Hu S, Lin T, Mikkilineni R, Pendergrass RK, Pickering C, Prins B, Baird D, Chen CY, Ward LD, Deaton AM, Welsh S, Willis CM, Lehner N, Arnold M, Wörheide MA, Suhre K, Kastenmüller G, Sethi A, Cule M, Raj A, Kang HM, Burkitt-Gray L, Melamud E, Black MH, Fauman EB, Howson JMM, Kang HM, McCarthy MI, Nioi P, Petrovski S, Scott RA, Smith EN, Szalma S, Waterworth DM, Mitnaul LJ, Szustakowski JD, Gibson BW, Miller MR, Whelan CD. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature*. 2023;622:329–338.

13. Ribeiro AH, Ribeiro MH, Paixão GMM, Oliveira DM, Gomes PR, Canazart JA, Ferreira MPS, Andersson CR, Macfarlane PW, Meira Jr. W, Schön TB, Ribeiro ALP. Automatic diagnosis of the 12-lead ECG using a deep neural network. *Nat Commun*. 2020;11:1760.

14. Bai W, Sinclair M, Tarroni G, Oktay O, Rajchl M, Vaillant G, Lee AM, Aung N, Lukaschuk E, Sanghvi MM, Zemrak F, Fung K, Paiva JM, Carapella V, Kim YJ, Suzuki H, Kainz B, Matthews PM, Petersen SE, Piechnik SK, Neubauer S, Glocker B, Rueckert D. Automated cardiovascular magnetic resonance image analysis with fully convolutional networks. *Journal of Cardiovascular Magnetic Resonance*. 2018;20:65.

15. Bai W, Suzuki H, Huang J, Francis C, Wang S, Tarroni G, Guitton F, Aung N, Fung K, Petersen SE, Piechnik SK, Neubauer S, Evangelou E, Dehghan A, O'Regan DP, Wilkins MR, Guo Y, Matthews PM, Rueckert D. A population-based phenome-wide association study of cardiac and aortic structure and function. *Nat Med*. 2020;26:1654–1662.

16. Evangelou E, Suzuki H, Bai W, Pazoki R, Gao H, Matthews PM, Elliott P. Alcohol consumption in the general population is associated with structural changes in multiple organ systems. *Elife*. 2021;10:e65325.

17. Backman JD, Li AH, Marcketta A, Sun D, Mbatchou J, Kessler MD, Benner C, Liu D, Locke AE, Balasubramanian S, Yadav A, Banerjee N, Gillies CE, Damask A, Liu S, Bai X, Hawes A, Maxwell E, Gurski L, Watanabe K, Kosmicki JA, Rajagopal V, Mighty J, Jones M, Mitnaul L, Stahl E, Coppola G, Jorgenson E, Habegger L, Salerno WJ, Shuldiner AR, Lotta LA, Overton JD, Cantor MN, Reid JG, Yancopoulos G, Kang HM, Marchini J, Baras A, Abecasis GR, Ferreira MAR. Exome sequencing and analysis of 454,787 UK Biobank participants. *Nature*. 2021;599:628–634.

18. Yang J, Lee SH, Goddard ME, Visscher PM. GCTA: A tool for genome-wide complex trait analysis. *Am J Hum Genet*. 2011;88:76–82.

19. Evangelou E, Warren HR, Mosen-Ansorena D, Mifsud B, Pazoki R, Gao H, Ntritsos G, Dimou N, Cabrera CP, Karaman I, Ng FL, Evangelou M, Witkowska K, Tzanis E, Hellwege JN, Giri A, Velez Edwards DR, Sun Y V., Cho K, Gaziano JM, Wilson PWF, Tsao PS, Kovesdy CP, Esko T, Mägi R, Milani L, Almgren P, Boutin T, Debette S, Ding J, Giulianini F, Holliday EG, Jackson AU, Li-Gao R, Lin WY, Luan J, Mangino M, Oldmeadow C, Prins BP, Qian Y, Sargurupremraj M, Shah N, Surendran P, Thériault S, Verweij N, Willems SM, Zhao JH, Amouyel P, Connell J, de Mutsert R, Doney ASF, Farrall M, Menni C, Morris AD, Noordam R, Paré G, Poulter NR, Shields DC, Stanton A, Thom S, Abecasis G, Amin N, Arking DE, Ayers KL, Barbieri CM, Batini C, Bis JC, Blake T, Bochud M, Boehnke M, Boerwinkle E, Boomsma DI, Bottinger EP, Braund PS, Brumat M, Campbell A, Campbell H, Chakravarti A, Chambers JC, Chauhan G, Ciullo M, Cocca M, Collins F, Cordell HJ, Davies G, Borst MH de, Geus EJ de, Deary IJ, Deelen J, Del Greco M F, Demirkale CY, Dörr M, Ehret GB, Elosua R, Enroth S, Erzurumluoglu AM, Ferreira T, Frånberg M, et al. Genetic analysis of over 1 million people identifies 535 new loci associated with blood pressure traits. *Nat Genet*. 2018;50:1412–1425.

20. Henry A, Mo X, Finan C, Chaffin MD, Speed D, Issa H, Denaxas S, Ware JS, Zheng SL, Malarstig A, Gratton J, Bond I, Roselli C, Miller D, Chopade S, Schmidt AF, Abner E, Adams L, Andersson C, Aragam KG, Ärnlöv J, Asselin G, Raja AA, Backman JD, Bartz TM, Biddinger KJ, Biggs ML, Bloom HL, Boersma E, Brandimarto J, Brown MR, Brunak S, Bruun MT, Buckbinder L, Bundgaard H, Carey DJ, Chasman DI, Chen X, Cook JP, Czuba T, Denus S de, Dehghan A, Delgado GE, Doney AS, Dörr M, Dowsett J, Dudley SC, Engström G, Erikstrup C, Esko T, Farber-Eger EH, Felix SB, Finer S, Ford I, Ghanbari M, Ghasemi S, Ghouse J, Giedraitis V, Giulianini F, Gottdiener JS, Gross S, Guðbjartsson DF, Gui H, Gutmann R, Hägg S, Haggerty CM, Hedman ÅK, Helgadottir A, Hemingway H, Hillege H, Hyde CL, Jensen BA, Jukema JW, Kardys I, Karra R, Kavousi M, Kizer JR, Kleber ME, Køber L, Koekemoer A, Kuchenbaecker K, Lai Y-P, Lanfear D, Langenberg C, Lin H, Lind L, Lindgren CM, Liu PP, London B, Lowery BD, Luan J, Lubitz SA, Magnusson P, Margulies KB, Marston NA, Martin H, März W, Melander O, et al. Mapping the aetiological foundations of the heart failure spectrum using human genetics. *medRxiv*. 2023;2023.10.01.23296379.

21. Mbatchou J, Barnard L, Backman J, Marcketta A, Kosmicki JA, Ziyatdinov A, Benner C, O'Dushlaine C, Barber M, Boutkov B, Habegger L, Ferreira M, Baras A, Reid J, Abecasis G, Maxwell E, Marchini J. Computationally efficient whole-genome regression for quantitative and binary traits. *Nat Genet*. 2021;53:1097–1103.

22. de Marvao A, McGurk KA, Zheng SL, Thanaj M, Bai W, Duan J, Biffi C, Mazzarotto F, Statton B, Dawes TJW, Savioli N, Halliday BP, Xu X, Buchan RJ, Baksi AJ, Quinlan M, Tokarczuk P, Tayal U, Francis C, Whiffin N, Theotokis PI, Zhang X, Jang M, Berry A, Pantazis A, Barton PJR, Rueckert D, Prasad SK, Walsh R, Ho CY, Cook SA, Ware JS, O'Regan DP. Phenotypic Expression and Outcomes in Individuals With Rare Genetic Variants of Hypertrophic Cardiomyopathy. *J Am Coll Cardiol*. 2021;78:1097–1110.

23. Zheng SL, Jurgens SJ, McGurk KA, Xu X, Grace C, Theotokis PI, Buchan RJ, Francis C, Marvao A de, Curran L, Bai W, Pua CJ, Chiaw TH, Jorda P, Slegtenhorst MA van, Verhagen JMA, Harper AR, Ormondroyd E, Chin CWL, Consortium GER, Collaborators HCMG, Pantazis A, Baksi J, Halliday BP, Matthews P, Pinto YM, Walsh R, Amin AS, Wilde AAM, Cook SA, Prasad SK, Barton PJR, O'Regan DP, Lumbers RT, Goel A, Tadros R, Michels M, Watkins H, Bezzina CR, Ware JS. Evaluation of polygenic score for hypertrophic cardiomyopathy in the general population and across clinical settings. *medRxiv*. 2023;2023.03.14.23286621.

24. Whiffin N, Minikel E, Walsh R, O'Donnell-Luria AH, Karczewski K, Ing AY, Barton PJR, Funke B, Cook SA, Macarthur D, Ware JS. Using high-resolution variant frequencies to empower clinical genome interpretation. *Genetics in Medicine*. 2017;19:1151–1158.

25. Bai W, Suzuki H, Huang J, Francis C, Wang S, Tarroni G, Guitton F, Aung N, Fung K, Petersen SE, Piechnik SK, Neubauer S, Evangelou E, Dehghan A, O'Regan DP, Wilkins MR, Guo Y, Matthews PM, Rueckert D. A population-based phenome-wide association study of cardiac and aortic structure and function. *Nat Med*. 2020;26:1654–1662.

26. Fournier N, Fabre A. Smooth muscle motility disorder phenotypes: A systematic review of cases associated with seven pathogenic genes (ACTG2, MYH11, FLNA, MYLK, RAD21, MYL9 and LMOD1). Intractable Rare Dis Res. 2022;11:113–119.

27. Mahajan A, Wessel J, Willems SM, Zhao W, Robertson NR, Chu AY, Gan W, Kitajima H, Taliun D, Rayner NW, Guo X, Lu Y, Li M, Jensen RA, Hu Y, Huo S, Lohman KK, Zhang W, Cook JP, Prins BP, Flannick J, Grarup N, Trubetskoy VV, Kravic J, Kim YJ, Rybin D V., Yaghootkar H, Müller-Nurasyid M, Meidtner K, Li-Gao R, Varga T V., Marten J, Li J, Smith AV, An P, Ligthart S, Gustafsson S, Malerba G, Demirkan A, Tajes JF, Steinthorsdottir V, Wuttke M, Lecoeur C, Preuss M, Bielak LF, Graff M, Highland HM, Justice AE, Liu DJ, Marouli E, Peloso GM, Warren HR, Afaq S, Afzal S, Ahlqvist E, Bang LB, Bertoni AG, Bombieri C, Bork-Jensen J, Brandslund I, Brody JA, Burtt NP, Canouil M, Chen YDI, Cho YS, Christensen C, Eastwood S V., Eckardt KU, Fischer K, Gambaro G, Giedraitis V, Grove ML, De Haan HG, Hackinger S, Hai Y, Han S, Tybjærg-Hansen A, Hivert MF, Isomaa B, Jäger S, Jørgensen ME, Jørgensen T, Käräjämäki A, Kim BJ, Kim SS, Koistinen HA, Kovacs P, Kriebel J,

Kronenberg F, Läll K, Lange LA, Lee JJ, Lehne B, Li H, Lin KH, Linneberg A, Liu CT, Liu J, et al. Refining the accuracy of validated target identification through coding variant fine-mapping in type 2 diabetes article. *Nat Genet*. 2018;50:559–571.

28. Tadros R, Francis C, Xu X, Vermeer AMC, Harper AR, Huurman R, Kelu Bisabu K, Walsh R, Hoorntje ET, te Rijdt WP, Buchan RJ, van Velzen HG, van Slegtenhorst MA, Vermeulen JM, Offerhaus JA, Bai W, de Marvao A, Lahrouchi N, Beekman L, Karper JC, Veldink JH, Kayvanpour E, Pantazis A, Baksi AJ, Whiffin N, Mazzarotto F, Sloane G, Suzuki H, Schneider-Luftman D, Elliott P, Richard P, Ader F, Villard E, Lichtner P, Meitinger T, Tanck MWT, van Tintelen JP, Thain A, McCarty D, Hegele RA, Roberts JD, Amyot J, Dubé M-P, Cadrin-Tourigny J, Giraldeau G, L'Allier PL, Garceau P, Tardif J-C, Boekholdt SM, Lumbers RT, Asselbergs FW, Barton PJR, Cook SA, Prasad SK, O'Regan DP, van der Velden J, Verweij KJH, Talajic M, Lettre G, Pinto YM, Meder B, Charron P, de Boer RA, Christiaans I, Michels M, Wilde AAM, Watkins H, Matthews PM, Ware JS, Bezzina CR. Shared genetic pathways contribute to risk of hypertrophic and dilated cardiomyopathies with opposite directions of effect. *Nat Genet*. 2021;53:128–134.

29. Pietzner M, Wheeler E, Carrasco-Zanini J, Cortes A, Koprulu M, Wörheide MA, Oerton E, Cook J, Stewart ID, Kerrison ND, Luan J, Raffler J, Arnold M, Arlt W, O'Rahilly S, Kastenmüller G, Gamazon ER, Hingorani AD, Scott RA, Wareham NJ, Langenberg C. Mapping the proteo-genomic convergence of human diseases. *Science (1979)*. 2021;374:eabj1541.

30. Salo PP, Havulinna AS, Tukiainen T, Raitakari O, Lehtimäki T, Kähönen M, Kettunen J, Männikkö M, Eriksson JG, Jula A, Blankenberg S, Zeller T, Salomaa V, Kristiansson K, Perola M. Genome-Wide Association Study Implicates Atrial Natriuretic Peptide Rather Than B-Type Natriuretic Peptide in the Regulation of Blood Pressure in the General Population. *Circ Cardiovasc Genet*. 2017;10:e001713.

31. Harper AR, Goel A, Grace C, Thomson KL, Petersen SE, Xu X, Waring A, Ormondroyd E, Kramer CM, Ho CY, Neubauer S, Kolm P, Kwong R, Dolman SF, Desvigne-Nickens P, Dimarco JP, Geller N, Kim D-Y, Zhang C, Weintraub W, Abraham T, Anderson L, Appelbaum E, Autore C, Berry C, Biagini E, Bradlow W, Bucciarelli-Ducci C, Chiribiri A, Choudhury L, Crean A, Dawson D, Desai MY, Elstein E, Flett A, Friedrich M, Heitner S, Helms A, Jacoby DL, Kim H, Kim B, Larose E, Mahmod M, Mahrholdt H, Maron M, McCann G, Michels M, Mohiddin S, Nagueh S, Newby D, Olivotto I, Owens A, Pierre-Mongeon F, Prasad S, Rimoldi O, Salerno M, Schulz-Menger J, Sherrid M, Swoboda P, van Rossum A, Weinsaft J, White J, Williamson E, Tadros R, Ware JS, Bezzina CR, Farrall M, Watkins H, Investigators H. Common genetic variants and modifiable risk factors underpin hypertrophic cardiomyopathy susceptibility and expressivity. *Nat Genet*. 2021;53:135–142.

32. Roselli C, Chaffin MD, Weng LC, Aeschbacher S, Ahlberg G, Albert CM, Almgren P, Alonso A, Anderson CD, Aragam KG, Arking DE, Barnard J, Bartz TM, Benjamin EJ, Bihlmeyer NA, Bis JC, Bloom HL, Boerwinkle E, Bottinger EB, Brody JA, Calkins H, Campbell A, Cappola TP, Carlquist J, Chasman DI, Chen LY, Chen YDI, Choi EK, Choi SH, Christophersen IE, Chung MK, Cole JW, Conen D, Cook J, Crijns HJ, Cutler MJ, Damrauer SM, Daniels BR, Darbar D, Delgado G, Denny JC, Dichgans M, Dörr M, Dudink EA, Dudley SC, Esa N, Esko T, Eskola M, Fatkin D, Felix SB, Ford I, Franco OH, Geelhoed B, Grewal RP, Gudnason V, Guo X, Gupta N, Gustafsson S, Gutmann R, Hamsten A, Harris TB, Hayward C, Heckbert SR, Hernesniemi J, Hocking LJ, Hofman A, Horimoto ARVR, Huang J, Huang PL, Huffman J, Ingelsson E, Ipek EG, Ito K, Jimenez-Conde J, Johnson R, Jukema JW, Kääb S, Kähönen M, Kamatani Y, Kane JP, Kastrati A, Kathiresan S, Katschnig-Winter P, Kavousi M, Kessler T, Kietselaer BL, Kirchhof P, Kleber ME, Knight S, Krieger JE, Kubo M, Launer LJ, Laurikka J, Lehtimäki T, Leineweber K, Lemaitre RN, Li M, Lim HE, et al. Multi-ethnic genome-wide association study for atrial fibrillation. *Nat Genet*. 2018;50:1225–1233.

33. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, Saleheen D, Kyriakou T, Nelson CP, CHopewell J, Webb TR, Zeng L, Dehghan A, Alver M, MArmasu S, Auro K, Bjonnes A, Chasman DI, Chen S, Ford I, Franceschini N, Gieger C, Grace C, Gustafsson S, Huang J, Hwang SJ, Kim YK, Kleber ME, Lau KW, Lu X, Lu Y, Lyytikäinen LP, Mihailov E, Morrison AC, Pervjakova N, Qu L, Rose LM, Salfati E, Saxena R, Scholz M, Smith A V., Tikkanen E, Uitterlinden A, Yang X, Zhang W, Zhao W, De Andrade M, De Vries PS, Van Zuydam NR, Anand SS, Bertram L, Beutner F, Dedoussis G, Frossard P, Gauguier D, Goodall AH, Gottesman O, Haber M, Han BG, Huang J, Jalilzadeh S, Kessler T, König IR, Lannfelt L, Lieb W, Lind L, MLindgren C, Lokki ML, Magnusson PK, Mallick NH, Mehra N, Meitinger T, Memon FUR, Morris AP, Nieminen MS, Pedersen NL, Peters A, Rallidis LS, Rasheed A, Samuel M, Shah SH, Sinisalo J, EStirrups K, Trompet S, Wang L, Zaman KS, Ardissino D, Boerwinkle E, Borecki IB, Bottinger EP, Buring JE, Chambers JC, Collins R, Cupples L, Danesh J, Demuth I, Elosua R, Epstein SE, et al. A comprehensive 1000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat Genet*. 2015;47:1121–1130.

34. Carrasco-Zanini J, Pietzner M, Davitte J, Surendran P, Croteau-Chonka DC, Robins C, Torralbo A, Tomlinson C, Grünschläger F, Fitzpatrick N, Ytsma C, Kanno T, Gade S, Freitag D, Ziebell F, Haas S, Denaxas S, Betts JC, Wareham NJ, Hemingway H, Scott RA, Langenberg C. Proteomic signatures improve risk prediction for common and rare diseases. *Nat Med*. 2024;30:2489–2498.

35. Heijink IH, Hackett TL, Pouwels SD. Effects of cigarette smoking on SARS-CoV-2 receptor ACE2 expression in the respiratory epithelium†. Journal of Pathology. 2021;253:351–354.

36. Smith JC, Sausville EL, Girish V, Yuan M Lou, Vasudevan A, John KM, Sheltzer JM. Cigarette Smoke Exposure and Inflammatory Signaling Increase the Expression of the SARS-CoV-2 Receptor ACE2 in the Respiratory Tract. *Dev Cell*. 2020;53:514–529.

37. Friske MM, Giannone F, Senger M, Seitz R, Hansson AC, Spanagel R. Chronic alcohol intake regulates expression of SARS-CoV2 infection-relevant genes in an organ-specific manner. *Alcohol: Clinical and Experimental Research*. 2023;47:76–86.

38. Solopov PA, Colunga Biancatelli RML, Catravas JD. Alcohol Increases Lung Angiotensin-Converting Enzyme 2 Expression and Exacerbates Severe Acute Respiratory Syndrome Coronavirus 2 Spike Protein Subunit 1–Induced Acute Lung Injury in K18-hACE2 Transgenic Mice. *American Journal of Pathology*. 2022;192:990–1000.

39. Balasubramanian N, James TD, Pushpavathi SG, Marcinkiewcz CA. Repeated ethanol exposure and withdrawal alters angiotensin-converting enzyme 2 expression in discrete brain regions: Implications for SARS-CoV-2 neuroinvasion. *Alcohol: Clinical and Experimental Research*. 2022;47:219–239.

40. Peter JG, Ntusi NAB, Ntsekhe M. Are Recommendations That Favor Other Agents Over Angiotensin-Converting Enzyme Inhibitors in Africans With Hypertension Justified? Circulation. 2024;149:804–806.

41. Bindom SM, Lazartigues E. The sweeter side of ACE2: Physiological evidence for a role in diabetes. Mol Cell Endocrinol. 2009;302:193–202.

42. Man JCK, Van Duijvenboden K, Krijger PHL, Hooijkaas IB, Van Der Made I, De Gier-De Vries C, Wakker V, Creemers EE, De Laat W, Boukens BJ, Christoffels VM. Genetic Dissection of a Super Enhancer Controlling the Nppa-Nppb Cluster in the Heart. *Circ Res*. 2021;128:115–129.

43. Yasuno S, Usami S, Kuwahara K, Nakanishi M, Arai Y, Kinoshita H, Nakagawa Y, Fujiwara M, Murakami M, Ueshima K, Harada M, Nakao K. Endogenous cardiac natriuretic peptides protect the heart in a mouse model of dilated cardiomyopathy and sudden death. *Am J Physiol Heart Circ Physiol*. 2009;296:H1804-10.

44. Holditch SJ, Schreiber CA, Nini R, Tonne JM, Peng KW, Geurts A, Jacob HJ, Burnett JC, Cataliotti A, Ikeda Y. B-type natriuretic peptide deletion leads to progressive hypertension, associated organ damage, and reduced survival: Novel model for human hypertension. *Hypertension*. 2015;66:199–210.

45. Chang P, Zhang X, Zhang J, Wang J, Wang X, Li M, Wang R, Yu J, Fu F. BNP protects against diabetic cardiomyopathy by promoting Opa1-mediated mitochondrial fusion via activating the PKG-STAT3 pathway. *Redox Biol*. 2023;62:102702.

46. Ho CY, Day SM, Axelsson A, Russell MW, Zahka K, Lever HM, Pereira AC, Colan SD, Margossian R, Murphy AM, Canter C, Bach RG, Wheeler MT, Rossano JW, Owens AT, Bundgaard H, Benson L, Mestroni L, Taylor MRG, Patel AR, Wilmot I, Thrush P, Vargas JD, Soslow JH, Becker JR, Seidman CE, Lakdawala NK, Cirino AL, Krieger JE, Sacilotto L, Arteaga E, Antunes MO, Hall EK, Choudhury L, Pahl E, Lin KY, Lewis GD, Desai AS, Burns KM, McMurray JJV, MacRae CA, Solomon SD, Orav EJ, Braunwald E. Valsartan in early-stage hypertrophic cardiomyopathy: a randomized phase 2 trial. *Nat Med*. 2021;27:1818–1824.

47. Velicki L, Popovic D, Okwose NC, Preveden A, Tesic M, Tafelmeier M, Charman SJ, Macgowan GA, Filipovic N, Ristic A, Olivotto I, Maier LS, Jakovljevic DG. Sacubitril/valsartan for treatment of symptomatic non-obstructive hypertrophic cardiomyopathy: a randomised, controlled, phase II clinical trial (SILICOFCM). *Eur J Heart Fail*. 2024;26:1361–1368.

48. Solomon SD, Vaduganathan M, Claggett BL, Packer M, Zile M, Swedberg K, Rouleau J, Pfeffer MA, Desai A, Lund LH, Kober L, Anand I, Sweitzer N, Linssen G, Merkely B, Luis Arango J, Vinereanu D, Chen CH, Senni M, Sibulo A, Boytsov S, Shi V, Rizkala A, Lefkowitz M, Mcmurray JJV. Sacubitril/Valsartan Across the Spectrum of Ejection Fraction in Heart Failure. *Circulation*. 2020;141:352–361.

49. Kara K, Lehmann N, Neumann T, Kälsch H, Möhlenkamp S, Dykun I, Broecker-Preuss M, Pundt N, Moebus S, Jöckel KH, Erbel R, Mahabadi AA. NT-proBNP is superior to BNP for predicting first cardiovascular events in the general population: The Heinz Nixdorf Recall study. *Int J Cardiol*. 2015;183:155–161.

50. Gill D, Seidler T, Troughton RW, Yandle TG, Frampton CM, Richards M, Lainchbury JG, Nicholls G. Vigorous response in plasma N-terminal pro-brain natriuretic peptide (NT-BNP) to acute myocardial infarction. *Clin Sci*. 2004;106:135–139.

51. Rørth R, Jhund PS, Yilmaz MB, Kristensen SL, Welsh P, Desai AS, Køber L, Prescott MF, Rouleau JL, Solomon SD, Swedberg K, Zile MR, Packer M, McMurray JJV. Comparison of BNP and NT-proBNP in Patients With Heart Failure and Reduced Ejection Fraction. *Circ Heart Fail*. 2020;13:e006541.

52. Welsh P, Campbell RT, Mooney L, Kimenai DM, Hayward C, Campbell A, Porteous D, Mills NL, Lang NN, Petrie MC, Januzzi JL, McMurray JJV, Sattar N. Reference Ranges for NT-proBNP (N-Terminal Pro-B-Type Natriuretic Peptide) and Risk Factors for Higher NT-proBNP Concentrations in a Large General Population Cohort. *Circ Heart Fail*. 2022;15:e009427.

53. Gupta DK, Daniels LB, Cheng S, deFilippi CR, Criqui MH, Maisel AS, Lima JA, Bahrami H, Greenland P, Cushman M, Tracy R, Siscovick D, Bertoni AG, Cannone V, Burnett JC, Carr JJ, Wang TJ. Differences in Natriuretic

Peptide Levels by Race/Ethnicity (From the Multi-Ethnic Study of Atherosclerosis). *American Journal of Cardiology*. 2017;120:1008–1015.

54. Gupta DK, Claggett B, Wells Q, Cheng S, Li M, Maruthur N, Selvin E, Coresh J, Konety S, Butler KR, Mosley T, Boerwinkle E, Hoogeveen R, Ballantyne CM, Solomon SD. Racial Differences in Circulating Natriuretic Peptide Levels: The Atherosclerosis Risk in Communities Study. *J Am Heart Assoc*. 2015;4:e001831.

55. Chouairi F, Pacor J, Miller PE, Fuery MA, Caraballo C, Sen S, Leifer ES, Felker GM, Fiuzat M, O'Connor CM, Januzzi JL, Friedman DJ, Desai NR, Ahmad T, Freeman J V. Effects of Atrial Fibrillation on Heart Failure Outcomes and NT-proBNP Levels in the GUIDE-IT Trial. *Mayo Clin Proc Innov Qual Outcomes*. 2021;5:447–455.

56. Zhao X, Li H, Liu C, Ren Y, Sun C. NT Pro-BNP can be used as a risk predictor of clinical atrial fibrillation with or without left atrial enlargement. *Clin Cardiol*. 2022;45:68–74.

57. Lee YS. Bradyarrhythmia Can Increase the Plasma Level of N-Terminal Pro-Brain Natriuretic Peptide. *International Journal of Arrhythmia*. 2016;17:112–117.

58. Kováts T, Wettstein A, Nagy E, Tomcsányi J. Bradycardia can induce increased serum natriuretic peptide-level. *Int J Cardiol*. 2008;123:e43–e44.

Supplementary information

Supplementary figures

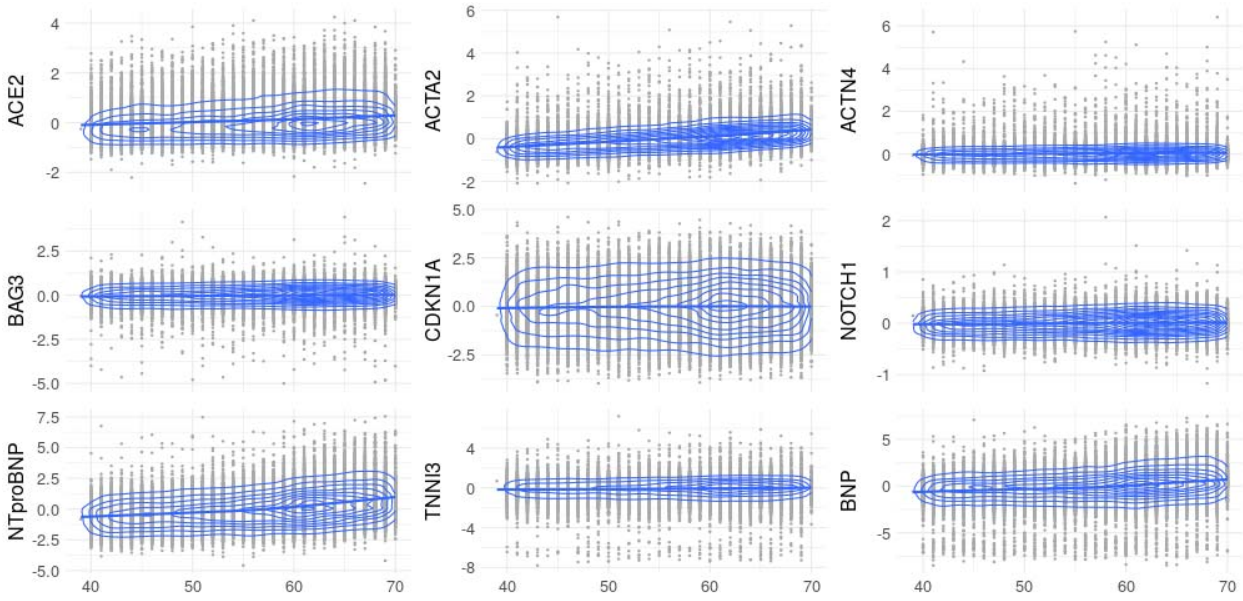


Figure S1 The relationships with age. The plots depict the relationships between age at recruitment (x-axis, years) and the nine plasma protein levels. Pearson's correlation coefficient between the protein levels and age were as follows; R=0.16 ACE2, R=0.42 ACTA2, R=0.34 NT-proBNP, R=0.06 ACTN2, R=0.06 BAG3, R=0.02 CDKN1A, R=0.06 NOTCH1, R=0.06 TNNI3, and R=0.23 BNP.

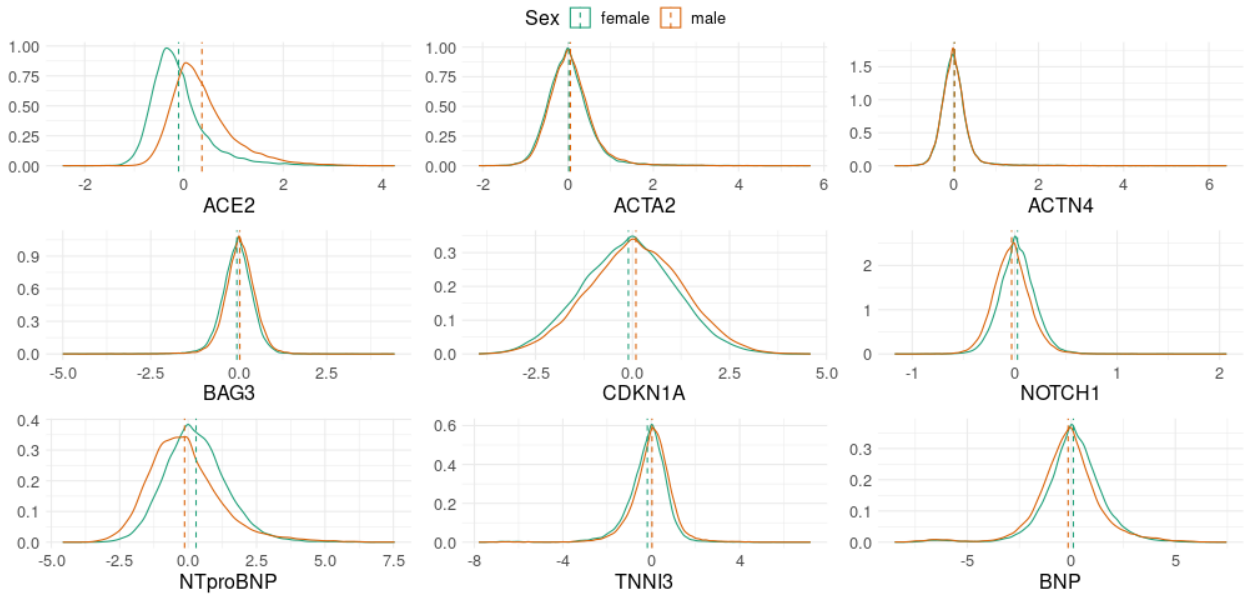
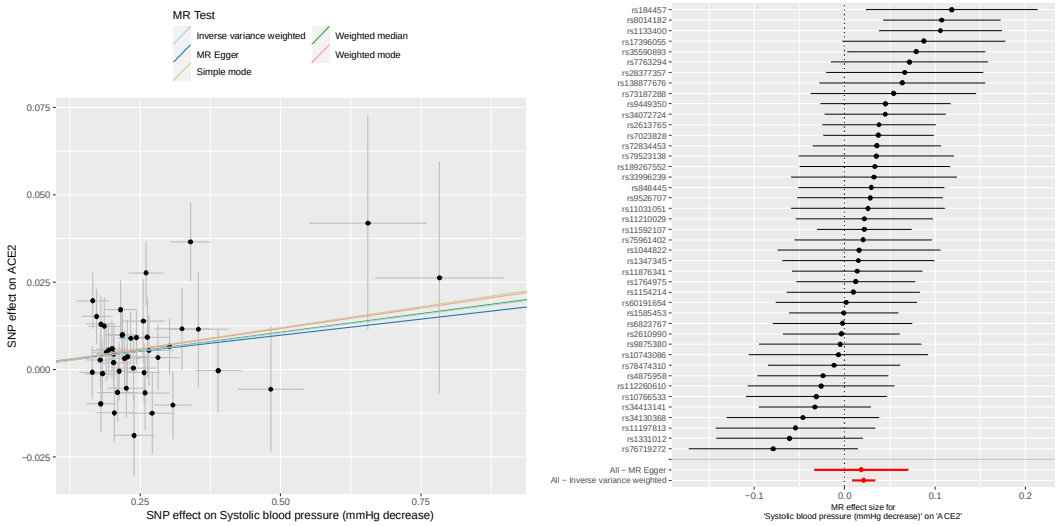


Figure S2 The relationships with sex. The distributions depict the relationships between sex and the nine plasma protein levels. ACE2 ($P=1.0x10^{-16}$, β=0.47), ACTA2 ($P=1.08x10^{-13}$, β=0.04), BAG3 ($P=4.97x10^{-85}$, β=0.08), CDKN1A ($P=6.61x10^{-68}$, β=0.20), and TNNI3 ($P=3.97x10^{-127}$, β=0.21) were significantly increased with male sex. NOTCH1 ($P=8.61x10^{-258}$, β=0.06), NT-proBNP ($P=1.62x10^{-278}$, β=0.42), and BNP ($P=3.50x10^{-69}$, β=0.25), were significantly increased with female sex.

a) **b)**



c) **d)**

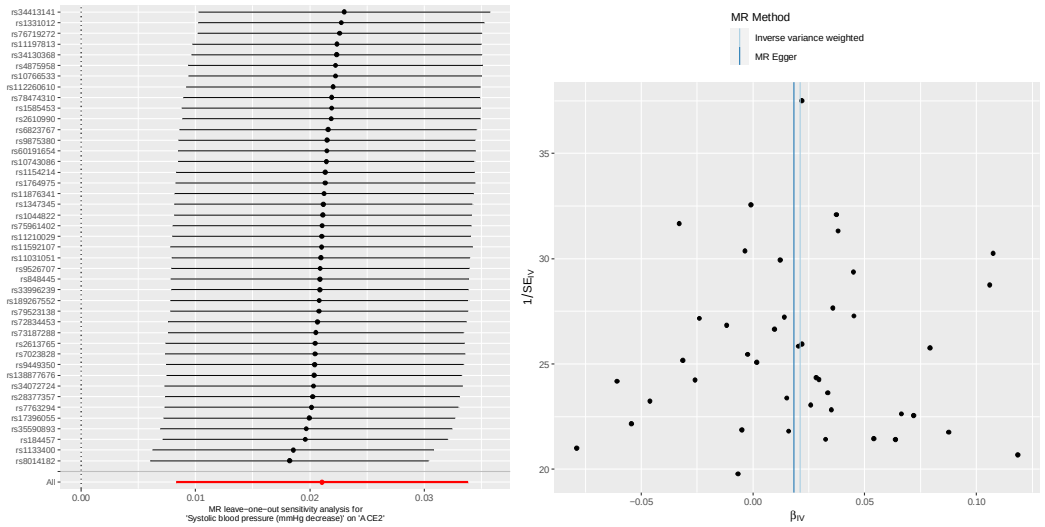
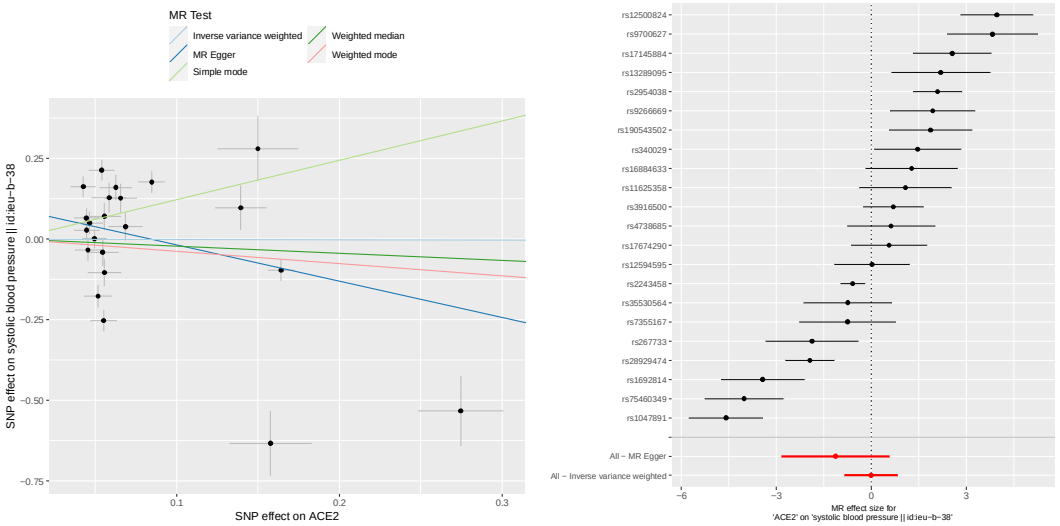


Figure S3 Mendelian randomization analysis of decreased systolic blood pressure as an exposure for ACE2 levels. The plots show summary information on the analyses, performed as per the TwoSampleMR R package. **a)** Mendelian randomization scatter plot for decreased systolic blood pressure as an exposure for ACE2. **b)** Mendelian randomization single SNP funnel plot for decreased systolic blood

pressure as an exposure for ACE2. **c)** Mendelian randomization single SNP forest plot for decreased systolic blood pressure as an exposure for ACE2. **d)** Mendelian randomization leave one out plot for decreased systolic blood pressure as an exposure for ACE2.

a) **b)**



c) **d)**

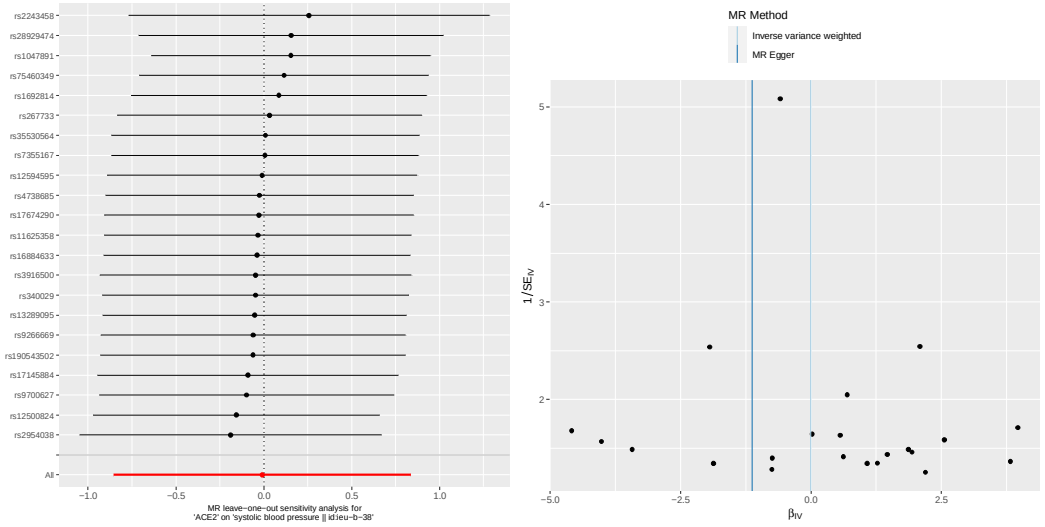
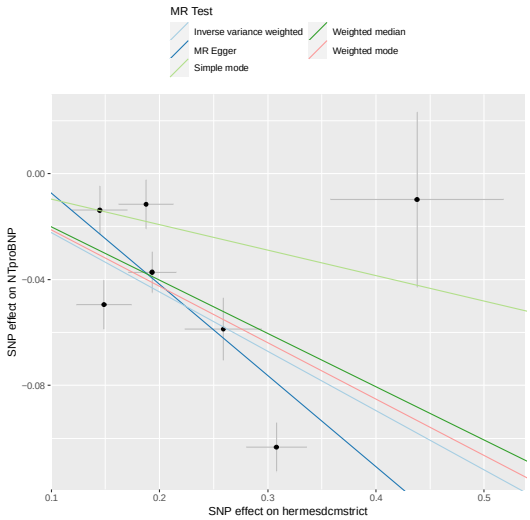


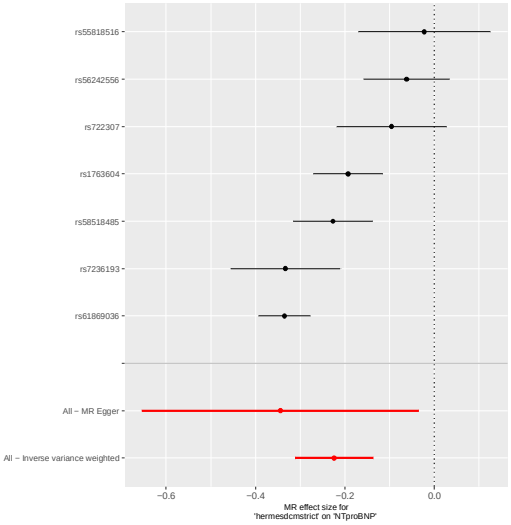
Figure S4 Mendelian randomization analysis of ACE2 levels as an exposure for systolic blood pressure. The plots show summary information on the analyses, performed as per the TwoSampleMR R package. **a)** Mendelian randomization scatter plot for systolic blood pressure as an outcome of ACE2 exposure. **b)** Mendelian randomization single SNP funnel plot for systolic blood pressure as an outcome of

ACE2 exposure. **c)** Mendelian randomization single SNP forest plot for systolic blood pressure as an outcome of ACE2 exposure. **d)** Mendelian randomization leave one out plot for systolic blood pressure as an outcome of ACE2 exposure.

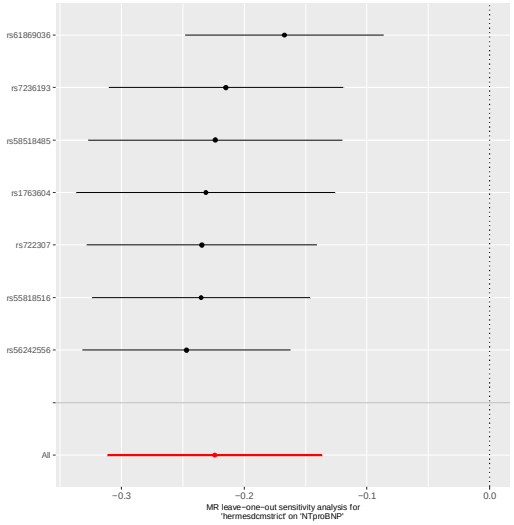
a)


MR Test
Inverse variance weighted
MR Egger
Simple mode
Weighted median
Weighted mode
SNP effect on NTproBNP
SNP effect on hermesdcmstrict

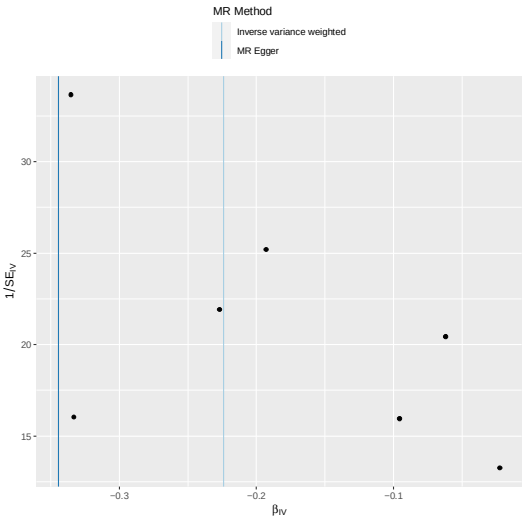
b)


rs55818516
rs56242556
rs722307
rs1763604
rs58518485
rs7236193
rs61869036
All − MR Egger
All − Inverse variance weighted
MR effect size for
'hermesdcmstrict' on 'NTproBNP'

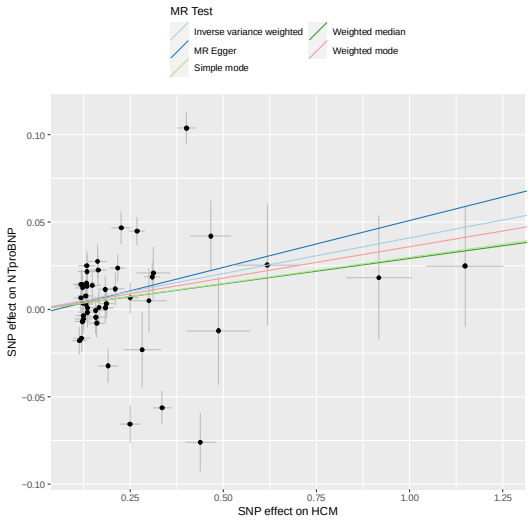
c)


rs61869036
rs7236193
rs58518485
rs1763604
rs722307
rs55818516
rs56242556
All
MR leave−one−out sensitivity analysis for
'hermesdcmstrict' on 'NTproBNP'

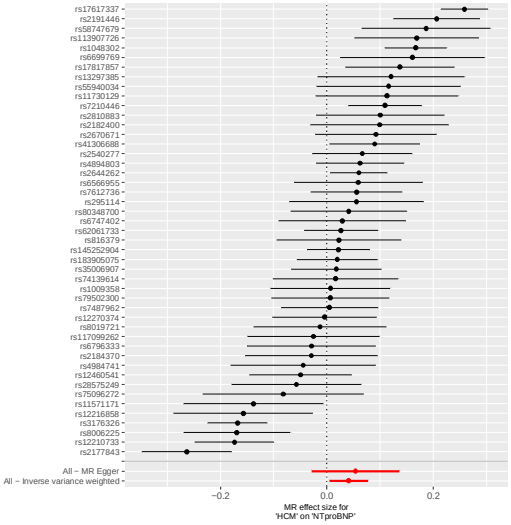
d)


MR Method
Inverse variance weighted
MR Egger

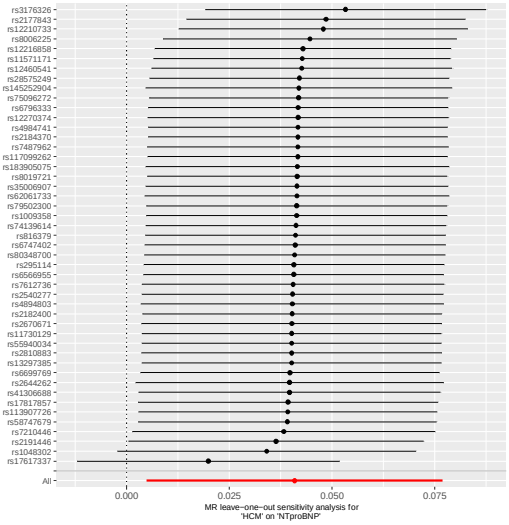
e)


MR Test
Inverse variance weighted
MR Egger
Simple mode
Weighted median
Weighted mode
SNP effect on NTproBNP
SNP effect on HCM

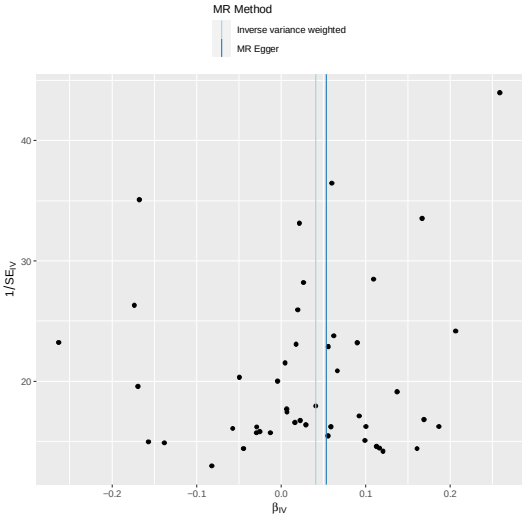
f)


All − MR Egger
All − Inverse variance weighted
MR effect size for
'HCM' on 'NTproBNP'

g)


All
MR leave−one−out sensitivity analysis for
'HCM' on 'NTproBNP'

h)


MR Method
Inverse variance weighted
MR Egger
1/SE_IV
β_IV

i) **j)**

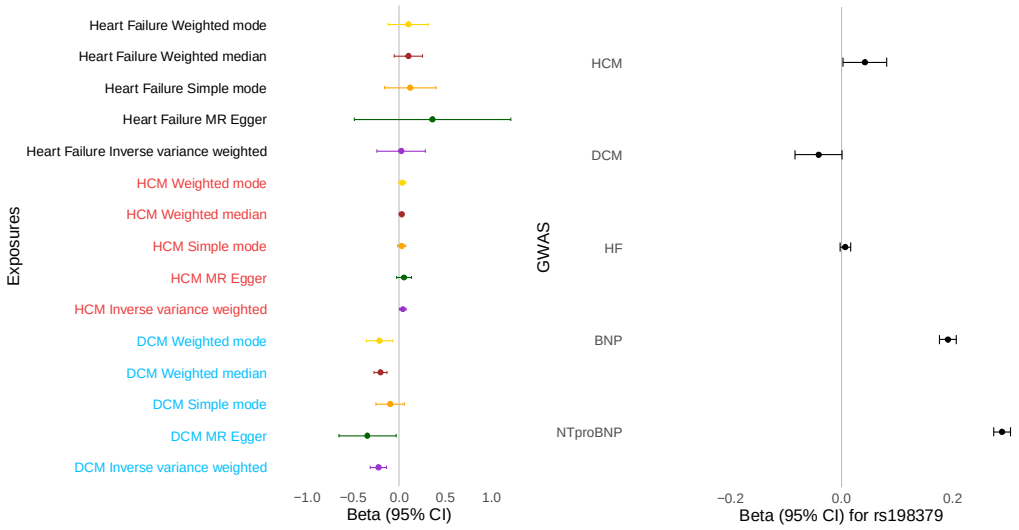
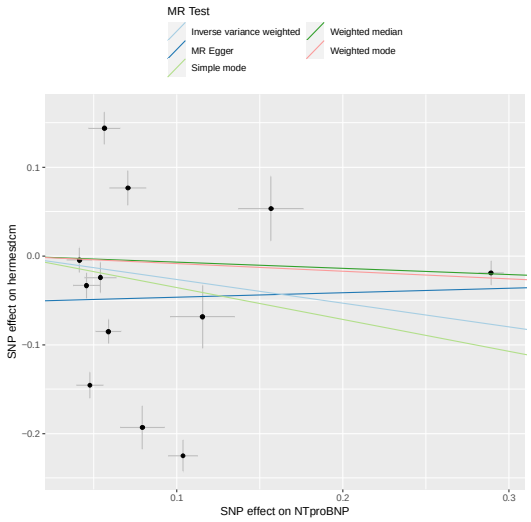


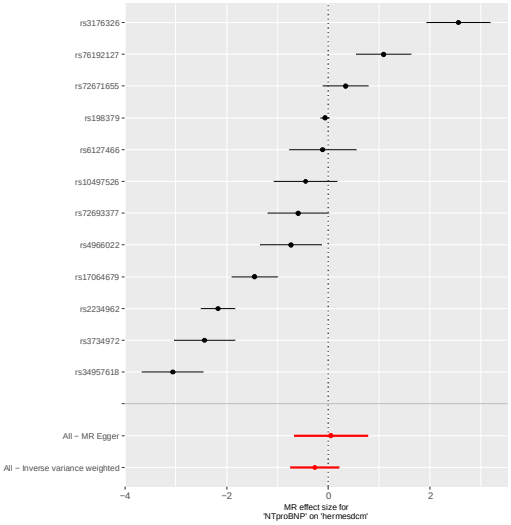
Figure S5 Mendelian randomization analysis of cardiomyopathies as an exposure for NT-proBNP levels. Mendelian randomisation for cardiomyopathy exposures: **a)-d)** DCM and **e)-h)** HCM, for NT-proBNP outcome. Summary plots are presented in **i)-j)**. Left ventricular hypertrophy increases NT-proBNP and BNP circulating levels, with the effect mainly through variants in *BAG3* and *CLCNKA.* The plots show summary information on the analyses, performed as per the TwoSampleMR R package. **a)** Mendelian randomization scatter plot for dilated cardiomyopathy (DCM) as an exposure for NT-proBNP. **b)** Mendelian randomization single SNP funnel plot for DCM as an exposure for NT-proBNP. **c)** Mendelian randomization single SNP forest plot for DCM an exposure for NT-proBNP. **d)** Mendelian randomization leave one out plot for DCM as an exposure for NT-proBNP. **e)** Mendelian randomization scatter plot for hypertrophic cardiomyopathy (HCM) as an exposure for NT-proBNP. **f)** Mendelian randomization single SNP funnel plot for HCM as an exposure for NT-proBNP. **g)** Mendelian randomization single SNP forest plot for HCM as an exposure for NT-proBNP. **h)**

Mendelian randomization leave one out plot for HCM as an exposure for NT-proBNP. **i)** Mendelian randomization genetic determination model of CM and HF genetic instruments as exposures for NT-proBNP outcome. **j)** The effect size of the lead *NPPB* SNP (downstream variant rs198379) identified here to associate at GWAS with increased NT-proBNP and BNP, and from GWAS summary statistics of published case-control studies of cardiomyopathies and heart failure. The eQTL variant increases *NPPB* expression in the atria (GTEx). The results suggest that the variant increases NT-proBNP and BNP production, the risk of HCM and heart failure, and decreases the risk of DCM. See **Table S5** for further details.

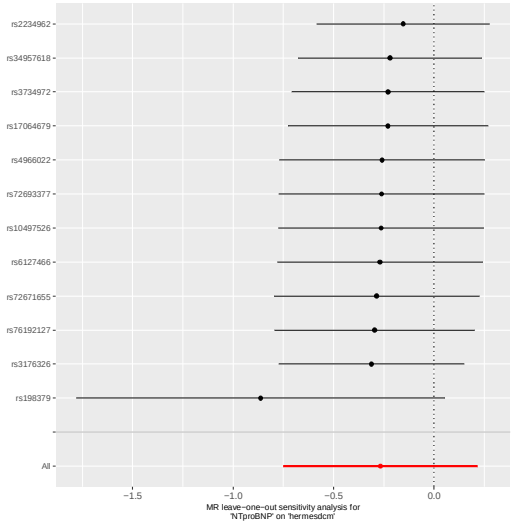
a)


MR Test
Inverse variance weighted
MR Egger
Simple mode
Weighted median
Weighted mode
SNP effect on hermesdcm
SNP effect on NTproBNP

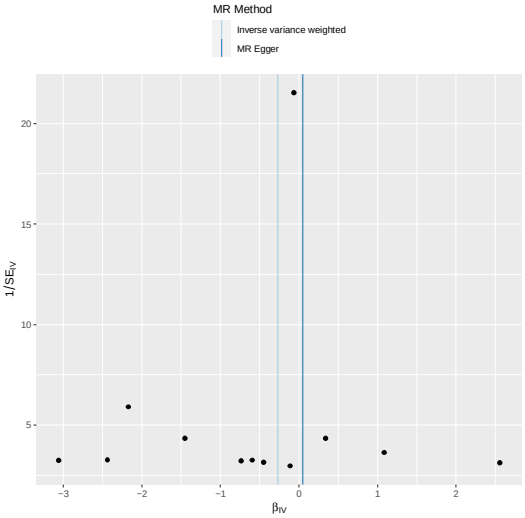
b)


rs3176326
rs76192127
rs72671655
rs198379
rs6127466
rs10497526
rs72693377
rs4966022
rs17064679
rs2234962
rs3734972
rs34957618
All − MR Egger
All − Inverse variance weighted
MR effect size for
'NTproBNP' on 'hermesdcm'

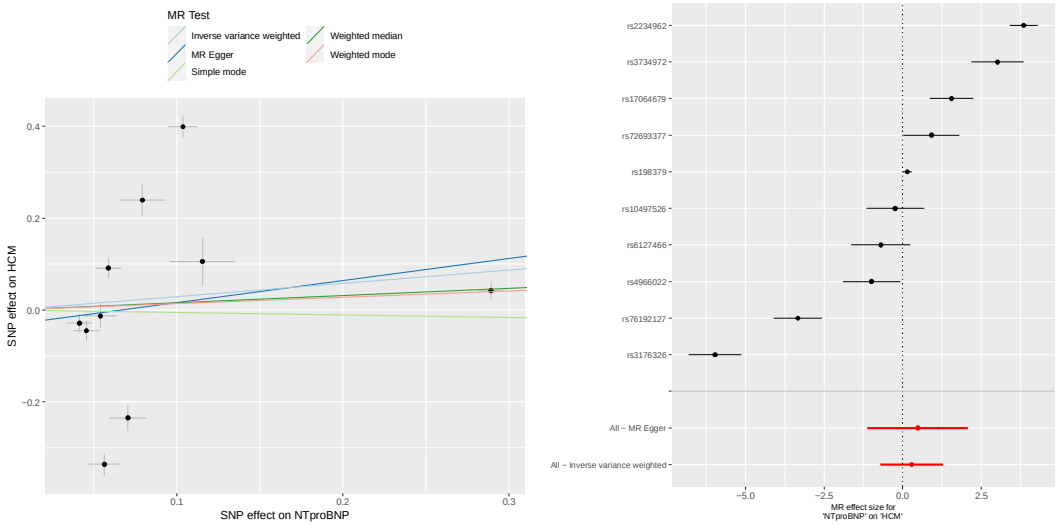
c)


rs2234962
rs34957618
rs3734972
rs17064679
rs4966022
rs72693377
rs10497526
rs6127466
rs72671655
rs76192127
rs3176326
rs198379
All
MR leave−one−out sensitivity analysis for
'NTproBNP' on 'hermesdcm'

d)


MR Method
Inverse variance weighted
MR Egger
1/SE_IV
β_IV

e) **f)**



g) **h)**

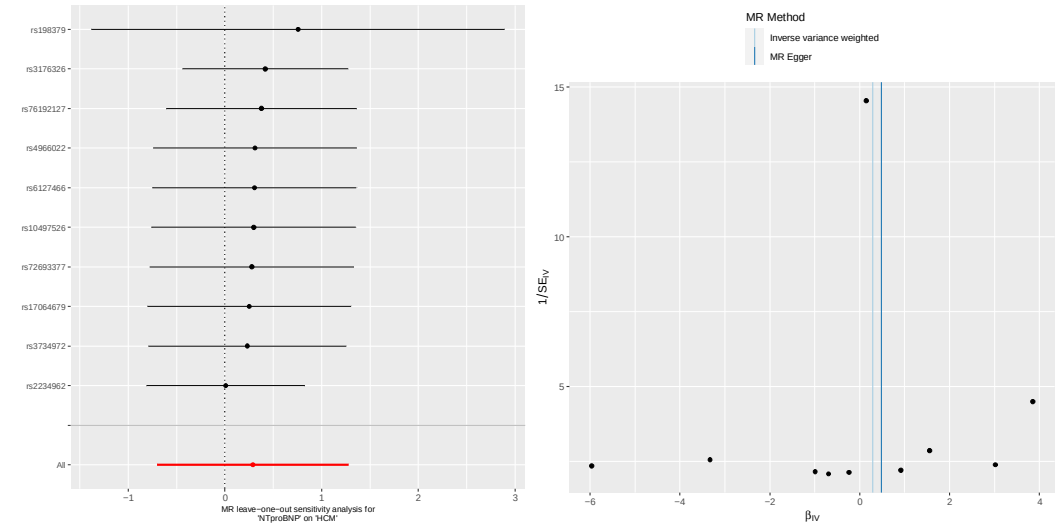
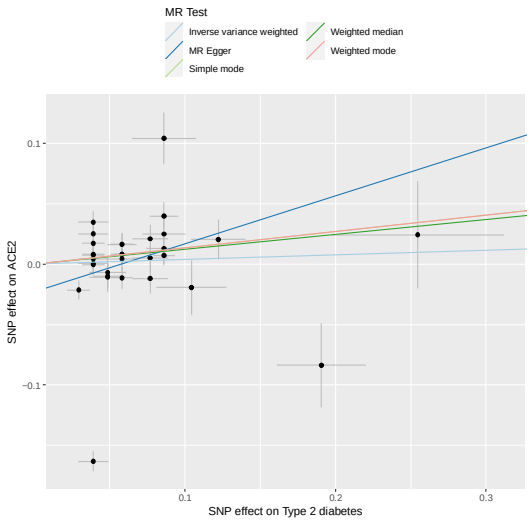


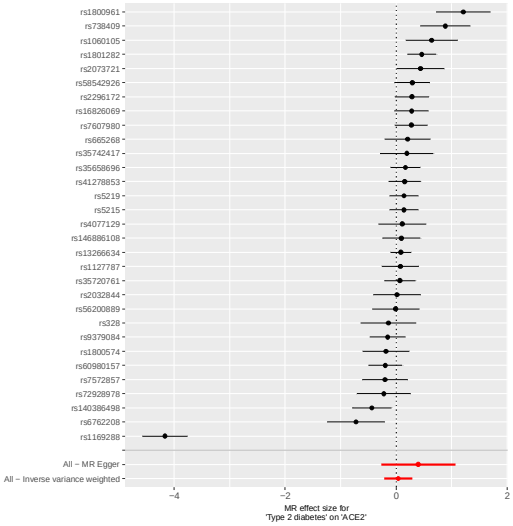
Figure S6 Mendelian randomization analysis of NT-proBNP as an exposure for cardiomyopathies. Mendelian randomisation for cardiomyopathy outcomes: **a)-d)** DCM and **e)-h)** HCM, for NT-proBNP exposure levels. The plots show summary information on the analyses, performed as per the TwoSampleMR R package. **a)** Mendelian randomization scatter plot for NT-proBNP as an exposure for dilated cardiomyopathy

(DCM). **b)** Mendelian randomization single SNP funnel plot for NT-proBNP as an exposure for DCM. **c)** Mendelian randomization single SNP forest plot for NT-proBNP as an exposure for DCM. **d)** Mendelian randomization leave one out plot for NT-proBNP as an exposure for DCM. **e)** Mendelian randomization scatter plot for NT-proBNP as an exposure for hypertrophic cardiomyopathy (HCM). **f)** Mendelian randomization single SNP funnel plot for NT-proBNP as an exposure for HCM. **g)** Mendelian randomization single SNP forest plot for NT-proBNP as an exposure for HCM. **h)** Mendelian randomization leave one out plot for NT-proBNP as an exposure for HCM.

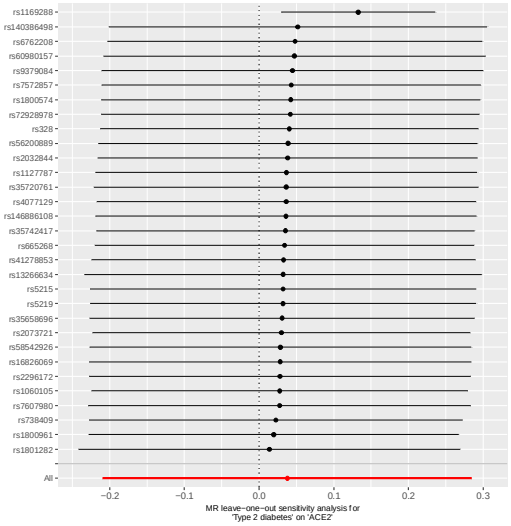
a)


MR Test
Inverse variance weighted
MR Egger
Simple mode
Weighted median
Weighted mode
SNP effect on ACE2
SNP effect on Type 2 diabetes

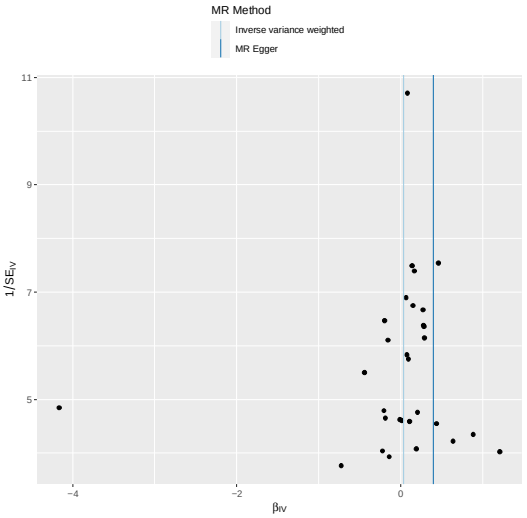
b)


All – MR Egger
All – Inverse variance weighted
MR effect size for
'Type 2 diabetes' on 'ACE2'

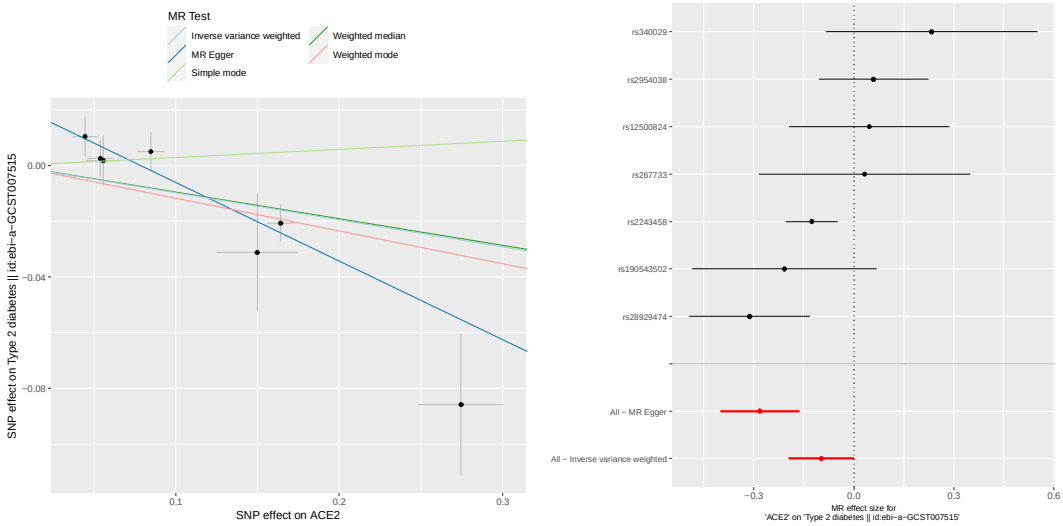
c)


All
MR leave−one−out sensitivity analysis for
'Type 2 diabetes' on 'ACE2'

d)


MR Method
Inverse variance weighted
MR Egger
1/SEIV
βIV

e) **f)**



g) **h)**

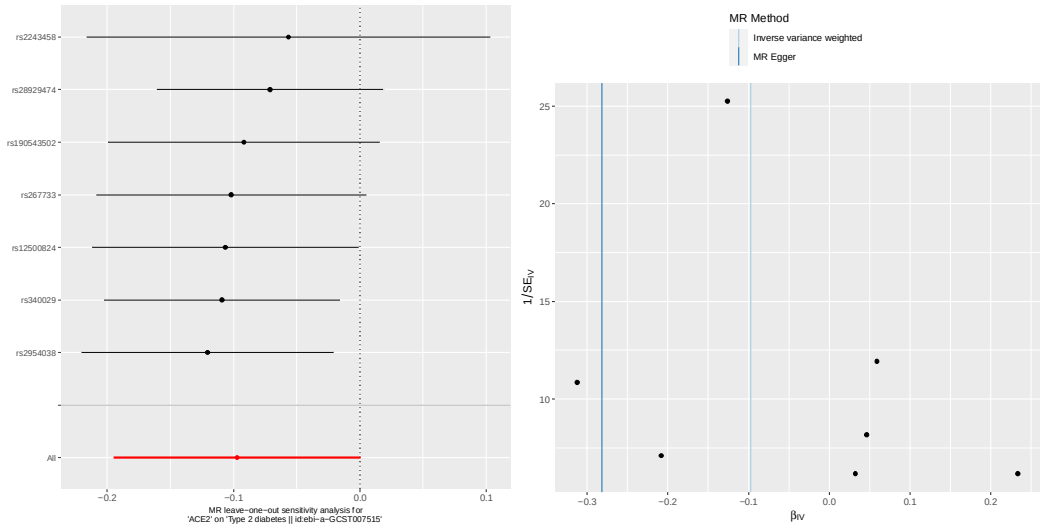
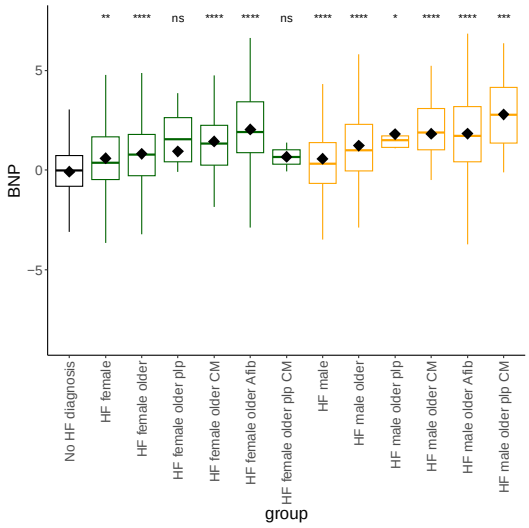


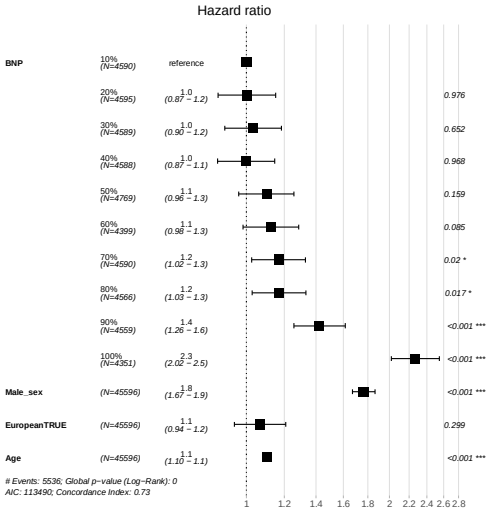
Figure S7 Mendelian randomization analysis of type-2 diabetes and ACE2 levels. Mendelian randomisation for type-2 diabetes as an exposure **a)-d)** and outcome **e)-h)** for ACE2 levels. The plots show summary information on the analyses, performed as per the TwoSampleMR R package. **a)** Mendelian randomization scatter plot for type-2 diabetes (T2D) as an exposure for ACE2. **b)** Mendelian randomization single SNP

funnel plot for T2D as an exposure for ACE2. **c)** Mendelian randomization single SNP forest plot for T2D as an exposure for ACE2. **d)** Mendelian randomization leave one out plot for T2D as an exposure for ACE2. **e)** Mendelian randomization scatter plot for ACE2 as an exposure for T2D. **f)** Mendelian randomization single SNP funnel plot for ACE2as an exposure for T2D. **g)** Mendelian randomization single SNP forest plot for ACE2as an exposure for T2D. **h)** Mendelian randomization leave one out plot for ACE2 as an exposure for T2D.

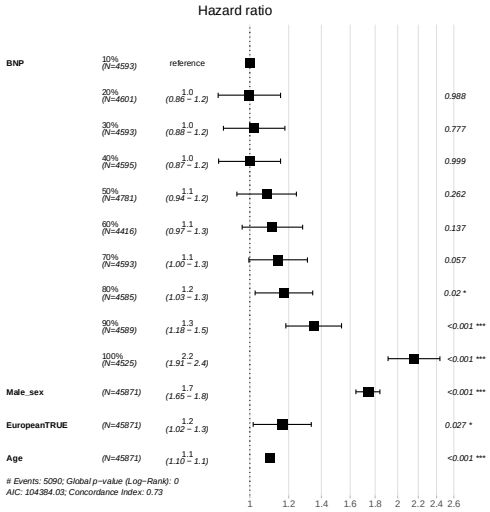
a)


** **** ns **** **** ns **** **** * **** **** ***
BNP
5
0
−5
No HF diagnosis
HF female
HF female older
HF female older plp
HF female older CM
HF female older Afib
HF female older plp CM
HF male
HF male older
HF male older plp
HF male older CM
HF male older Afib
HF male older plp CM
group

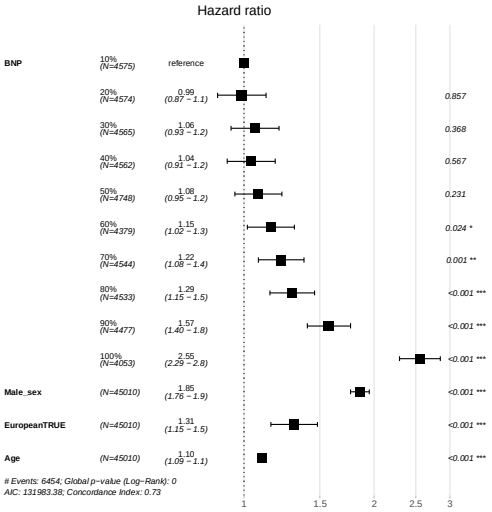
b)


Hazard ratio
BNP
10% (N=4590) reference
20% (N=4595) 1.0 (0.87 − 1.2) 0.976
30% (N=4589) 1.0 (0.90 − 1.2) 0.652
40% (N=4588) 1.0 (0.87 − 1.1) 0.968
50% (N=4769) 1.1 (0.96 − 1.3) 0.159
60% (N=4399) 1.1 (0.98 − 1.3) 0.085
70% (N=4590) 1.2 (1.02 − 1.3) 0.02 *
80% (N=4566) 1.2 (1.03 − 1.3) 0.017 *
90% (N=4559) 1.4 (1.26 − 1.6) <0.001 ***
100% (N=4351) 2.3 (2.02 − 2.5) <0.001 ***
Male_sex (N=45596) 1.8 (1.67 − 1.9) <0.001 ***
EuropeanTRUE (N=45596) 1.1 (0.94 − 1.2) 0.299
Age (N=45596) 1.1 (1.10 − 1.1) <0.001 ***
Events: 5536; Global p−value (Log−Rank): 0
AIC: 113490; Concordance Index: 0.73
1 1.2 1.4 1.6 1.8 2 2.2 2.4 2.6 2.8

c)


Hazard ratio
BNP
10% (N=4593) reference
20% (N=4601) 1.0 (0.86 − 1.2) 0.988
30% (N=4593) 1.0 (0.88 − 1.2) 0.777
40% (N=4595) 1.0 (0.87 − 1.2) 0.999
50% (N=4781) 1.1 (0.94 − 1.2) 0.262
60% (N=4416) 1.1 (0.97 − 1.3) 0.137
70% (N=4593) 1.1 (1.00 − 1.3) 0.057
80% (N=4585) 1.2 (1.03 − 1.3) 0.02 *
90% (N=4589) 1.3 (1.18 − 1.5) <0.001 ***
100% (N=4525) 2.2 (1.91 − 2.4) <0.001 ***
Male_sex (N=45871) 1.7 (1.65 − 1.8) <0.001 ***
EuropeanTRUE (N=45871) 1.2 (1.02 − 1.3) 0.027 *
Age (N=45871) 1.1 (1.10 − 1.1) <0.001 ***
Events: 5090; Global p−value (Log−Rank): 0
AIC: 104384.03; Concordance Index: 0.73
1 1.2 1.4 1.6 1.8 2 2.2 2.4 2.6

d)


Hazard ratio
BNP
10% (N=4575) reference
20% (N=4574) 0.99 (0.87 − 1.1) 0.857
30% (N=4565) 1.06 (0.93 − 1.2) 0.368
40% (N=4562) 1.04 (0.91 − 1.2) 0.567
50% (N=4748) 1.08 (0.95 − 1.2) 0.231
60% (N=4379) 1.15 (1.02 − 1.3) 0.024 *
70% (N=4544) 1.22 (1.08 − 1.4) 0.001 **
80% (N=4533) 1.29 (1.15 − 1.5) <0.001 ***
90% (N=4477) 1.57 (1.40 − 1.8) <0.001 ***
100% (N=4053) 2.55 (2.29 − 2.8) <0.001 ***
Male_sex (N=45010) 1.85 (1.76 − 1.9) <0.001 ***
EuropeanTRUE (N=45010) 1.31 (1.15 − 1.5) <0.001 ***
Age (N=45010) 1.10 (1.09 − 1.1) <0.001 ***
Events: 6454; Global p−value (Log−Rank): 0
AIC: 131983.38; Concordance Index: 0.73
1 1.5 2 2.5 3

e)

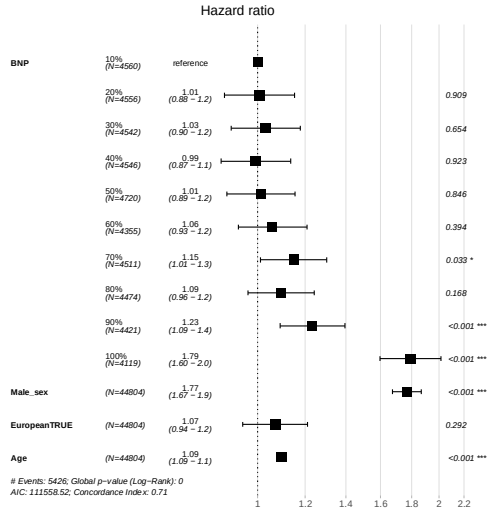


Figure S8 Increasing levels of BNP are observed in participants diagnosed with incident heart failure, cardiomyopathies, atrial fibrillation, and myocardial infarction. a) The figure shows the sequential increase in mean BNP with a heart failure diagnosis alongside sex, age, a diagnosis of cardiomyopathy or atrial fibrillation, and carriers of pathogenic cardiomyopathy-associated variants. The Student's t-test significance was using "No HF diagnosis" as the reference group. The groups contained the following sample sizes, respectively: (No HF diagnosis) 44050, (female groups:) 103, 286, 12, 33, 234, 2, (male groups:) 202, 496, 7, 42, 470, 14. NT-proBNP units are Olink's arbitrary unit in $\log_2$ scale. HF, heart failure; plp, P/LP variant carrier; Afib, atrial fibrillation; CM, cardiomyopathy diagnosis. **b, c, d)** Forest plots of Cox proportional hazards regression models for deciles of BNP levels with incident b) heart failure, c) cardiomyopathy, d) atrial fibrillation, and e) myocardial infarction, since recruitment. The forest plots were created assessing death or diagnosis from recruitment with those diagnosed before recruitment excluded. Sex (increasing risk is male), European

ancestry (increasing risk is European), and age at recruitment, were added to this multivariable analysis.

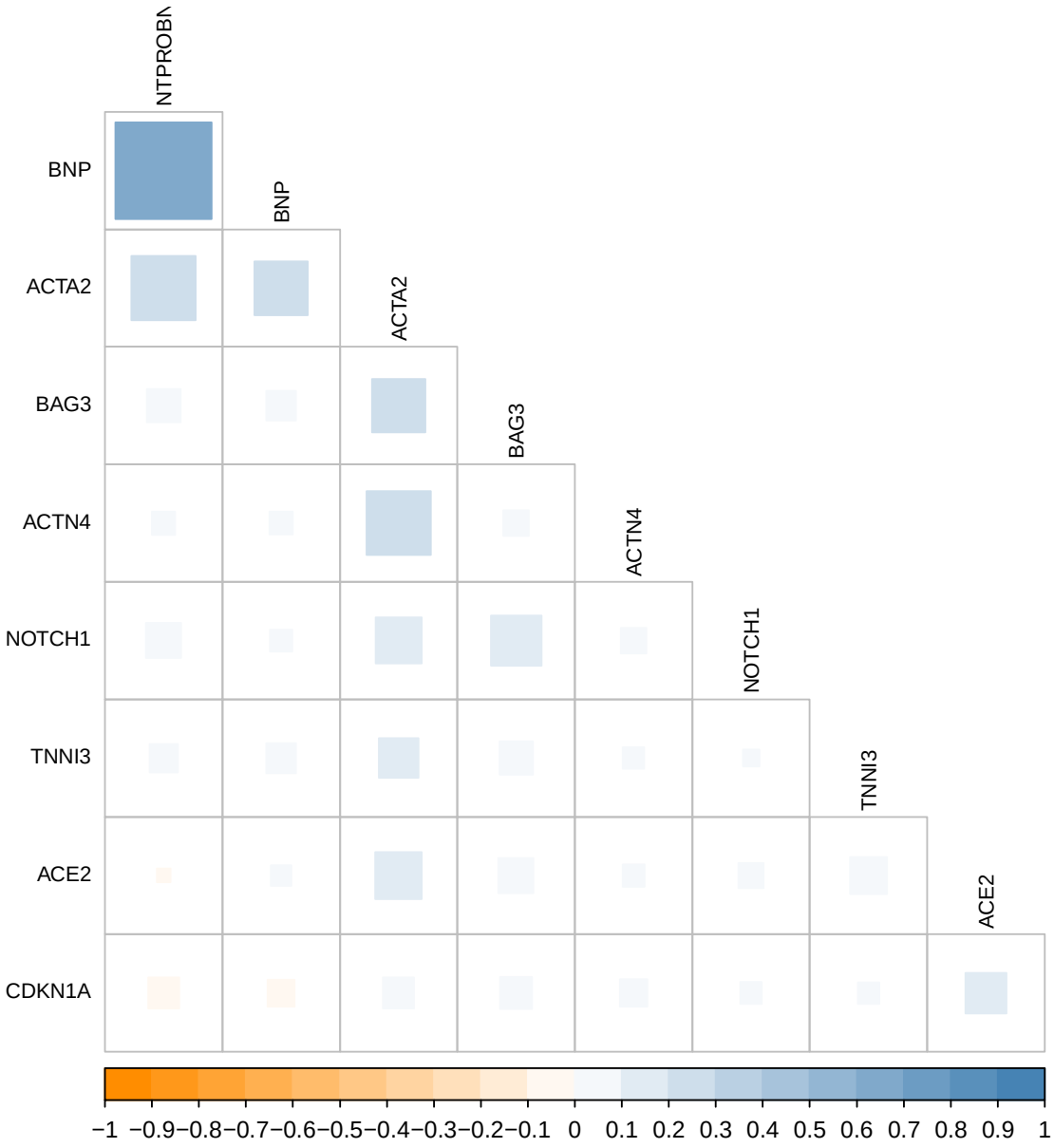
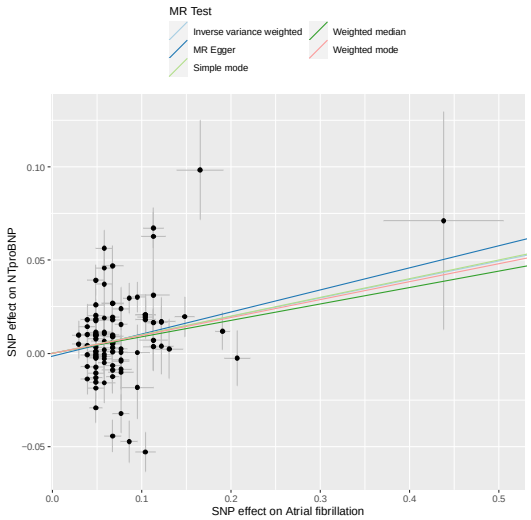
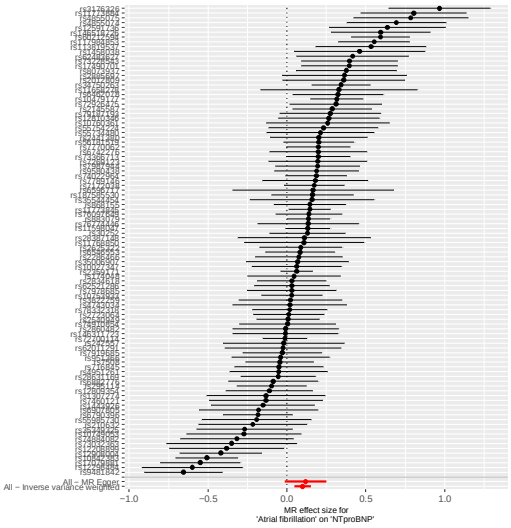


Figure S9 Correlation between nine circulating protein levels. The plot depicts the Pearson's correlation coefficient (R) between the raw levels of the circulating proteins analysed. ACTA2 had the most relationships. BNP and NT-proBNP correlated with the strongest relationship to R=0.67.

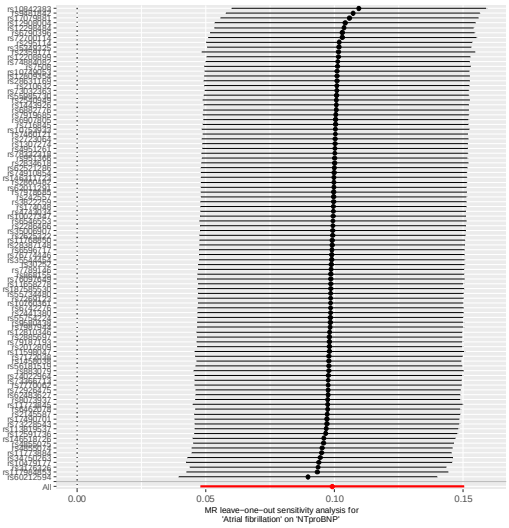
a)


MR Test
Inverse variance weighted
MR Egger
Simple mode
Weighted median
Weighted mode
SNP effect on NTproBNP
SNP effect on Atrial fibrillation

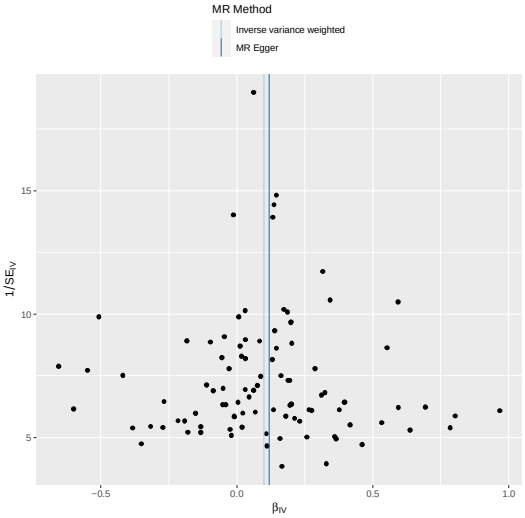
b)


All − MR Egger
All − Inverse variance weighted
MR effect size for
'Atrial fibrillation' on 'NTproBNP'

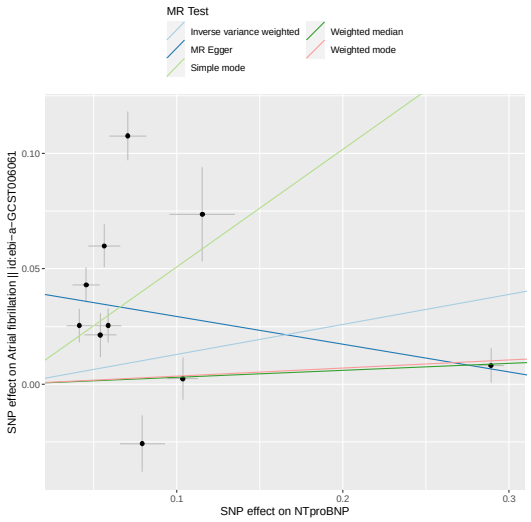
c)


All −
MR leave−one−out sensitivity analysis for
'Atrial fibrillation' on 'NTproBNP'

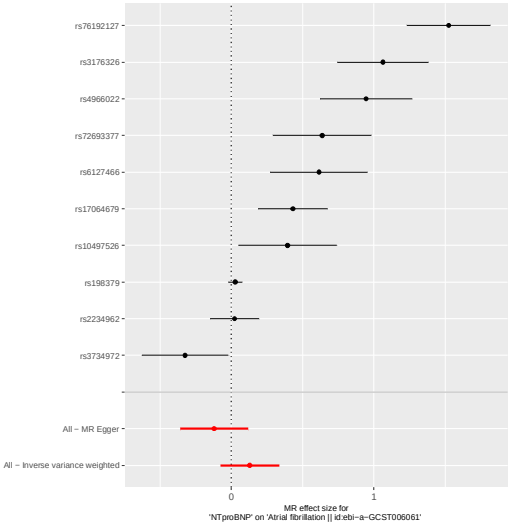
d)


MR Method
Inverse variance weighted
MR Egger
1/SE_IV
β_IV

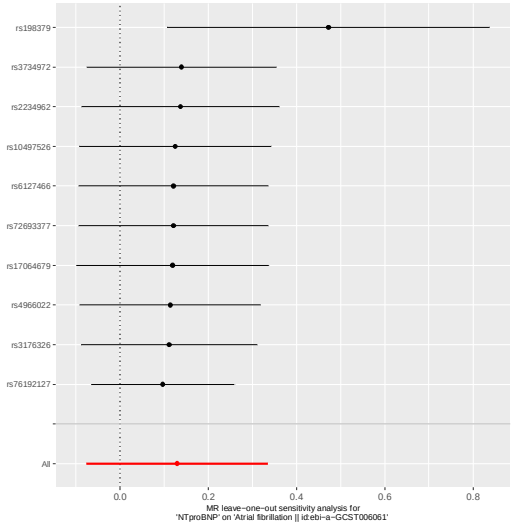
e)


MR Test
Inverse variance weighted
MR Egger
Simple mode
Weighted median
Weighted mode
SNP effect on Atrial fibrillation || id:ebi−a−GCST006061
SNP effect on NTproBNP

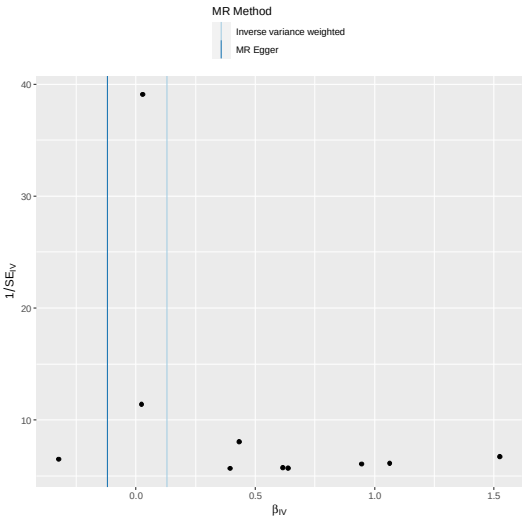
f)


rs76192127
rs3176326
rs4966022
rs72693377
rs6127466
rs17064679
rs10497526
rs198379
rs2234962
rs3734972
All − MR Egger
All − Inverse variance weighted
MR effect size for
'NTproBNP' on 'Atrial fibrillation || id:ebi−a−GCST006061'

g)


rs198379
rs3734972
rs2234962
rs10497526
rs6127466
rs72693377
rs17064679
rs4966022
rs3176326
rs76192127
All
MR leave−one−out sensitivity analysis for
'NTproBNP' on 'Atrial fibrillation || id:ebi−a−GCST006061'

h)


MR Method
Inverse variance weighted
MR Egger
1/SE_IV
β_IV

i)

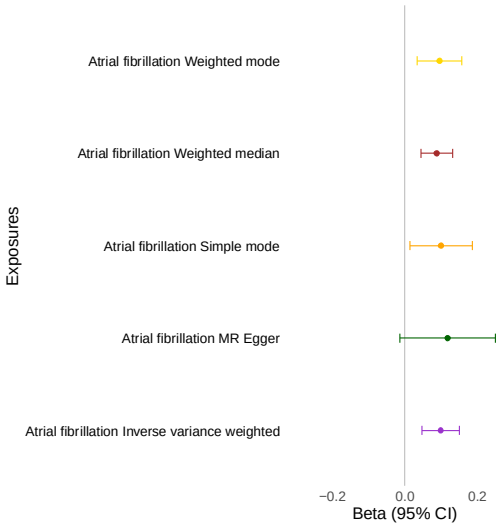
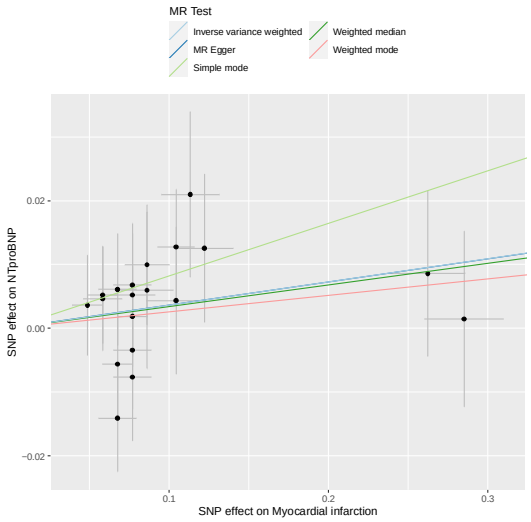
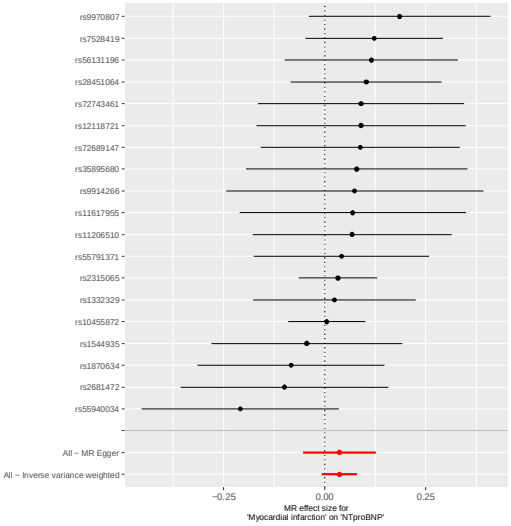


Figure S10 Mendelian randomization analysis of atrial fibrillation and NT-proBNP levels. Mendelian randomisation for atrial fibrillation (afib) as an exposure **a)-d)** and outcome **e)-h)** for NT-proBNP levels. The plots show summary information on the analyses, performed as per the TwoSampleMR R package. **a)** Mendelian randomization scatter plot for afib as an exposure for NT-proBNP. **b)** Mendelian randomization single SNP funnel plot for afib as an exposure for NT-proBNP. **c)** Mendelian randomization single SNP forest plot for afib as an exposure for NT-proBNP. **d)** Mendelian randomization leave one out plot for afib as an exposure for NT-proBNP. **e)** Mendelian randomization scatter plot for NT-proBNP as an exposure for afib. **f)** Mendelian randomization single SNP funnel plot for NT-proBNP as an exposure for afib. **g)** Mendelian randomization single SNP forest plot for NT-proBNP as an exposure for afib. **h)** Mendelian randomization leave one out plot for NT-proBNP as an exposure for afib. **i)** Summary plot of Mendelian randomization genetic determination model of atrial fibrillation genetic instruments as exposures for NT-proBNP outcome.

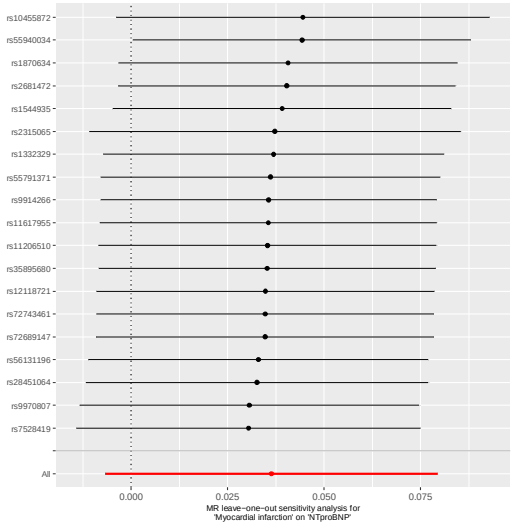
a)


MR Test
Inverse variance weighted
MR Egger
Simple mode
Weighted median
Weighted mode
SNP effect on NTproBNP
SNP effect on Myocardial infarction

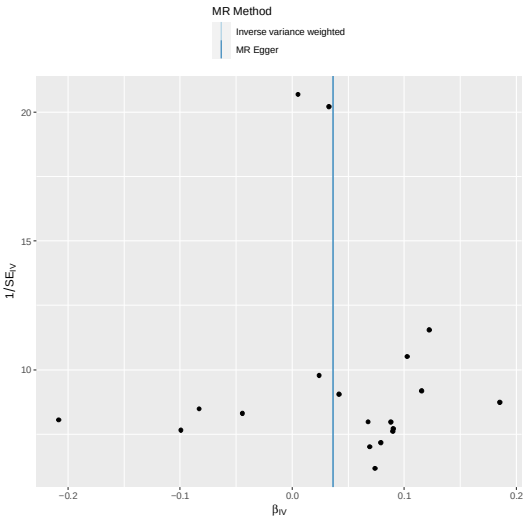
b)


rs9970807
rs7528419
rs56131196
rs28451064
rs72743461
rs12118721
rs72689147
rs35895680
rs9914266
rs11617955
rs11206510
rs55791371
rs2315065
rs1332329
rs10455872
rs1544935
rs1870634
rs2681472
rs55940034
All − MR Egger
All − Inverse variance weighted
MR effect size for
'Myocardial infarction' on 'NTproBNP'

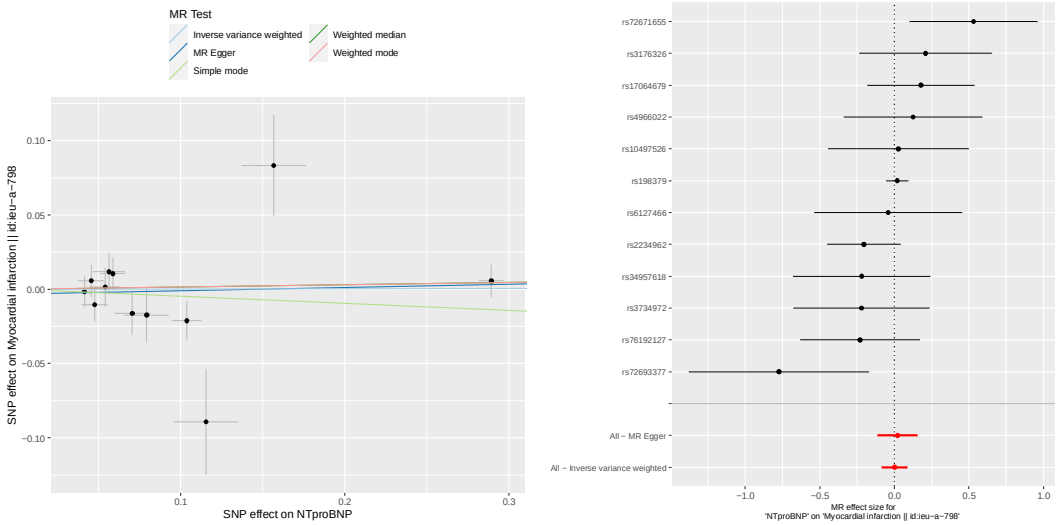
c)


rs10455872
rs55940034
rs1870634
rs2681472
rs1544935
rs2315065
rs1332329
rs55791371
rs9914266
rs11617955
rs11206510
rs35895680
rs12118721
rs72743461
rs72689147
rs56131196
rs28451064
rs9970807
rs7528419
All
MR leave−one−out sensitivity analysis for
'Myocardial infarction' on 'NTproBNP'

d)


MR Method
Inverse variance weighted
MR Egger
1/SE_IV
β_IV

e) **f)**



g) **h)**

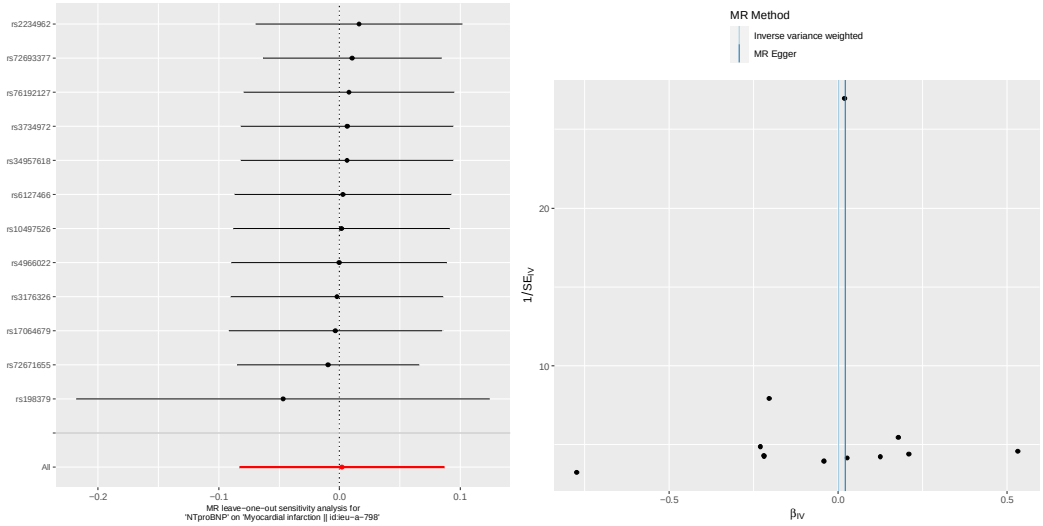


Figure S11 Mendelian randomization analysis of myocardial infarction and NT-proBNP levels. Mendelian randomisation for myocardial infarction (MI) as an exposure **a)-d)** and outcome **e)-h)** for NT-proBNP levels. The plots show summary information on the analyses, performed as per the TwoSampleMR R package. **a)** Mendelian randomization scatter plot for MI as an exposure for NT-proBNP. **b)** Mendelian

randomization single SNP funnel plot for MI as an exposure for NT-proBNP. **c)** Mendelian randomization single SNP forest plot for MI as an exposure for NT-proBNP. **d)** Mendelian randomization leave one out plot for MI as an exposure for NT-proBNP. **e)** Mendelian randomization scatter plot for NT-proBNP as an exposure for MI. **f)** Mendelian randomization single SNP funnel plot for NT-proBNP as an exposure for MI. **g)** Mendelian randomization single SNP forest plot for NT-proBNP as an exposure for MI. **h)** Mendelian randomization leave one out plot for NT-proBNP as an exposure for MI.

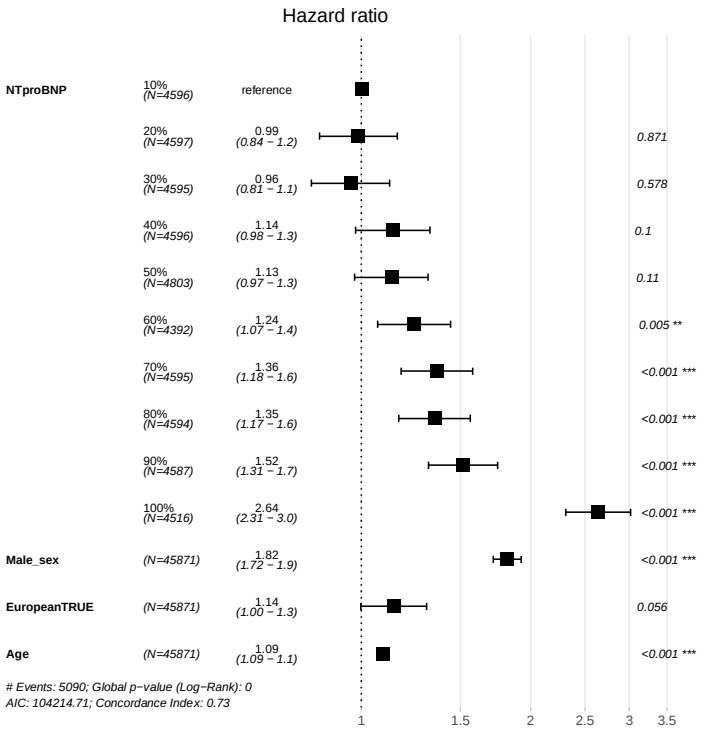


Figure S12 NT-proBNP increases with incident cardiomyopathy. Forest plot of cardiomyopathies by NT-proBNP deciles. The forest plots of Cox proportional hazards regression models were created assessing death or diagnosis from recruitment with those diagnosed before recruitment excluded. Sex (increasing risk is male), European ancestry (increasing risk is European), and age at recruitment, were added to this multivariable analysis.

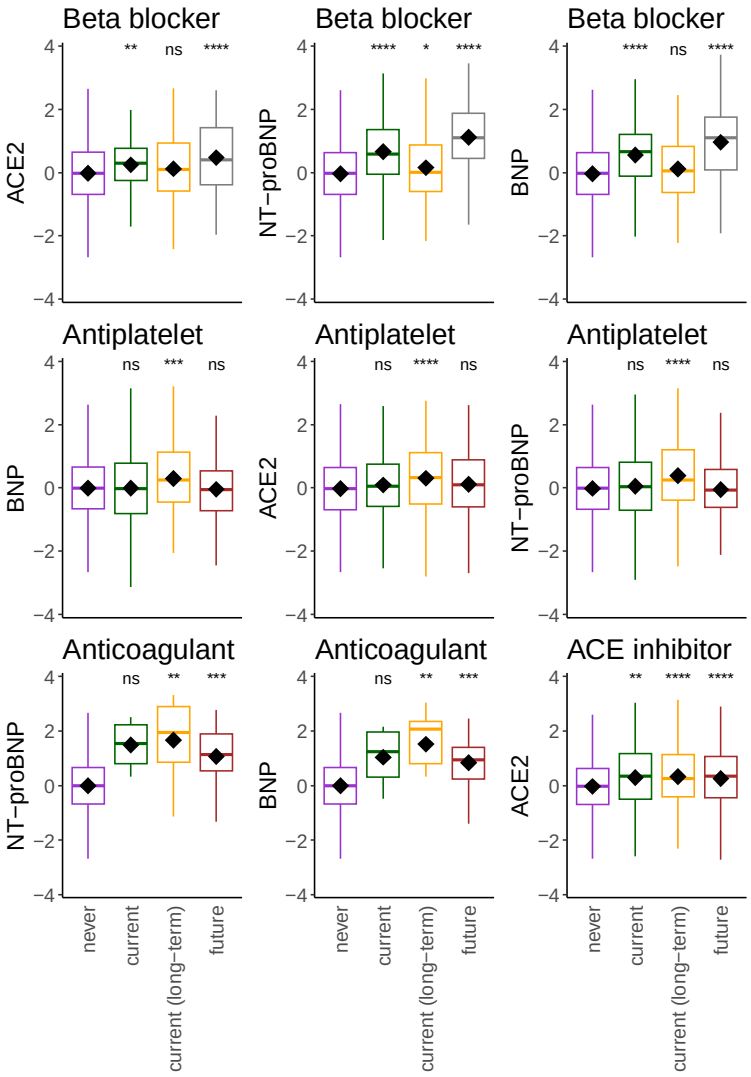


Figure S13 Relationships with medication intake adjusted for time between recruitment and imaging visit. The plots depict the proteins measured at recruitment and adjusted for days to the imaging visit that were significantly increased with medication intake reported only at recruitment (current), or the imaging visit on average 8 years later (future), or reported during both visits (current (long-term)). The significance of differences in means as derived by Student's t-test are denoted as stars compared to individuals with no report of the medication (never). The data only includes those with proteomics who attended the imaging visit (n=5,324).

a) **b)**

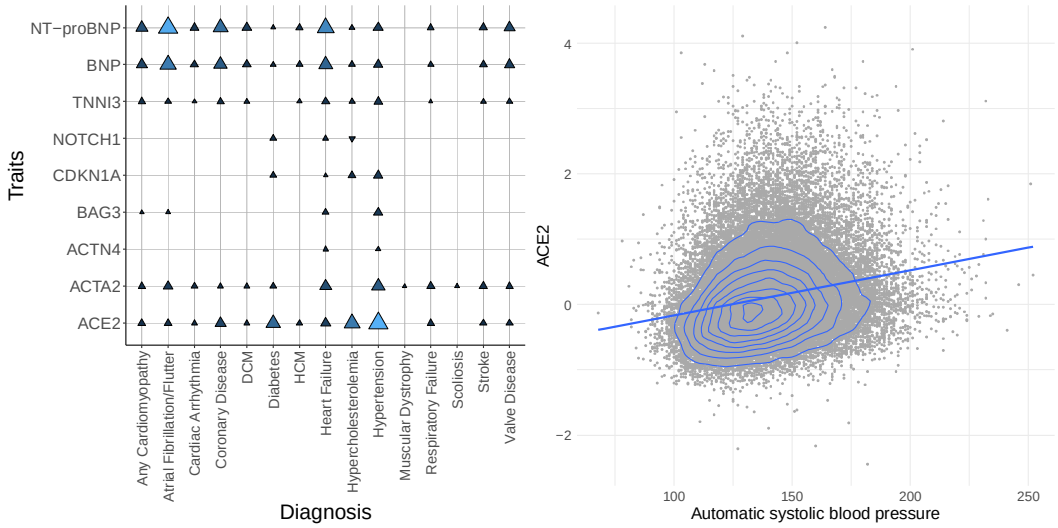
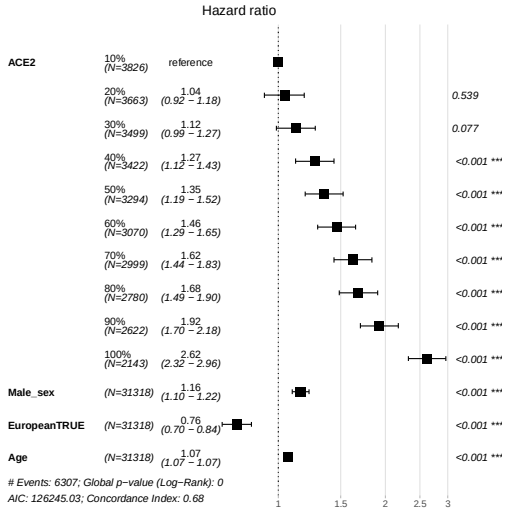
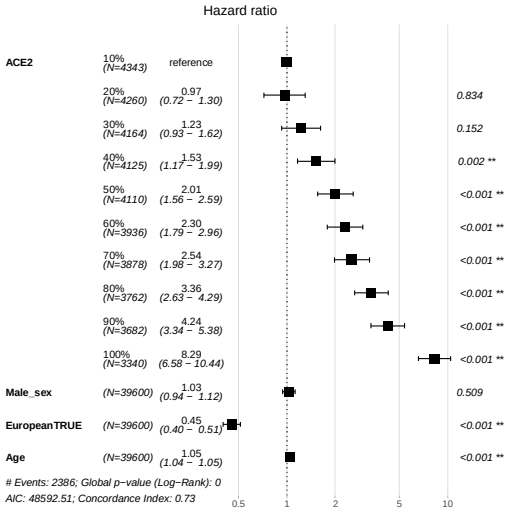


Figure S14 Curated cardiovascular disease association study and ACE2's relationship with systolic blood pressure. **a)** ICD codes as diagnoses are described on the x-axis and the protein traits are on the y-axis. Each point denotes a significant association. The shape and colour denote the direction of effect and odds ratio. The size represents the significance (where larger is more significant). See **Table S2** for the full results. DCM, dilated cardiomyopathy; HCM, hypertrophic cardiomyopathy. **b)** The linear relationship between ACE2 levels and an automatic measure of systolic blood pressure measured at recruitment (R=0.21).

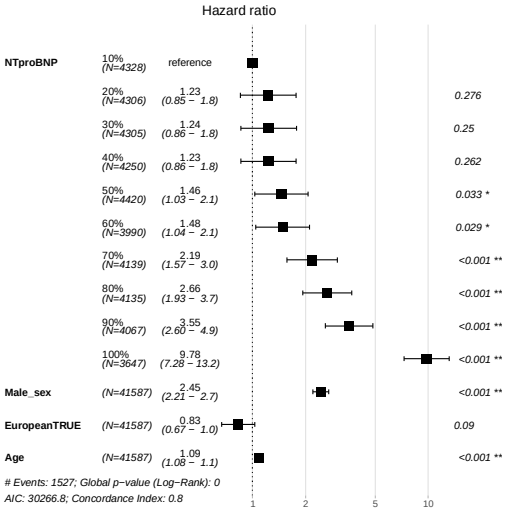
a)


Hazard ratio
ACE2
10% (N=3826) reference
20% (N=3663) 1.04 (0.92 − 1.18) 0.539
30% (N=3499) 1.12 (0.99 − 1.27) 0.077
40% (N=3422) 1.27 (1.12 − 1.43) <0.001 ***
50% (N=3294) 1.35 (1.19 − 1.52) <0.001 ***
60% (N=3070) 1.46 (1.29 − 1.65) <0.001 ***
70% (N=2999) 1.62 (1.44 − 1.83) <0.001 ***
80% (N=2780) 1.68 (1.49 − 1.90) <0.001 ***
90% (N=2622) 1.92 (1.70 − 2.18) <0.001 ***
100% (N=2143) 2.62 (2.32 − 2.96) <0.001 ***
Male_sex (N=31318) 1.16 (1.10 − 1.22) <0.001 ***
EuropeanTRUE (N=31318) 0.76 (0.70 − 0.84) <0.001 ***
Age (N=31318) 1.07 (1.07 − 1.07) <0.001 ***
Events: 6307; Global p−value (Log−Rank): 0
AIC: 126245.03; Concordance Index: 0.68
1
1.5
2
2.5
3

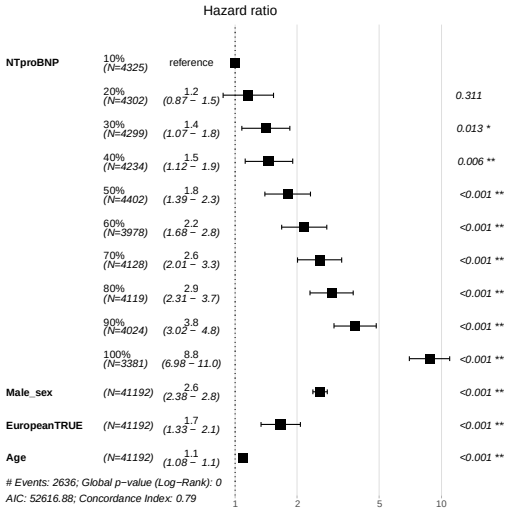
b)


Hazard ratio
ACE2
10% (N=4343) reference
20% (N=4260) 0.97 (0.72 − 1.30) 0.834
30% (N=4164) 1.23 (0.93 − 1.62) 0.152
40% (N=4125) 1.53 (1.17 − 1.99) 0.002 **
50% (N=4110) 2.01 (1.56 − 2.59) <0.001 **
60% (N=3936) 2.30 (1.79 − 2.96) <0.001 **
70% (N=3878) 2.54 (1.98 − 3.27) <0.001 **
80% (N=3762) 3.36 (2.63 − 4.29) <0.001 **
90% (N=3682) 4.24 (3.34 − 5.38) <0.001 **
100% (N=3340) 8.29 (6.58 − 10.44) <0.001 **
Male_sex (N=39600) 1.03 (0.94 − 1.12) 0.509
EuropeanTRUE (N=39600) 0.45 (0.40 − 0.51) <0.001 **
Age (N=39600) 1.05 (1.04 − 1.05) <0.001 **
Events: 2386; Global p−value (Log−Rank): 0
AIC: 48592.51; Concordance Index: 0.73
0.5
1
2
5
10

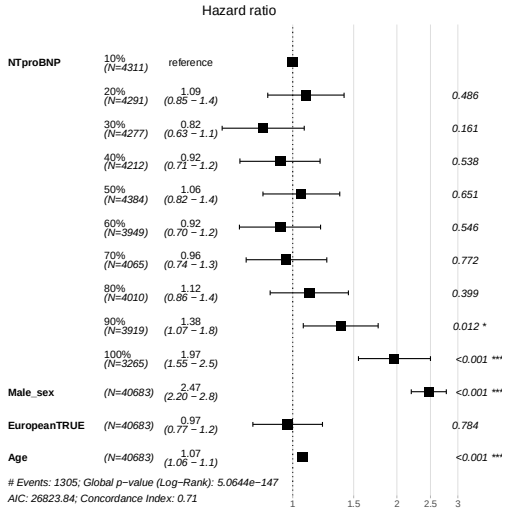
c)


Hazard ratio
NTproBNP
10% (N=4328) reference
20% (N=4306) 1.23 (0.85 − 1.8) 0.276
30% (N=4305) 1.24 (0.86 − 1.8) 0.25
40% (N=4250) 1.23 (0.86 − 1.8) 0.262
50% (N=4420) 1.46 (1.03 − 2.1) 0.033 *
60% (N=3990) 1.48 (1.04 − 2.1) 0.029 *
70% (N=4139) 2.19 (1.57 − 3.0) <0.001 **
80% (N=4135) 2.66 (1.93 − 3.7) <0.001 **
90% (N=4067) 3.55 (2.60 − 4.9) <0.001 **
100% (N=3647) 9.78 (7.28 − 13.2) <0.001 **
Male_sex (N=41587) 2.45 (2.21 − 2.7) <0.001 **
EuropeanTRUE (N=41587) 0.83 (0.67 − 1.0) 0.09
Age (N=41587) 1.09 (1.08 − 1.1) <0.001 **
Events: 1527; Global p−value (Log−Rank): 0
AIC: 30266.8; Concordance Index: 0.8
1
2
5
10

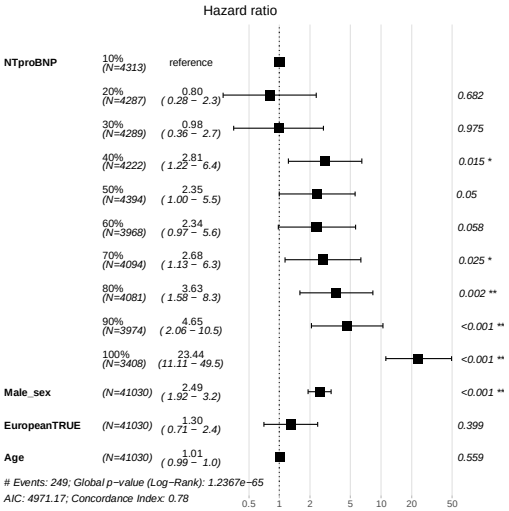
d)


Hazard ratio
NTproBNP
10% (N=4325) reference
20% (N=4302) 1.2 (0.87 − 1.5) 0.311
30% (N=4299) 1.4 (1.07 − 1.8) 0.013 *
40% (N=4234) 1.5 (1.12 − 1.9) 0.006 **
50% (N=4402) 1.8 (1.39 − 2.3) <0.001 **
60% (N=3978) 2.2 (1.68 − 2.8) <0.001 **
70% (N=4128) 2.6 (2.01 − 3.3) <0.001 **
80% (N=4119) 2.9 (2.31 − 3.7) <0.001 **
90% (N=4024) 3.8 (3.02 − 4.8) <0.001 **
100% (N=3381) 8.8 (6.98 − 11.0) <0.001 **
Male_sex (N=41192) 2.6 (2.38 − 2.8) <0.001 **
EuropeanTRUE (N=41192) 1.7 (1.33 − 2.1) <0.001 **
Age (N=41192) 1.1 (1.08 − 1.1) <0.001 **
Events: 2636; Global p−value (Log−Rank): 0
AIC: 52616.88; Concordance Index: 0.79
1
2
5
10

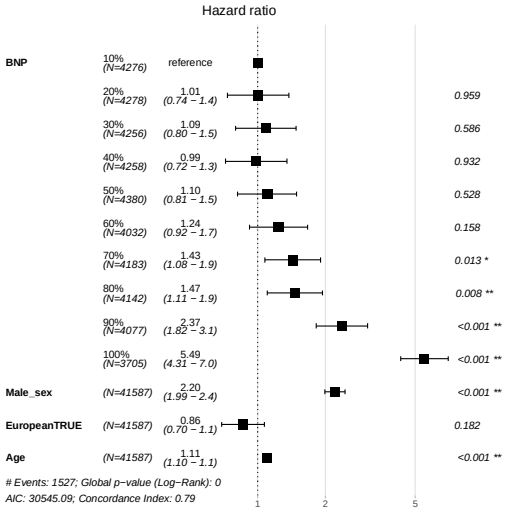
e)


Hazard ratio
NTproBNP
10% (N=4311) reference
20% (N=4291) 1.09 (0.85 − 1.4) 0.486
30% (N=4277) 0.82 (0.63 − 1.1) 0.161
40% (N=4212) 0.92 (0.71 − 1.2) 0.538
50% (N=4384) 1.06 (0.82 − 1.4) 0.651
60% (N=3949) 0.92 (0.70 − 1.2) 0.546
70% (N=4065) 0.96 (0.74 − 1.3) 0.772
80% (N=4010) 1.12 (0.86 − 1.4) 0.399
90% (N=3919) 1.38 (1.07 − 1.8) 0.012 *
100% (N=3265) 1.97 (1.55 − 2.5) <0.001 ***
Male_sex (N=40683) 2.47 (2.20 − 2.8) <0.001 ***
EuropeanTRUE (N=40683) 0.97 (0.77 − 1.2) 0.784
Age (N=40683) 1.07 (1.06 − 1.1) <0.001 ***
Events: 1305; Global p−value (Log−Rank): 5.0644e−147
AIC: 26823.84; Concordance Index: 0.71
1
1.5
2
2.5
3

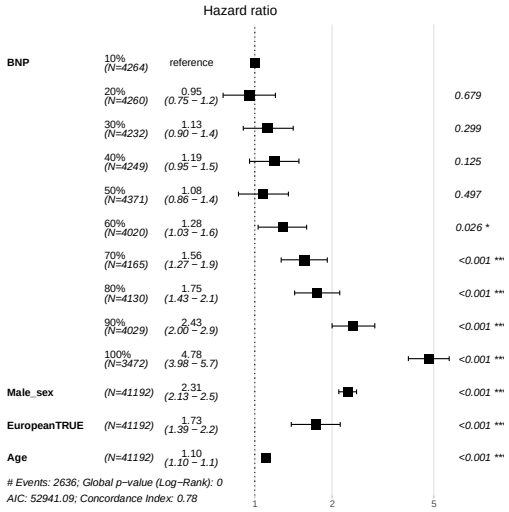
f)


Hazard ratio
NTproBNP
10% (N=4313) reference
20% (N=4287) 0.80 (0.28 − 2.3) 0.682
30% (N=4289) 0.98 (0.36 − 2.7) 0.975
40% (N=4222) 2.81 (1.22 − 6.4) 0.015 *
50% (N=4394) 2.35 (1.00 − 5.5) 0.05
60% (N=3968) 2.34 (0.97 − 5.6) 0.058
70% (N=4094) 2.68 (1.13 − 6.3) 0.025 *
80% (N=4081) 3.63 (1.58 − 8.3) 0.002 **
90% (N=3974) 4.65 (2.06 − 10.5) <0.001 **
100% (N=3408) 23.44 (11.11 − 49.5) <0.001 **
Male_sex (N=41030) 2.49 (1.92 − 3.2) <0.001 **
EuropeanTRUE (N=41030) 1.30 (0.71 − 2.4) 0.399
Age (N=41030) 1.01 (0.99 − 1.0) 0.559
Events: 249; Global p−value (Log−Rank): 1.2367e−65
AIC: 4971.17; Concordance Index: 0.78
0.5
1
2
5
10
20
50

g)


Hazard ratio
BNP
10% (N=4276) reference
20% (N=4278) 1.01 (0.74 − 1.4) 0.959
30% (N=4256) 1.09 (0.80 − 1.5) 0.586
40% (N=4258) 0.99 (0.72 − 1.3) 0.932
50% (N=4380) 1.10 (0.81 − 1.5) 0.528
60% (N=4032) 1.24 (0.92 − 1.7) 0.158
70% (N=4183) 1.43 (1.08 − 1.9) 0.013 *
80% (N=4142) 1.47 (1.11 − 1.9) 0.008 **
90% (N=4077) 2.37 (1.82 − 3.1) <0.001 **
100% (N=3705) 5.49 (4.31 − 7.0) <0.001 **
Male_sex (N=41587) 2.20 (1.99 − 2.4) <0.001 **
EuropeanTRUE (N=41587) 0.86 (0.70 − 1.1) 0.182
Age (N=41587) 1.11 (1.10 − 1.1) <0.001 **
Events: 1527; Global p−value (Log−Rank): 0
AIC: 30545.09; Concordance Index: 0.79
1
2
5

h)


Hazard ratio
BNP
10% (N=4264) reference
20% (N=4260) 0.95 (0.75 − 1.2) 0.679
30% (N=4232) 1.13 (0.90 − 1.4) 0.299
40% (N=4249) 1.19 (0.95 − 1.5) 0.125
50% (N=4371) 1.08 (0.86 − 1.4) 0.497
60% (N=4020) 1.28 (1.03 − 1.6) 0.026 *
70% (N=4165) 1.56 (1.27 − 1.9) <0.001 ***
80% (N=4130) 1.75 (1.43 − 2.1) <0.001 ***
90% (N=4029) 2.43 (2.00 − 2.9) <0.001 ***
100% (N=3472) 4.78 (3.98 − 5.7) <0.001 ***
Male_sex (N=41192) 2.31 (2.13 − 2.5) <0.001 ***
EuropeanTRUE (N=41192) 1.73 (1.39 − 2.2) <0.001 ***
Age (N=41192) 1.10 (1.10 − 1.1) <0.001 ***
Events: 2636; Global p−value (Log−Rank): 0
AIC: 52941.09; Concordance Index: 0.78
1
2
5

i) **j)**

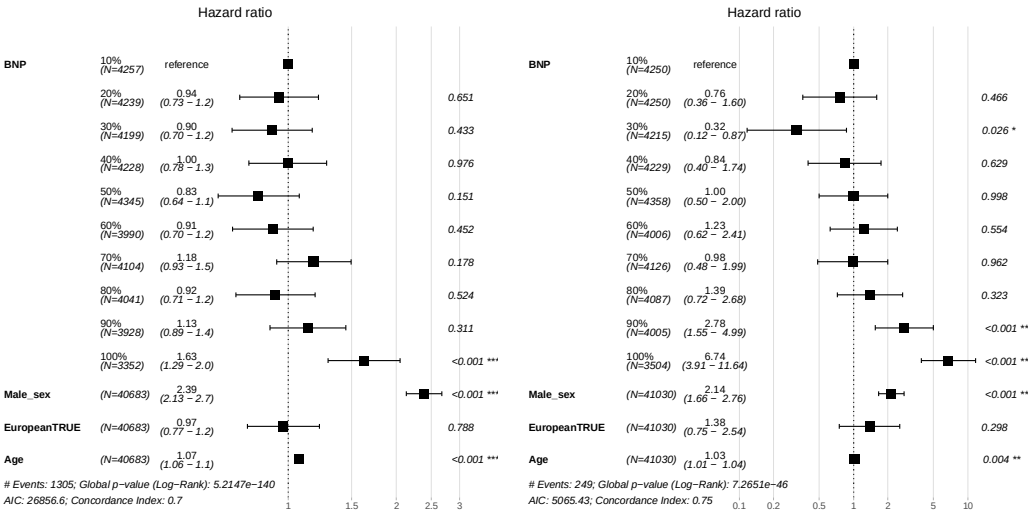


Figure S15 Sensitivity analysis of removing all-cause death from Cox proportional hazards regression models. Deciles of ACE2 levels with incident **a)** hypertension and **b)** diabetes, since recruitment. Deciles of NT-proBNP levels with incident **c)** heart failure, **d)** atrial fibrillation, **e)** myocardial infarction, and **f)** cardiomyopathy, since recruitment. Deciles of BNP levels with incident **g)** heart failure, **h)** atrial fibrillation, **i)** myocardial infarction, and **j)** cardiomyopathy, since recruitment. The forest plots were created assessing diagnosis from recruitment with those diagnosed before recruitment or died without a diagnosis excluded. Sex (increasing risk is male), European ancestry (increasing risk is European), and age at recruitment, were added to this multivariable analysis for comparison.

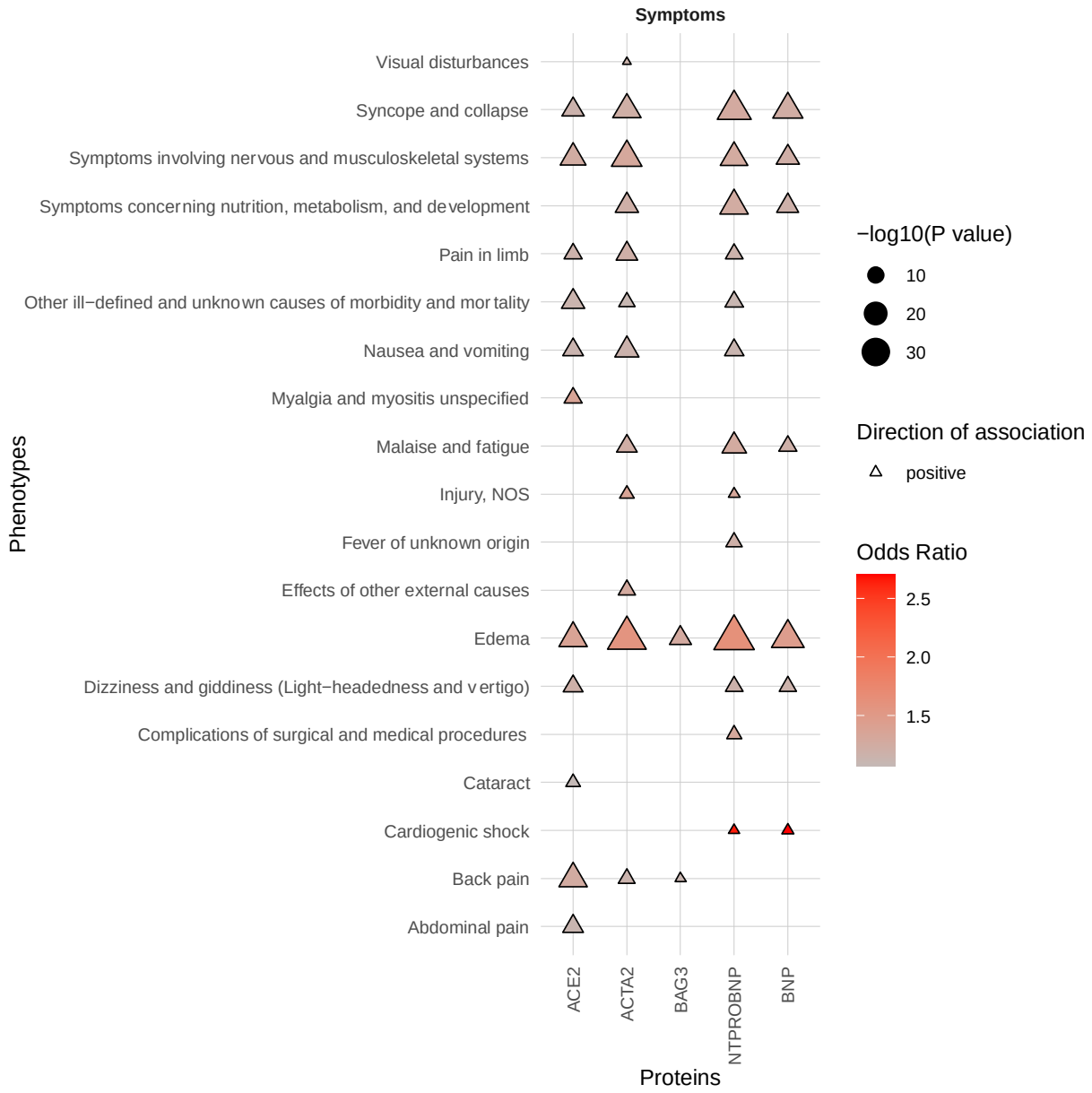


Figure S16 Phenome-wide association study results of the plasma protein levels with the symptoms category of phenotypes.

Phenotypes as phecodes are described on the y-axis and the protein traits on the x-axis. Each point denotes a significant PheWAS association with a Bonferroni correction for the number of analyzed phecodes. The shape and colour denote the direction of effect and odds ratio. See **Table S1** for the full PheWAS results.

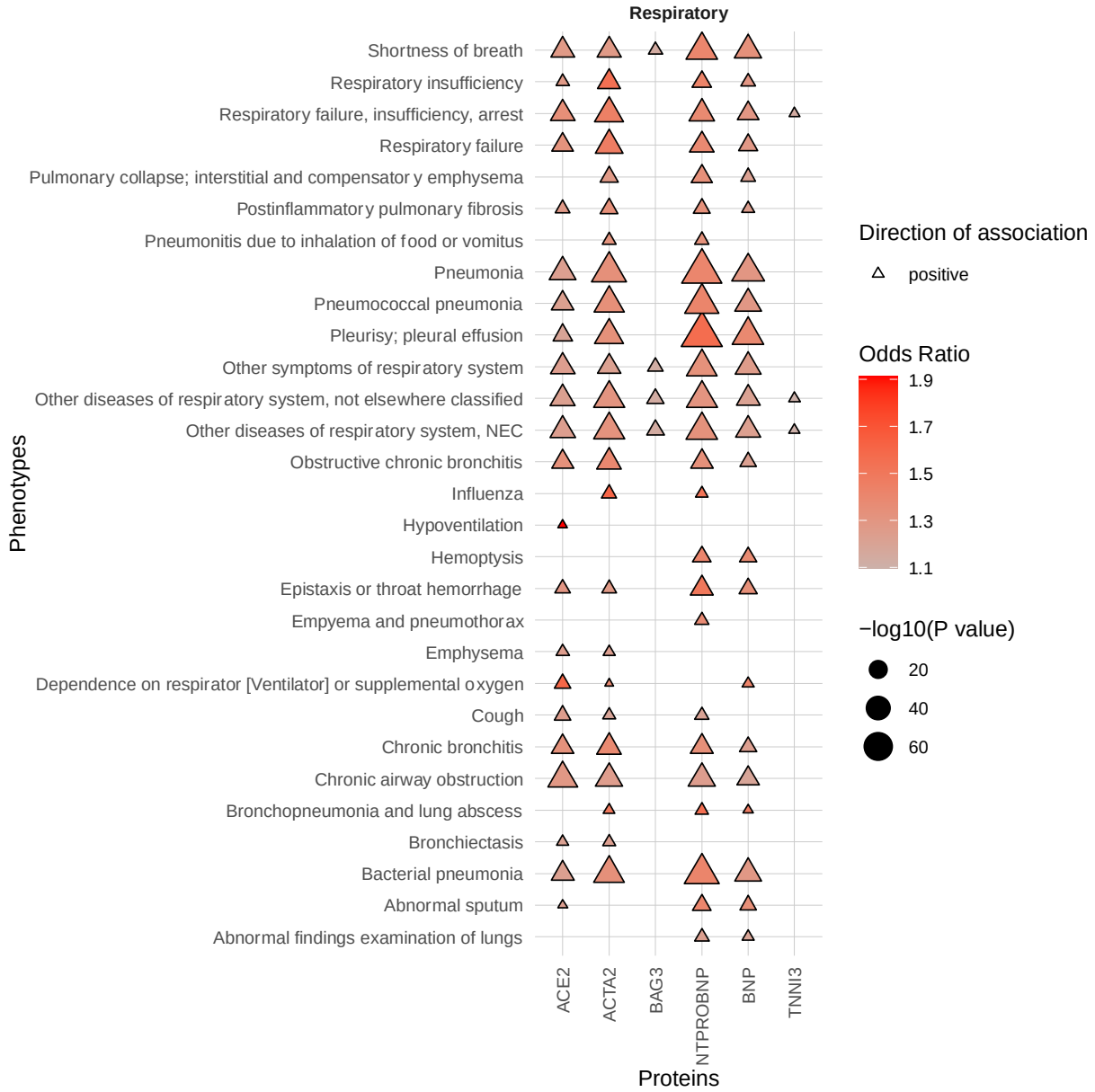


Figure S17 Phenome-wide association study results of the plasma protein levels with the respiratory category of phenotypes.

Phenotypes as phecodes are described on the y-axis and the protein traits on the x-axis. Each point denotes a significant PheWAS association with a Bonferroni correction for the number of analyzed phecodes. The shape and colour denote the direction of effect and odds ratio. See **Table S1** for the full PheWAS results.

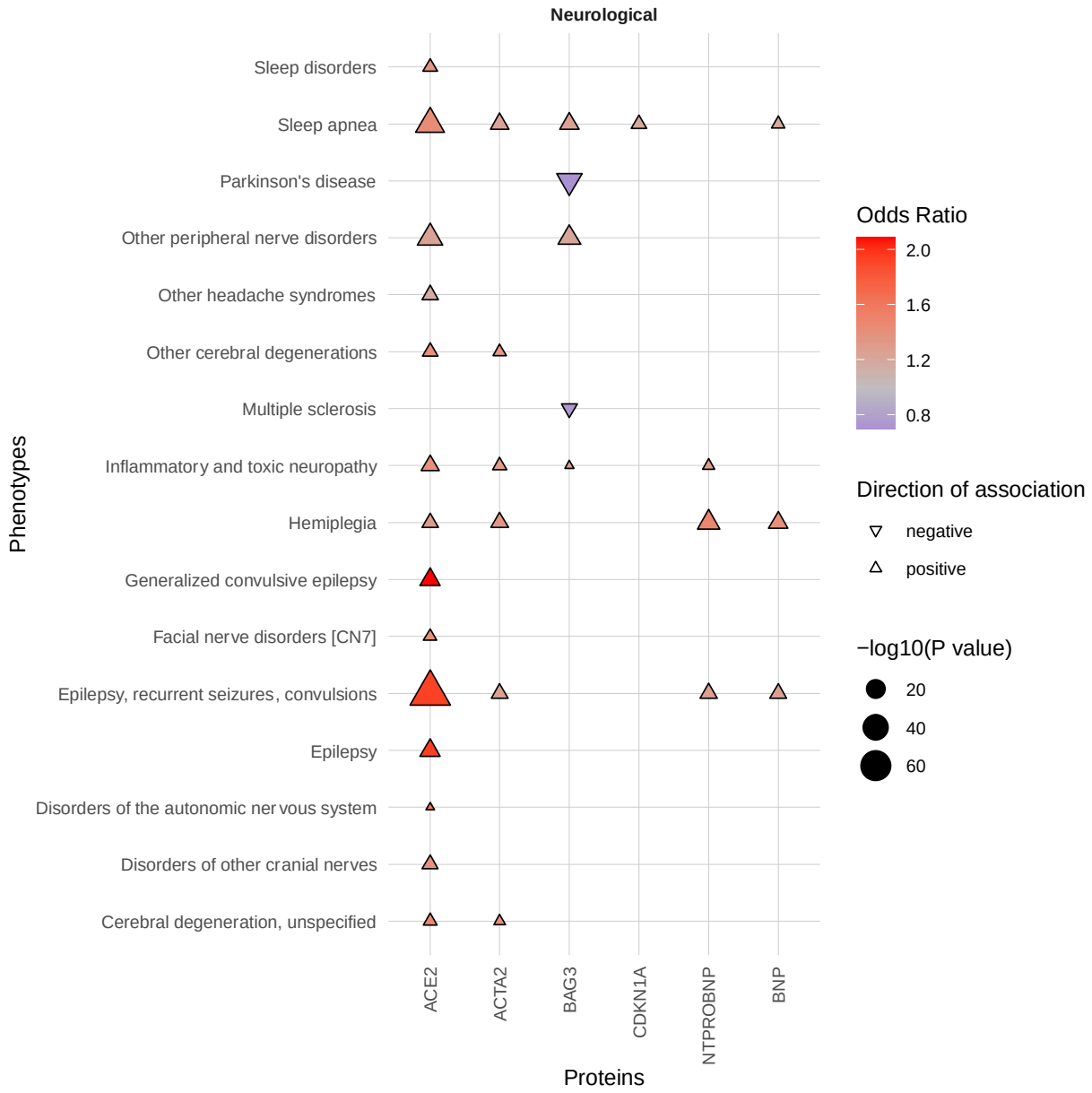


Figure S18 Phenome-wide association study results of the plasma protein levels with the neurological category of phenotypes.

Phenotypes as phecodes are described on the y-axis and the protein traits on the x-axis. Each point denotes a significant PheWAS association with a Bonferroni correction for the number of analyzed phecodes. The shape and colour denote the direction of effect and odds ratio. See **Table S1** for the full PheWAS results.

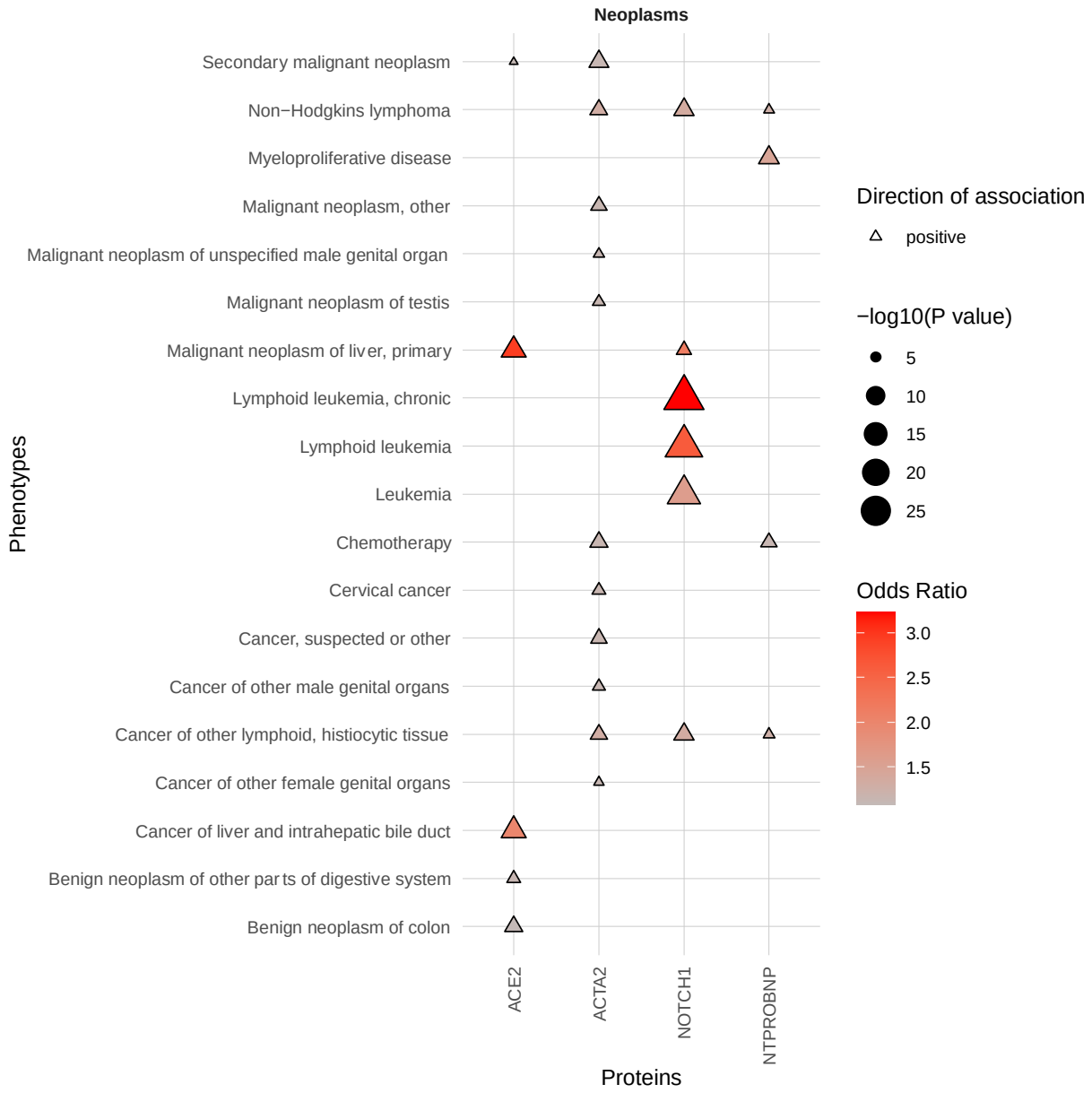


Figure S19 Phenome-wide association study results of the plasma protein levels with the neoplasms category of phenotypes.

Phenotypes as phecodes are described on the y-axis and the protein traits on the x-axis. Each point denotes a significant PheWAS association with a Bonferroni correction for the number of analyzed phecodes. The shape and colour denote the direction of effect and odds ratio. See **Table S1** for the full PheWAS results.

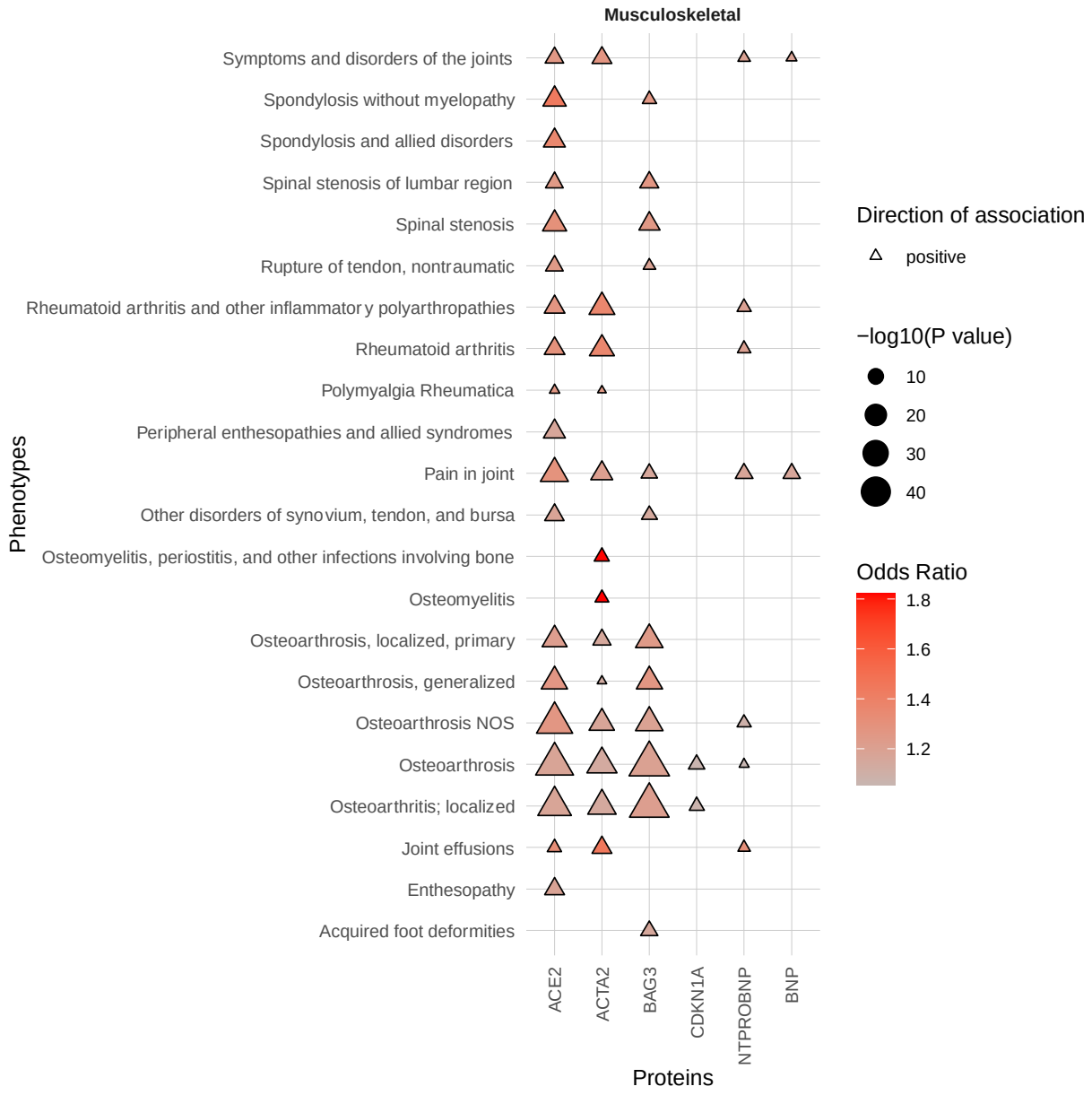


Figure S20 Phenome-wide association study results of the plasma protein levels with the musculoskeletal category of phenotypes.

Phenotypes as phecodes are described on the y-axis and the protein traits on the x-axis. Each point denotes a significant PheWAS association with a Bonferroni correction for the number of analyzed phecodes. The shape and colour denote the direction of effect and odds ratio. See **Table S1** for the full PheWAS results.

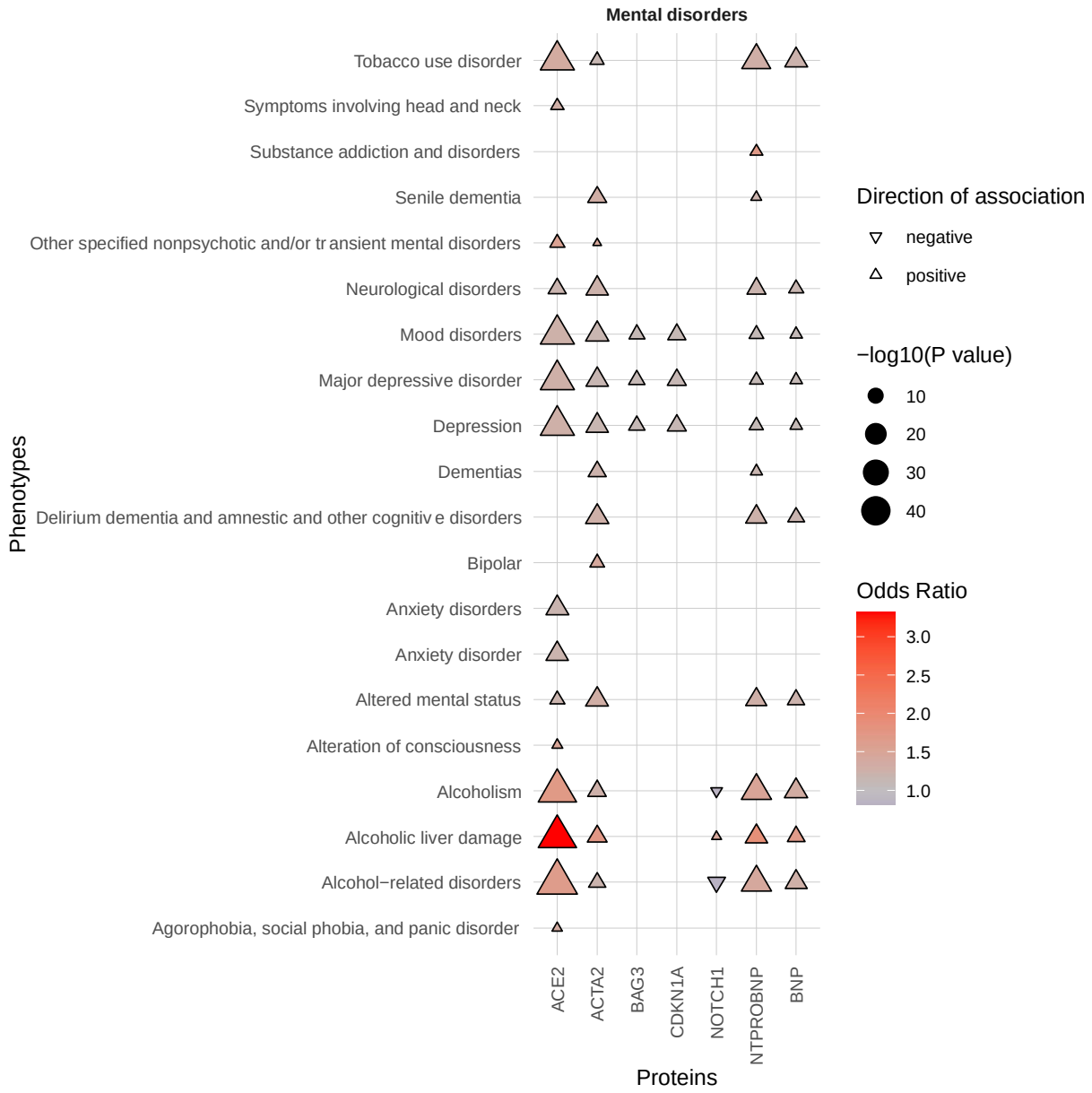


Figure S21 Phenome-wide association study results of the plasma protein levels with the mental disorders category of phenotypes.

Phenotypes as phecodes are described on the y-axis and the protein traits on the x-axis. Each point denotes a significant PheWAS association with a Bonferroni correction for the number of analyzed phecodes. The shape and colour denote the direction of effect and odds ratio. See **Table S1** for the full PheWAS results.

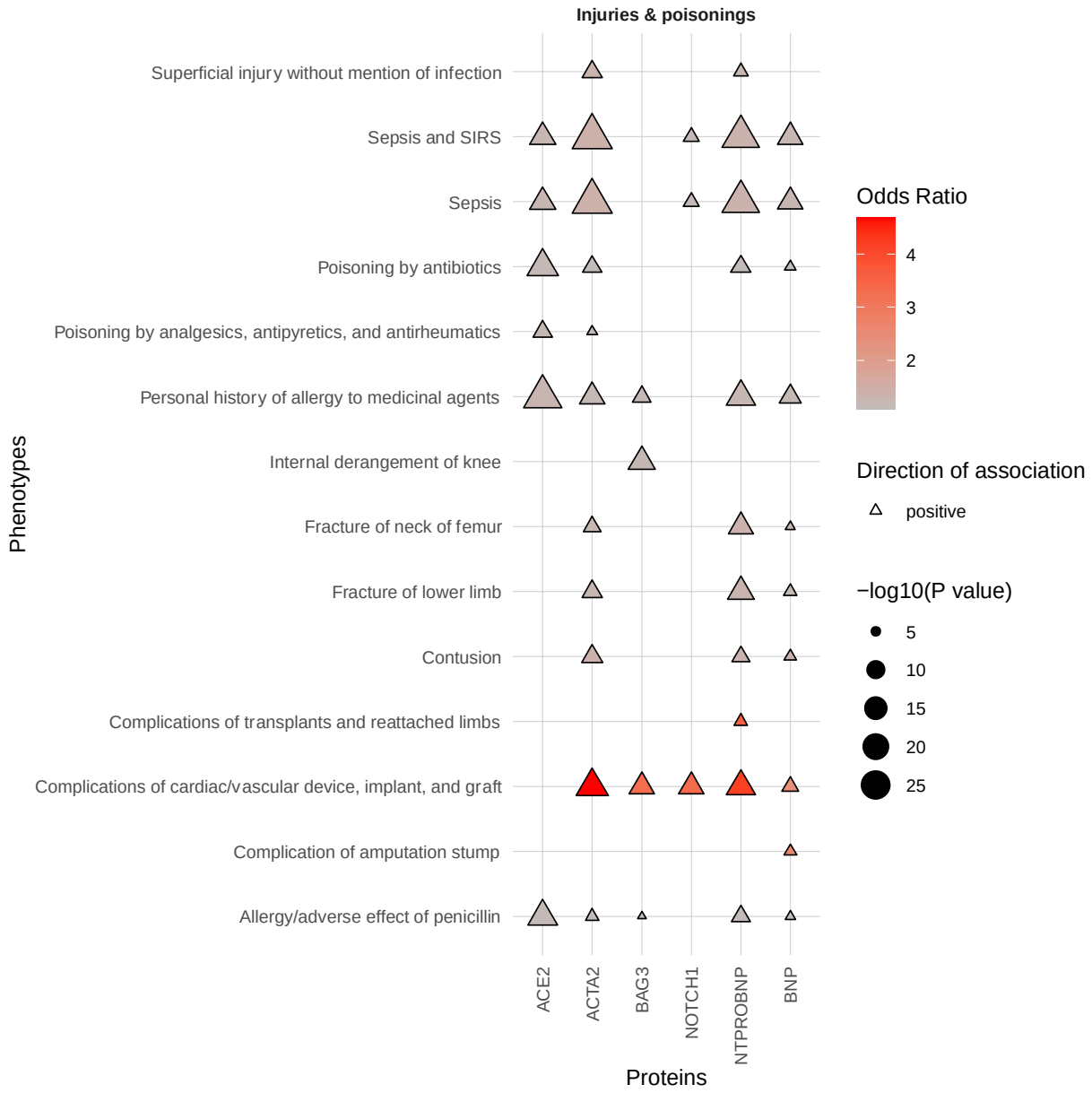


Figure S22 Phenome-wide association study results of the plasma protein levels with the injuries and poisonings category of phenotypes.

Phenotypes as phecodes are described on the y-axis and the protein traits on the x-axis. Each point denotes a significant PheWAS association with a Bonferroni correction for the number of analyzed phecodes. The shape and colour denote the direction of effect and odds ratio. See **Table S1** for the full PheWAS results.

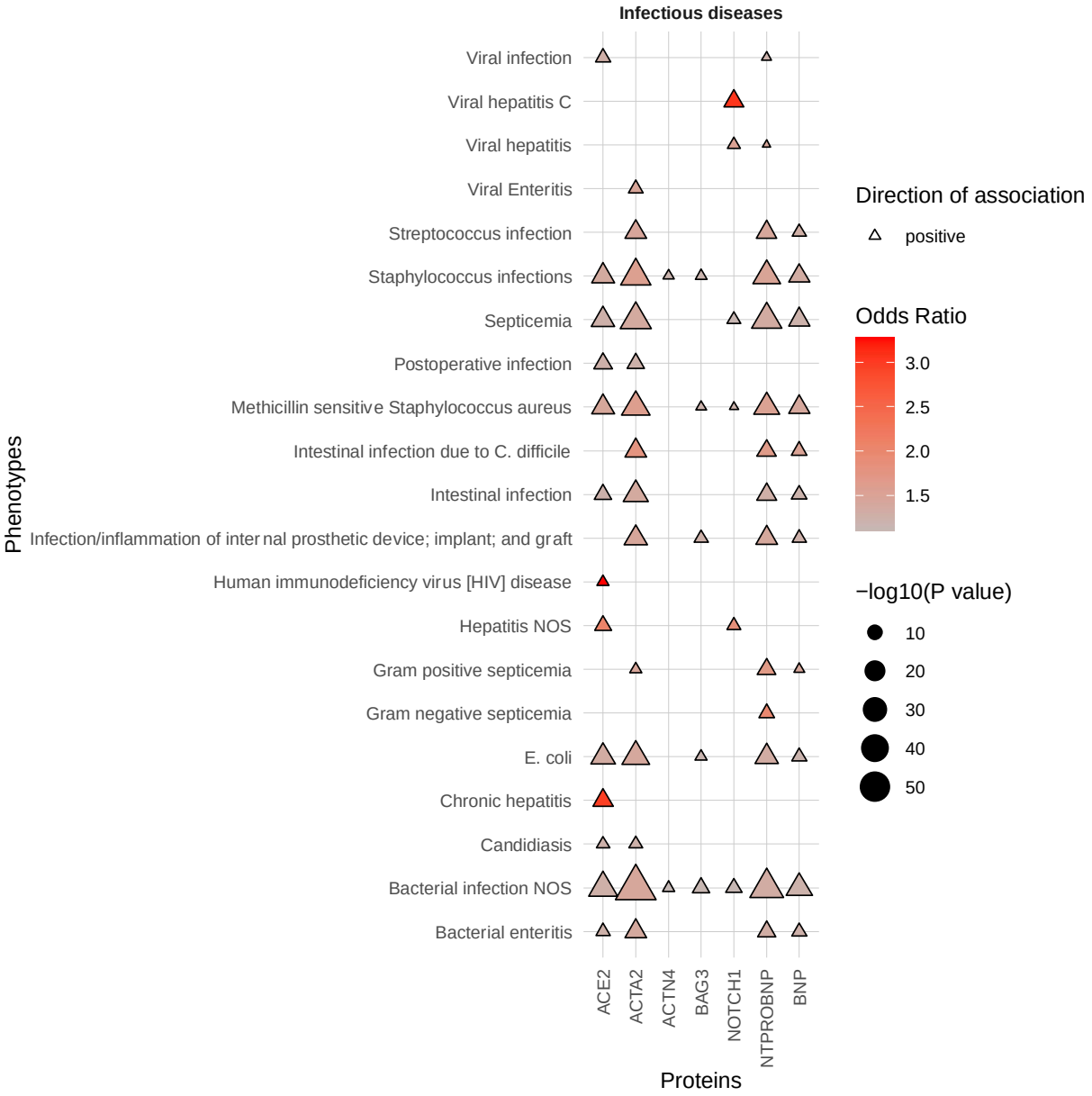


Figure S23 Phenome-wide association study results of the plasma protein levels with the infectious diseases category of phenotypes.

Phenotypes as phecodes are described on the y-axis and the protein traits on the x-axis. Each point denotes a significant PheWAS association with a Bonferroni correction for the number of analyzed phecodes. The shape and colour denote the direction of effect and odds ratio. See **Table S1** for the full PheWAS results.

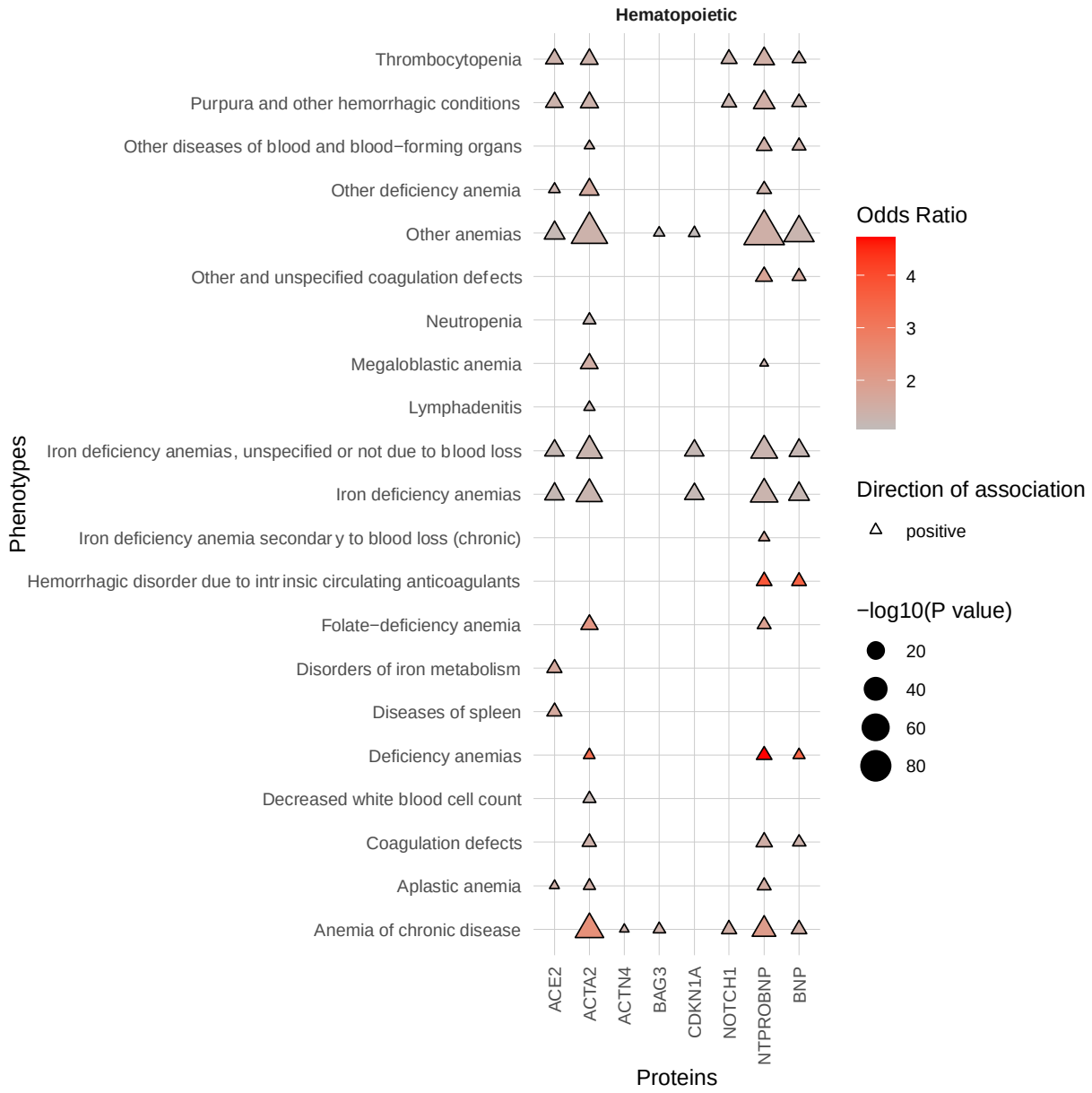


Figure S24 Phenome-wide association study results of the plasma protein levels with the hematopoietic category of phenotypes.

Phenotypes as phecodes are described on the y-axis and the protein traits on the x-axis. Each point denotes a significant PheWAS association with a Bonferroni correction for the number of analyzed phecodes. The shape and colour denote the direction of effect and odds ratio. See **Table S1** for the full PheWAS results.

Figure S25 Phenome-wide association study results of the plasma protein levels with the genitourinary category of phenotypes.

Phenotypes as phecodes are described on the y-axis and the protein traits on the x-axis. Each point denotes a significant PheWAS association with a Bonferroni correction for the number of analyzed phecodes. The shape and colour denote the direction of effect and odds ratio. See **Table S1** for the full PheWAS results.

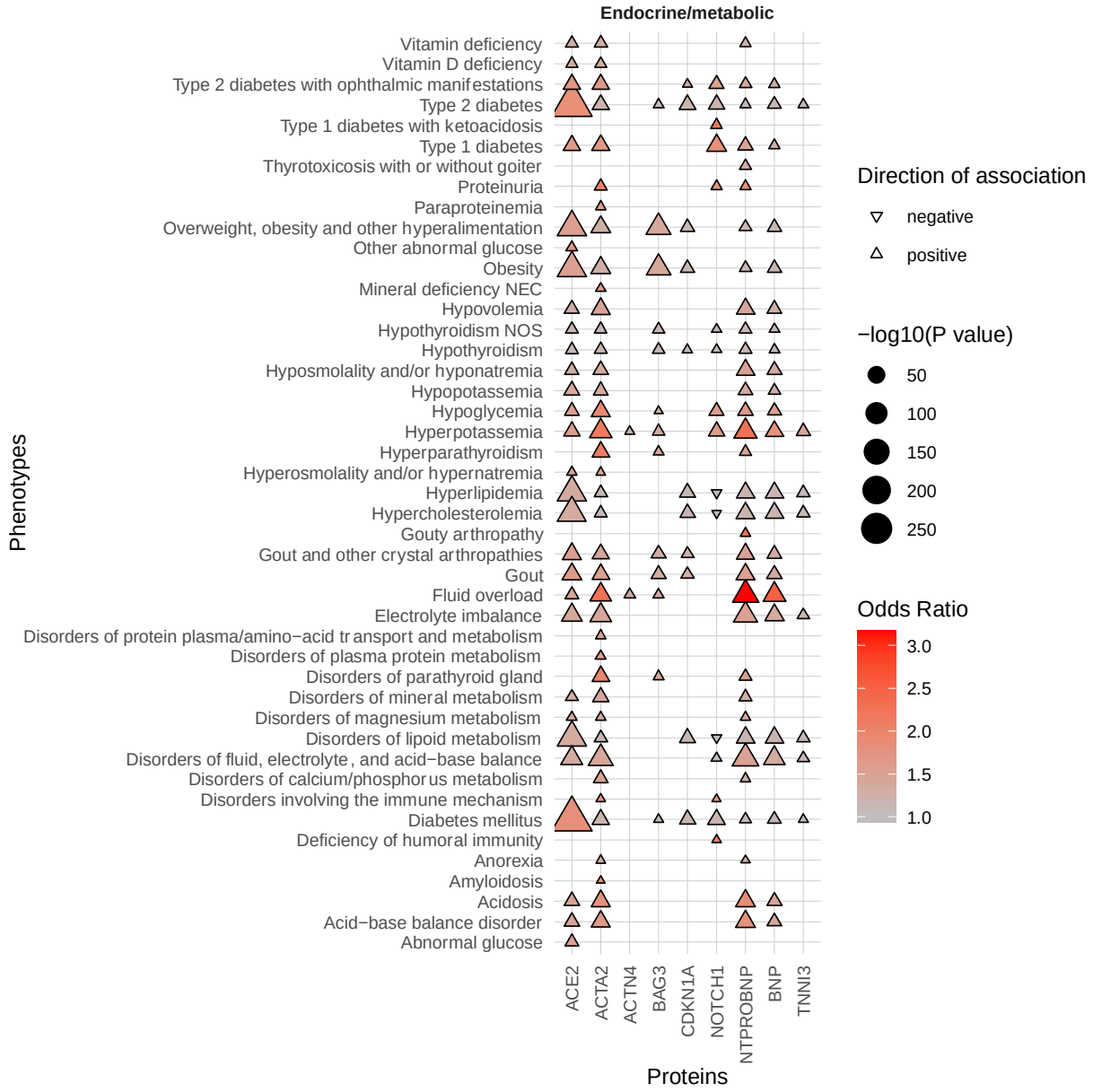


Figure S26 Phenome-wide association study results of the plasma protein levels with the endocrine and metabolic category of phenotypes.

Phenotypes as phecodes are described on the y-axis and the protein traits on the x-axis. Each point denotes a significant PheWAS association with a Bonferroni correction for the number of analyzed phecodes. The shape and colour denote the direction of effect and odds ratio. See **Table S1** for the full PheWAS results.

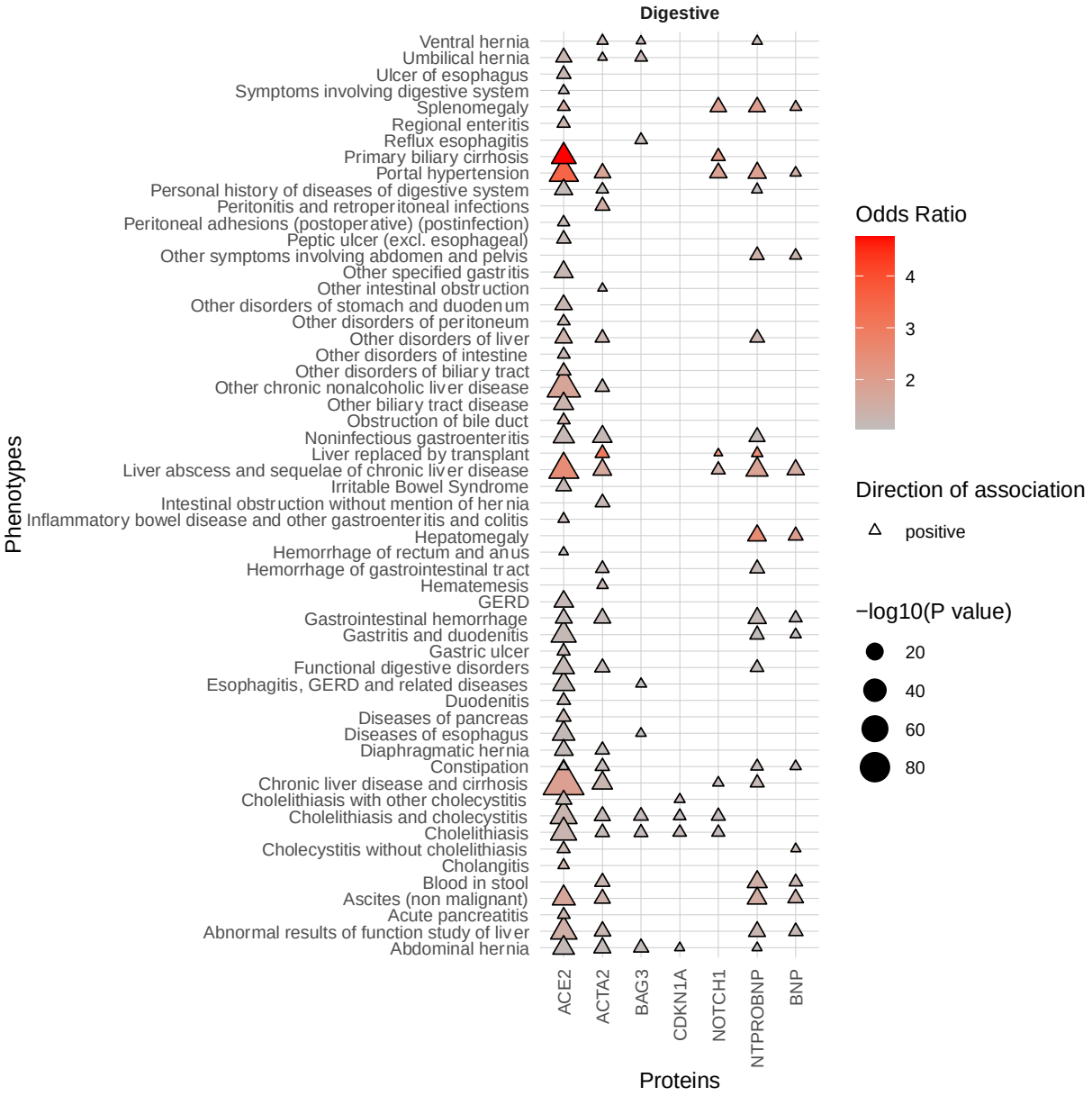


Figure S27 Phenome-wide association study results of the plasma protein levels with the digestive category of phenotypes.

Phenotypes as phecodes are described on the y-axis and the protein traits on the x-axis. Each point denotes a significant PheWAS association with a Bonferroni correction for the number of analyzed phecodes. The shape and colour denote the direction of effect and odds ratio. See **Table S1** for the full PheWAS results.

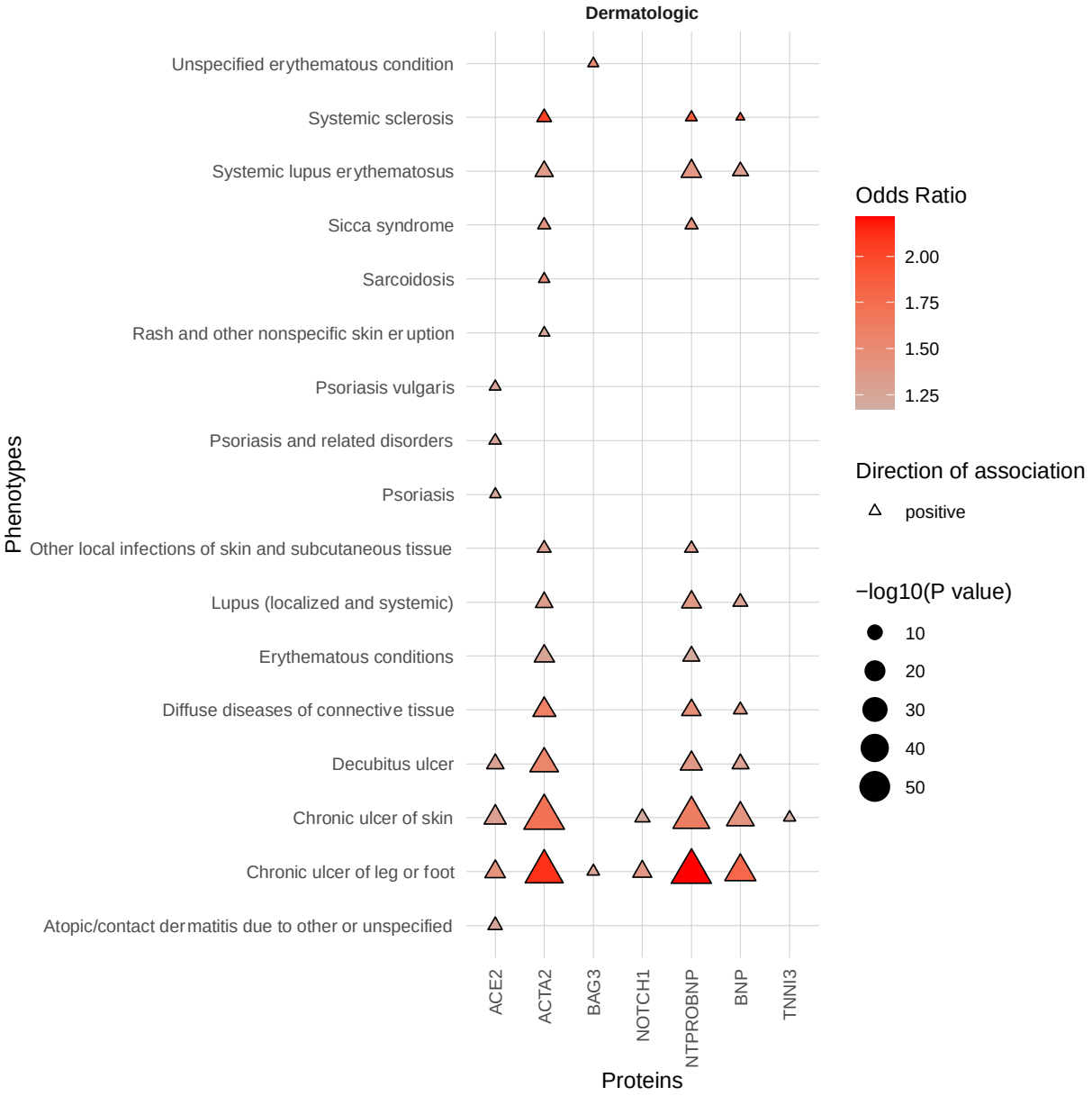


Figure S28 Phenome-wide association study results of the plasma protein levels with the dermatologic category of phenotypes.

Phenotypes as phecodes are described on the y-axis and the protein traits on the x-axis. Each point denotes a significant PheWAS association with a Bonferroni correction for the number of analyzed phecodes. The shape and colour denote the direction of effect and odds ratio. See **Table S1** for the full PheWAS results.

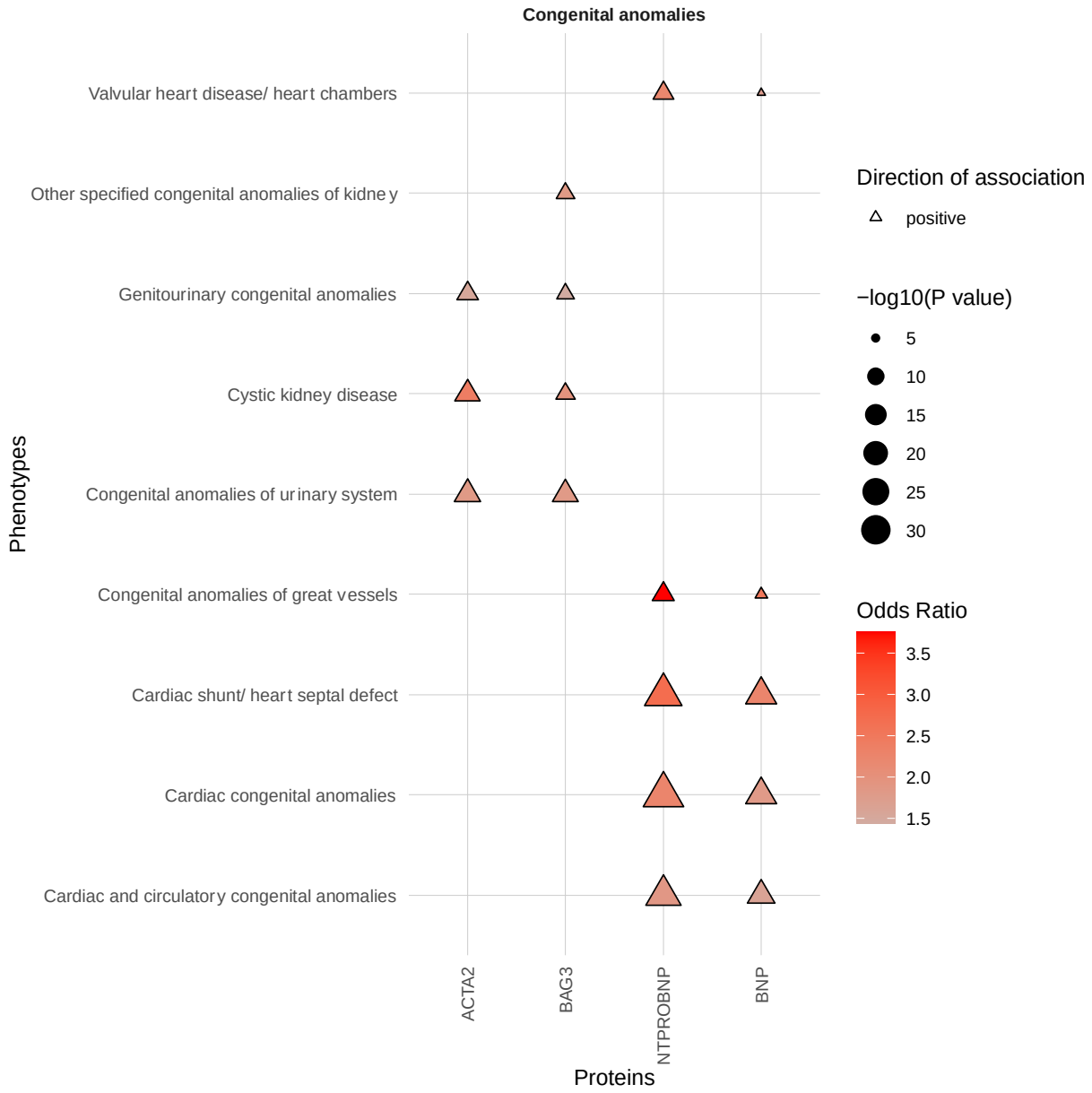
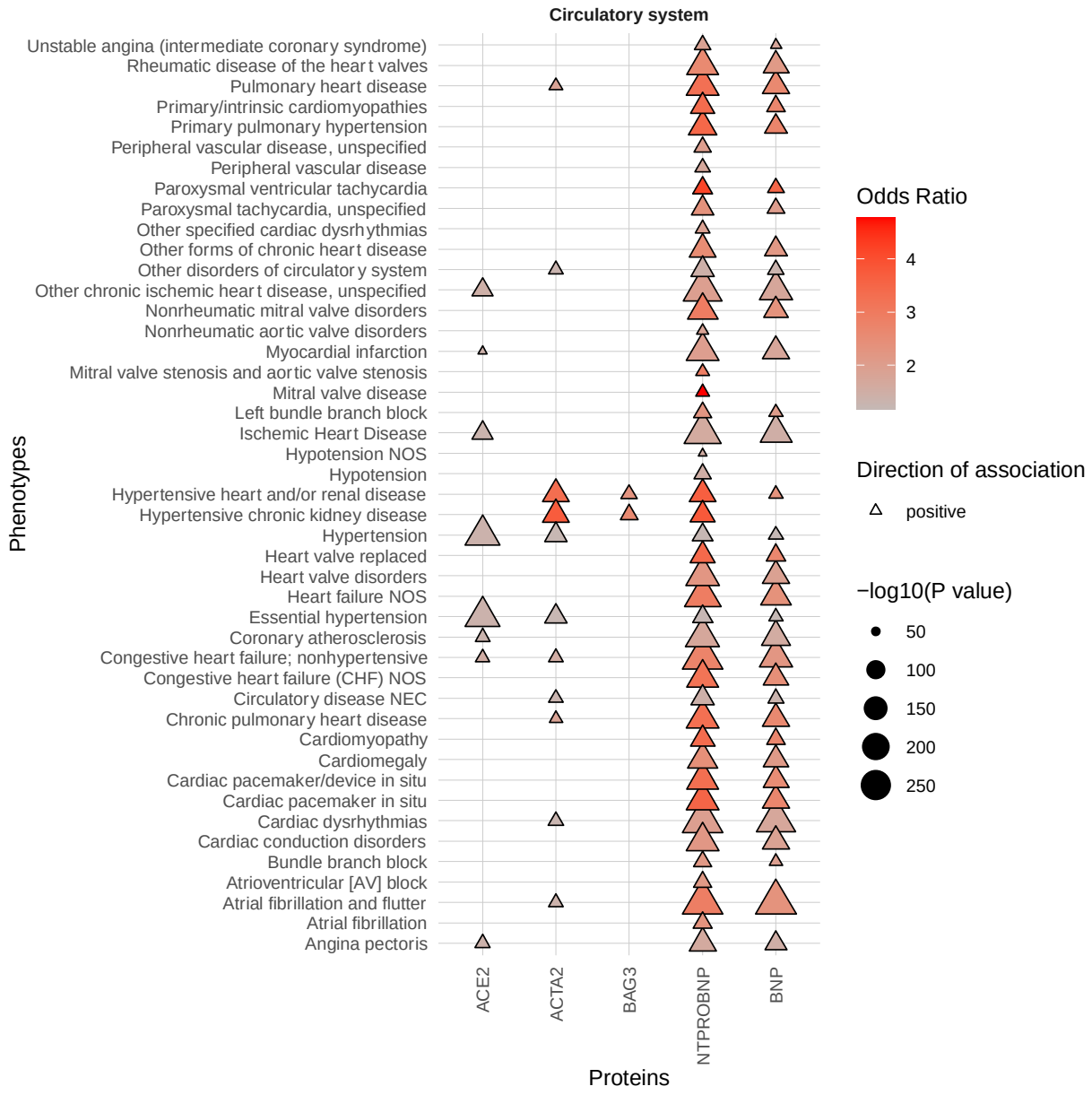


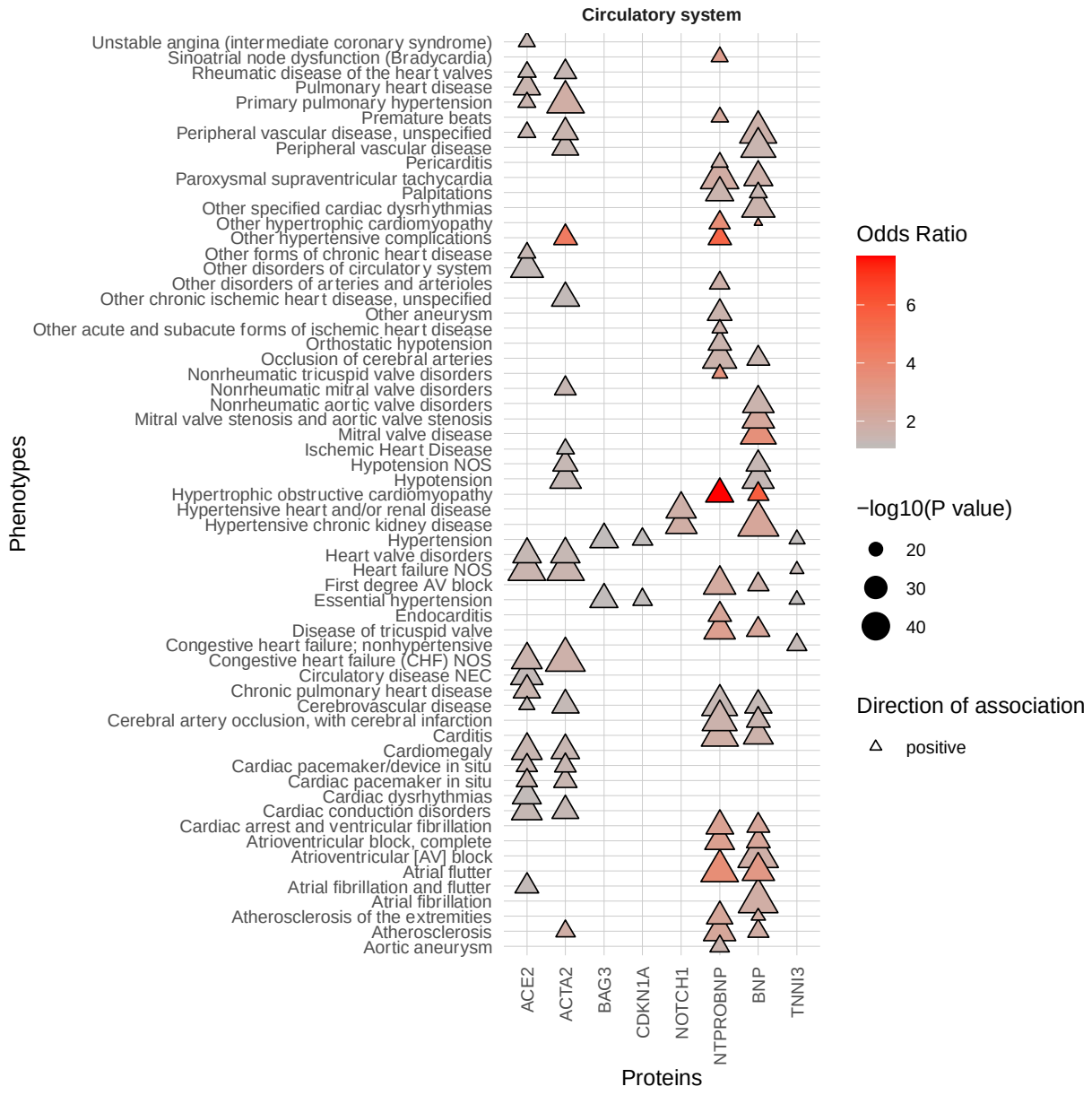
Figure S29 Phenome-wide association study results of the plasma protein levels with the congenital anomalies category of phenotypes.

Phenotypes as phecodes are described on the y-axis and the protein traits on the x-axis. Each point denotes a significant PheWAS association with a Bonferroni correction for the number of analyzed phecodes. The shape and colour denote the direction of effect and odds ratio. See **Table S1** for the full PheWAS results.

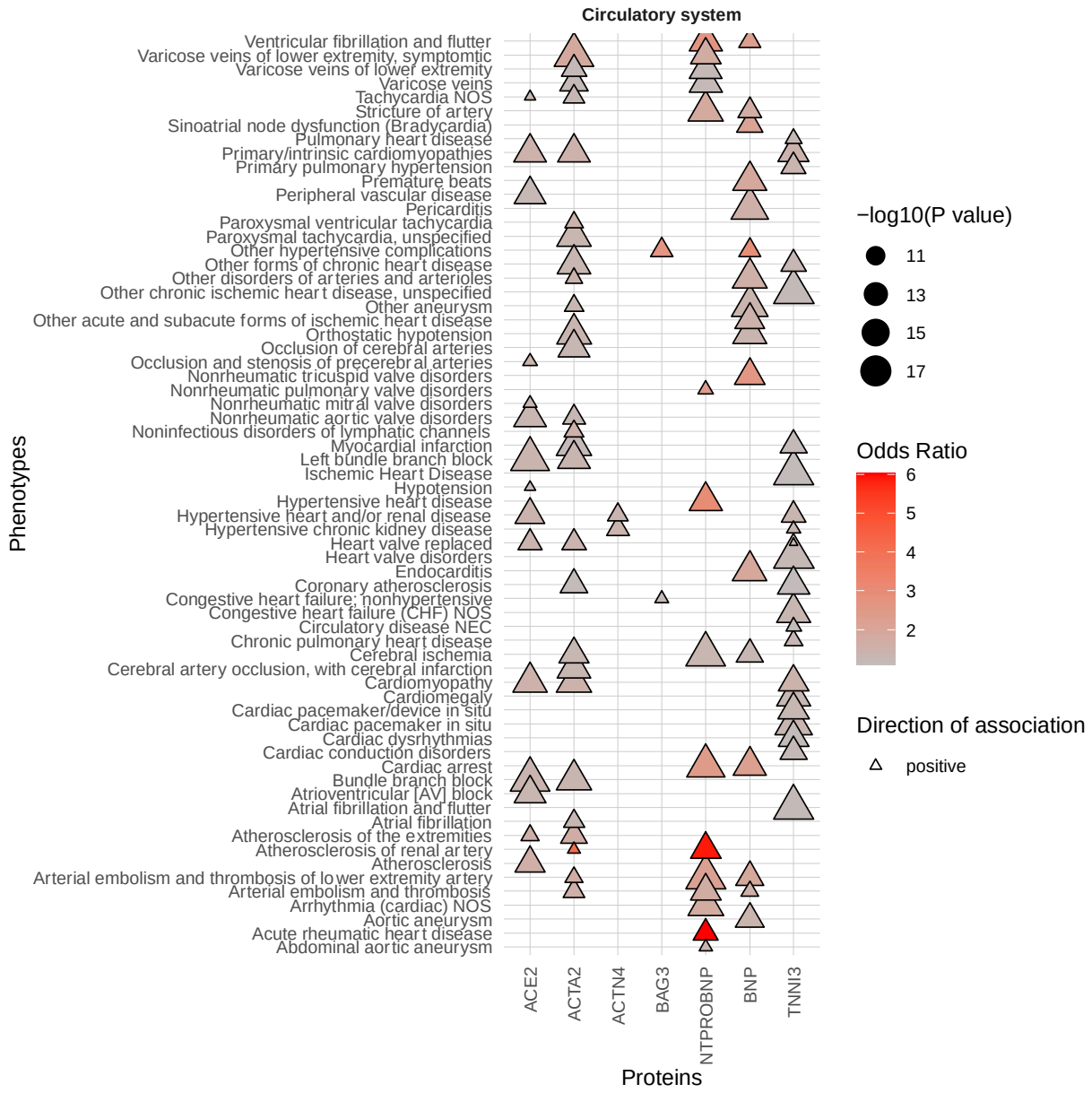
a)


Circulatory system
Unstable angina (intermediate coronary syndrome)
Rheumatic disease of the heart valves
Pulmonary heart disease
Primary/intrinsic cardiomyopathies
Primary pulmonary hypertension
Peripheral vascular disease, unspecified
Peripheral vascular disease
Paroxysmal ventricular tachycardia
Paroxysmal tachycardia, unspecified
Other specified cardiac dysrhythmias
Other forms of chronic heart disease
Other disorders of circulatory system
Other chronic ischemic heart disease, unspecified
Nonrheumatic mitral valve disorders
Nonrheumatic aortic valve disorders
Myocardial infarction
Mitral valve stenosis and aortic valve stenosis
Mitral valve disease
Left bundle branch block
Ischemic Heart Disease
Hypotension NOS
Hypotension
Hypertensive heart and/or renal disease
Hypertensive chronic kidney disease
Hypertension
Heart valve replaced
Heart valve disorders
Heart failure NOS
Essential hypertension
Coronary atherosclerosis
Congestive heart failure; nonhypertensive
Congestive heart failure (CHF) NOS
Circulatory disease NEC
Chronic pulmonary heart disease
Cardiomyopathy
Cardiomegaly
Cardiac pacemaker/device in situ
Cardiac pacemaker in situ
Cardiac dysrhythmias
Cardiac conduction disorders
Bundle branch block
Atrioventricular [AV] block
Atrial fibrillation and flutter
Atrial fibrillation
Angina pectoris
Phenotypes
ACE2
ACTA2
BAG3
NTPROBNP
BNP
Proteins
Odds Ratio
4
3
2
Direction of association
positive
−log10(P value)
50
100
150
200
250

b)


Circulatory system
Unstable angina (intermediate coronary syndrome)
Sinoatrial node dysfunction (Bradycardia)
Rheumatic disease of the heart valves
Pulmonary heart disease
Primary pulmonary hypertension
Premature beats
Peripheral vascular disease, unspecified
Peripheral vascular disease
Pericarditis
Paroxysmal supraventricular tachycardia
Palpitations
Other specified cardiac dysrhythmias
Other hypertrophic cardiomyopathy
Other hypertensive complications
Other forms of chronic heart disease
Other disorders of circulatory system
Other disorders of arteries and arterioles
Other chronic ischemic heart disease, unspecified
Other aneurysm
Other acute and subacute forms of ischemic heart disease
Orthostatic hypotension
Occlusion of cerebral arteries
Nonrheumatic tricuspid valve disorders
Nonrheumatic mitral valve disorders
Nonrheumatic aortic valve disorders
Mitral valve stenosis and aortic valve stenosis
Mitral valve disease
Ischemic Heart Disease
Hypotension NOS
Hypotension
Hypertrophic obstructive cardiomyopathy
Hypertensive heart and/or renal disease
Hypertensive chronic kidney disease
Hypertension
Heart valve disorders
Heart failure NOS
First degree AV block
Essential hypertension
Endocarditis
Disease of tricuspid valve
Congestive heart failure; nonhypertensive
Congestive heart failure (CHF) NOS
Circulatory disease NEC
Chronic pulmonary heart disease
Cerebrovascular disease
Cerebral artery occlusion, with cerebral infarction
Carditis
Cardiomegaly
Cardiac pacemaker/device in situ
Cardiac pacemaker in situ
Cardiac dysrhythmias
Cardiac conduction disorders
Cardiac arrest and ventricular fibrillation
Atrioventricular block, complete
Atrioventricular [AV] block
Atrial flutter
Atrial fibrillation and flutter
Atrial fibrillation
Atherosclerosis of the extremities
Atherosclerosis
Aortic aneurysm
Phenotypes
ACE2
ACTA2
BAG3
CDKN1A
NOTCH1
NTPROBNP
BNP
TNNI3
Proteins
Odds Ratio
6
4
2
−log10(P value)
20
30
40
Direction of association
positive

c)


Circulatory system
Ventricular fibrillation and flutter
Varicose veins of lower extremity, symptomtic
Varicose veins of lower extremity
Varicose veins
Tachycardia NOS
Stricture of artery
Sinoatrial node dysfunction (Bradycardia)
Pulmonary heart disease
Primary/intrinsic cardiomyopathies
Primary pulmonary hypertension
Premature beats
Peripheral vascular disease
Pericarditis
Paroxysmal ventricular tachycardia
Paroxysmal tachycardia, unspecified
Other hypertensive complications
Other forms of chronic heart disease
Other disorders of arteries and arterioles
Other chronic ischemic heart disease, unspecified
Other aneurysm
Other acute and subacute forms of ischemic heart disease
Orthostatic hypotension
Occlusion of cerebral arteries
Occlusion and stenosis of precerebral arteries
Nonrheumatic tricuspid valve disorders
Nonrheumatic pulmonary valve disorders
Nonrheumatic mitral valve disorders
Nonrheumatic aortic valve disorders
Noninfectious disorders of lymphatic channels
Myocardial infarction
Left bundle branch block
Ischemic Heart Disease
Hypotension
Hypertensive heart disease
Hypertensive heart and/or renal disease
Hypertensive chronic kidney disease
Heart valve replaced
Heart valve disorders
Endocarditis
Coronary atherosclerosis
Congestive heart failure; nonhypertensive
Congestive heart failure (CHF) NOS
Circulatory disease NEC
Chronic pulmonary heart disease
Cerebral ischemia
Cerebral artery occlusion, with cerebral infarction
Cardiomyopathy
Cardiomegaly
Cardiac pacemaker/device in situ
Cardiac pacemaker in situ
Cardiac dysrhythmias
Cardiac conduction disorders
Cardiac arrest
Bundle branch block
Atrioventricular [AV] block
Atrial fibrillation and flutter
Atrial fibrillation
Atherosclerosis of the extremities
Atherosclerosis of renal artery
Atherosclerosis
Arterial embolism and thrombosis of lower extremity artery
Arterial embolism and thrombosis
Arrhythmia (cardiac) NOS
Aortic aneurysm
Acute rheumatic heart disease
Abdominal aortic aneurysm
Phenotypes
ACE2
ACTA2
ACTN4
BAG3
NTPROBNP
BNP
TNNI3
Proteins
−log10(P value)
11
13
15
17
Odds Ratio
6
5
4
3
2
Direction of association
positive

d)

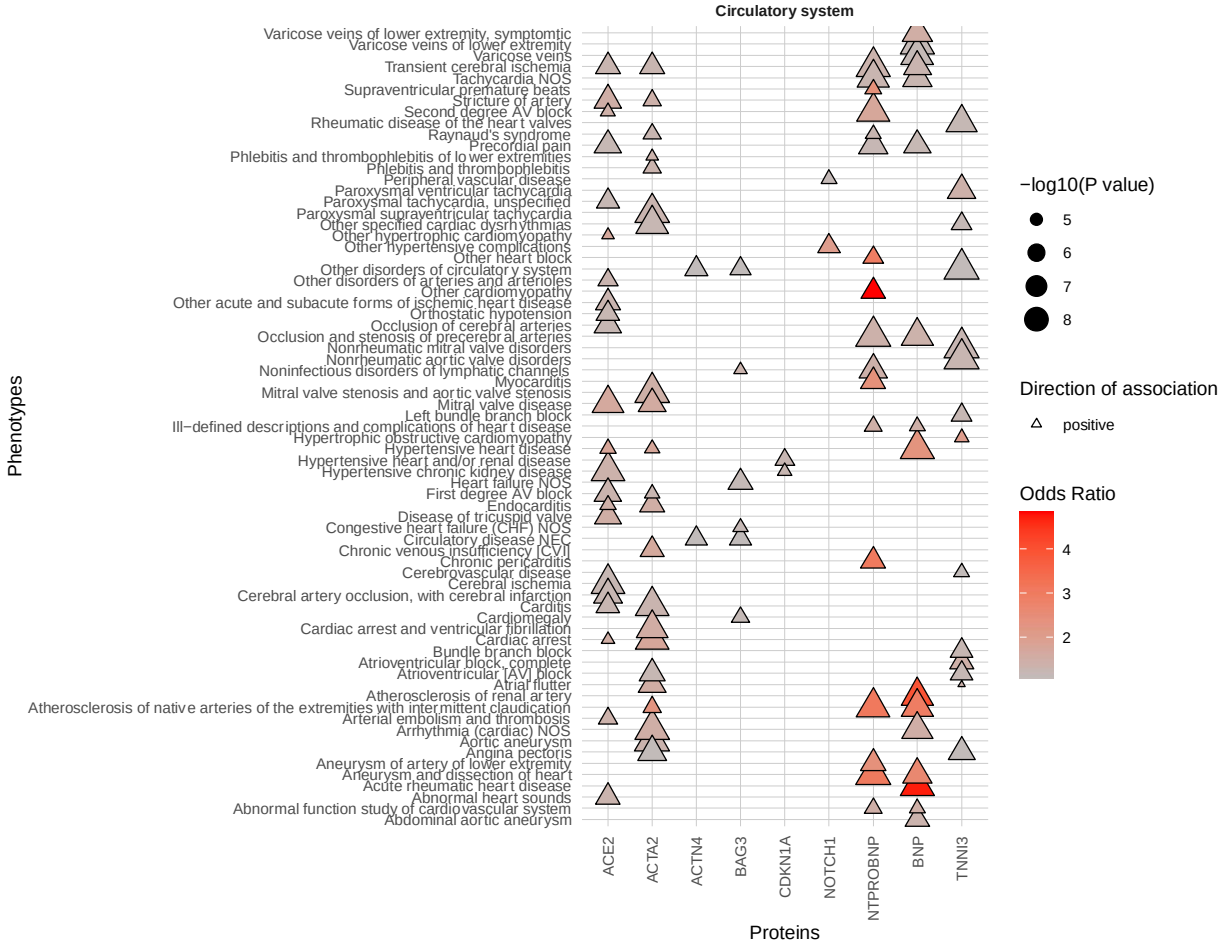


Figure S30 Phenome-wide association study results of the plasma protein levels with the circulatory system category of phenotypes.

a)-d) results are separated into four plots for clarity. Phenotypes as phecodes are described on the y-axis and the protein traits on the x-axis. Each point denotes a significant PheWAS association with a Bonferroni correction for the number of analyzed phecodes. The shape and colour denote the direction of effect and odds ratio. See **Table S1** for the full PheWAS results.