

Title:

Combining Genetic Proxies of Drug Targets and Time-to-event analyses From Longitudinal Observational Data To Identify Target Patient Populations

Authors:

Luke Zhang MSc¹, Prachi Kulkarni MSc^{1,2}, Farshad Farshidfar MD PhD¹, Whit Tingley MD PhD¹, Tim Hoey PhD¹, Whedy Wang MPH PhD¹, James R. Priest MD^{1,3}, Sylwia M. Figarska PhD¹

Affiliations:

1 Tenaya Therapeutics, South San Francisco, California, USA

2 University of California San Diego, San Diego, California, USA

3 Department of Pediatric, Stanford University, Stanford, California, USA

Short title: Identifying patient subpopulations through CVD genetic survival analyses

Abstract

Background: Human genetics is an important tool for identifying genes as potential drug targets, and the extensive genetic study of cardiovascular disease provides an opportunity to leverage genetics to match specific patient populations to specific drug targets to improve prioritization of patient selection for clinical studies.

Methods: We selected well described genetic variants in the region of *PCSK9* (rs11591147 and rs562556), *ADRB1* (rs7076938), *ACE* (rs4968782 and rs4363) and *BAG3* (rs2234962) for use as proxies for the effects of drugs. Time-to-event analyses were utilized to evaluate their effects on atrial fibrillation (AF) and heart failure (HF) death and/or re-hospitalization using real-world longitudinal dataset. To mitigate the effect of confounding factors for cardiovascular (CV) outcomes, we employed propensity score matching.

Results: After matching, a genetic proxy for PCSK9 inhibition (rs11591147) improved survival from CV death/heart transplant in individuals following a diagnosis of ischemic heart disease (Hazard Ratio (HR) 0.78, $P=0.04$). A genetic proxy for beta-blockade (rs7076938) improved freedom from rehospitalization or death in individuals with AF (HR 0.92, $P=0.001$), and a genetic proxy of ACE inhibition (rs7076938) improved freedom from death or rehospitalization for HF (HR 0.84, $P=0.017$) and AF (HR 0.85, $P=0.0014$). A protective variant in BAG3 (rs2234962) showed decreased risk of HF rehospitalization or CV death/heart transplant composite outcome within 10 years in HF patients (HR=0.96, $P=0.033$). Notably, despite smaller cohort sizes after matching, we often observed numerically smaller HRs and reduced P , indicating more pronounced effects and increased statistical association. However, not all genetic proxies replicated known treatment effects.

Conclusions: Genetic proxies for well-known drugs corroborate findings from clinical trials in cardiovascular disease. Our results may demonstrate a novel analytical approach that leverages genetic evidence from a large cohort to effectively select patient populations where specific drug targets may be most effective.

Introduction

Drug discovery and clinical development is a high-risk process that may require decades of capital investment prior to clinical use with an estimated median cost of \$1.1 billion for each new drug approved¹. Cardiovascular diseases with high heterogeneity such as heart failure (HF) frequently require complex and lengthy clinical trials enrolling thousands of patients to demonstrate efficacy². However in some examples, a small effect in a large group may be derived from a large effect in a smaller subgroup of affected individuals^{3,4}. This phenomenon has been observed in a number of cardiovascular therapies, which have shown greater efficacy or clinical utility for specific subgroups of disease following regulatory approval^{5,6}. Recognizing these nuances can guide the development of personalized medicine strategies, facilitating improved treatment effects in more targeted populations. This approach not only significantly reduces development costs but also optimizes therapeutic outcomes, thereby increasing the likelihood of success in drug development.

Regulatory bodies have issued guidance intended to increase the efficiency of drug development and support precision medicine by identifying subgroups of affected individuals where a specific therapeutic approach may be most effective⁷. Oncology has been at forefront of incorporating genetic information into drug discovery by identifying genetic targets which define prognosis in subtypes of large heterogeneous disease entities, and then designing therapies to specifically manipulate these targets^{8,9}. Outside of oncology, human genetics has proven to be a useful tool for selection of drug targets, and may also have the potential to enhance the design and conduct of clinical trials in cardiovascular medicine.

Here we describe efforts to develop a novel analytical framework intended to generate a strong pre-clinical hypothesis about the patient population(s) where a therapeutic intervention is most likely to be effective using a real-world clinical dataset and time-to-event analyses. To test and validate our analytical approach for cardiovascular disease, we leverage genetic variants with a known biological effect on gain, loss of function, or protein expression levels, as well as expected associations with cardiovascular phenotypes. With such variants we construct *in silico* trials using real world data and compare the results to the outcomes of well-established randomized clinical trials of approved therapies. We demonstrate the potential utility of our

analytical framework using known therapeutic targets across diverse areas of cardiovascular medicine, including lipid biology to atherosclerosis outcomes, beta-adrenergic signaling to HF and atrial fibrillation (AF) outcomes. Furthermore, we aim to explore whether a therapy originally developed for treating Mendelian cardiomyopathy, a specific genetic heart condition, might also offer therapeutics benefits to a broader population of individuals diagnosed with HF.

METHODS

Study Population, Clinical Data, and Outcomes

Genetic and clinical data in the UKB cohort were obtained from the UKB (<https://www.ukbiobank.ac.uk>) and is available to researchers through a streamlined application process. The UKB was approved by the North West Multi-Centre Research Ethics Committee and all participants provided written informed consent to participate in the study. We used International Classification of Diseases ICD-9 and ICD-10 codes to define patients' populations, co-morbidities and cardiovascular outcomes. Detailed ICD codes are provided in Supplemental Table 1. We considered two outcomes for this study: a) CV death or heart transplant (CV death/heart transplant), and b) rehospitalization or CV death/heart transplant - composite outcome. Details on selection of covariates (co-morbidities and medication use) considered for propensity score matching for the specific outcome are described in the supplementary material.

Statistical Approach

All statistical analyses were performed in R (version 4.3.3) using the *MatchIt* for propensity score matching, and *survival* packages for time-to-event analyses. For inclusion, exclusion, and matching of clinical covariates which can vary with time (such as medication usage, pre-existing conditions), individuals were matched based on covariates at the time of the initial diagnosis. For each genetic variant, individuals with effect allele were defined by variant allele (heterozygous or homozygous) and compared to individuals with the other allele. Covariate balance post-matching was assessed using standardized mean differences in Love plots, generated with the `love.plot()` function in *cobalt* R package, examining the balance measures before and after conditioning. Survival analyses were performed using Cox Proportional Hazards model and visualized with `survplot` as Kaplan-Meier or Cox Hazards curves as appropriate. For each variant

analyzed, we also performed a post-hoc power calculation based upon the available population size, number of events, and follow-up time, using the *SurvSNP* R package (details are provided in the Supplementary Material). The goal was to enhance our understanding of the reliability of our findings and to potentially guide future research endeavors. Additionally, for each SNP we conducted a PheWAS using PHENome Scan ANalysis Tool (PHESANT)¹⁰ and details of the analysis are provided in the Supplementary Material.

Genetic Variants

Similar to the approach for instrumental variable selection in Mendelian Randomization, we selected a set of genetic variants with clearly explainable biological impact on a gene which is the target of either an approved or proposed therapeutic modality. A missense variant in *PCSK9* (rs11591147 p.Arg46Leu) represents the effect of PCSK9 inhibition and has a protective effect on hyperlipidemia and coronary heart disease¹¹. We also tested another *PCSK9* missense variant, (rs562556 p.Val474Ile) that is likely to have an impact on PCSK9 activity and previously showed associations with LDL levels and CAD risk^{11,12}. Another variant of interest is rs7076938 located upstream of beta-adrenergic receptor beta-1 (*ADRB1*) (tagging for Gly389Arg rs1801253 $r^2=0.96$) that increases blood pressure and thus mimics effects of beta-adrenergic agonists. A genetic variant, rs4968782, located upstream of the Angiotensin-converting enzyme gene (*ACE*) likely affects transcription factor binding, and is a strong eQTL associated with increased *ACE* expression in the lung; therefore mimicking the effect of ACE inhibitors to modulate risk of hypertension. Also, effects of this SNP are stronger than effects of coding variants that were initially linked to circulating ACE levels and ACE activity¹³. We also tested an intronic variant in *ACE*, rs4363, that previously showed a very strong association¹⁴ with ACE levels reaching significance of $P=2 \times 10^{-257}$. This intronic SNP is highly correlated ($r^2=0.9$ in EUR) with a missense variant rs4343 associated with ACE activity¹⁵.

While no therapy is currently approved for *BAG3*, a coding variant in *BAG3* gene (C151R) is associated with decreased incidence of HF and has been mechanistically demonstrated to increase the protective effect of the BAG3 protein from proteo-toxicity in cardiomyocytes¹⁶.

Results:

As a first pass diagnostic analysis, we aimed to replicate the previously described case-control associations with the selected variants in the UKB, all of which showed associations with

expected phenotypes (Supplemental Table 2). The numbers of allele carriers per variant and event are presented in the figure panels (Figures 3-6) corresponding to each SNP. For patients who did not experience an event, follow-up period was calculated from the time of disease diagnosis until the most recent data update from the UKB, version September 2023. Events of non-cardiovascular death were right censored at the time of death. To evaluate the composite outcome of rehospitalization for disease of interest (HF or AF) or CV death/heart transplant, the follow-up period from the time of diagnosis was censored at the number of years where the impact of the SNP was most representative. All the tested variants met the underlying assumptions for proportional hazard risk over time as indicated by log-log plots.

The analytical framework (Figure 1) consists of selecting the individuals heterozygous or homozygous for the alternative allele for the variant of interest, selection of relevant known clinical risk factors including age at primary diagnosis, clinical comorbidities, and medication use. Cox Proportional Hazard models were performed prior to matching (with adjustments for age at disease diagnosis and sex), and after matching (without adjustments). Given the known associations for each of the selected variants and the intent of replicating known therapeutic effects, for the purposes of the methods development we defined statistical significance as a nominal P less than 0.05 and we did not perform correction for testing of multiple hypotheses. When the number of individuals homozygous for the effect allele was less than 25% of the number of heterozygous individuals, the groups were combined into a single group after matching. For each variant analyzed, we also performed a post-hoc power calculation based upon the available population size, number of events, and follow-up time. A summary of start and end timepoints for survival analyses conducted for each SNP, as well as the number of patients included in the study, can be seen in Figure 2. Additionally, power calculations for survival analyses conducted for each SNP can be seen in Supplemental Table 3.

PCSK9 rs11591147

PCSK9 is a well validated genetic target with a clearly defined mechanism of action to reduce circulating cholesterol levels and multiple successful clinical trials for prevention of atherosclerotic cardiovascular disease (ASCVD). The LoF variant rs11591147 disrupts the function of the *PCSK9* which results in higher amount of cell surface LDLR and internalization

of LDL, leading to lower serum LDL levels and subsequently confer protection against ischemic heart disease (IHD)^{17,18}. To replicate the successful trials for PCSK9 inhibitors for secondary prevention of ASCVD, we examined cardiovascular death or heart transplant among individuals after a diagnosis of IHD. The impact of rs11591147 on time to CV death/transplant among genetically determined Caucasians diagnosed with IHD did not reach statistical significance, exhibiting a hazard ratio (HR) of 0.95 and a *P* of 0.53. However, after matching on comorbidities existing at the time of diagnosis of ASCVD (Supplemental Figure 1), medication usage, and cardiovascular comorbidities, carriers of one or more effect alleles (T) demonstrated improved survival compared to individuals with no copies of the effect allele (Figure 3a). The HR markedly decreased to 0.78 with a significant *P* of 0.03, indicating a protective effect of PCSK9 rs11591147 against the development of severe cardiovascular outcomes from an IHD diagnosis. The difference between the unmatched and matched HRs underscore the importance of accounting for other causal factors, including some that were implemented as part of a treatment strategy, in determining the true impact of PCSK9 inhibition on cardiovascular outcomes (Figure 3). Results of covariate balancing for each matching term are presented in Supplemental Figure 1.

PCSK9 rs562556

The impact of rs562556 on time to CV death/transplant from date of first IHD diagnosis among Caucasians did not reach statistical significance, exhibiting a HR of 1.01 and *P* of 0.342. Even after matching on comorbidities (Supplemental Figure 6) existing at the time of diagnosis of ASCVD, medication usage, and cardiovascular comorbidities, no statistically significant impact was observed by rs562556 (HR=1.06, *P*=0.641, Supplemental Figure 5).

ADRB1 rs7076938

ADRB1 encodes the β_1 adrenergic receptor with primary cardiovascular effects in the heart sinus node and juxtaglomerular cells of the kidney to modulate contractility, heart rate, and blood pressure. It is the target of many selective beta blockers used in primary and secondary prevention of ASCVD, HF, hypertension, and AF¹⁹. The variant rs7076938 is a tagging SNP for rs1801253 (Gly389Arg, $r^2=0.96$ in EUR and $r^2=0.81$ in all populations) that has been shown to alter post-receptor signaling, with the Arg389 receptor (tagged by rs7076938 T) coupling more

efficiently with its corresponding G protein (Gs)^{20–22} and correlating with higher blood pressure^{23,24}. In our study we tested rs7076938, as it was more significantly associated with SBP in large genome wide association studies (GWAS)²⁴ and was used as a part of instrument variable in MR analysis representing antihypertensive medication proxy²⁵. Analogous to successful trials of beta blockers for treatment of AF, we examined AF progression with the combined outcome of rehospitalization for AF or CV death/heart transplant²⁶. Carriers of TT genotype had increased risk of CV death/heart transplant since first AF diagnosis (HR=1.169, $P=0.0012$) and the effect remained significant in the matched data (HR=1.172, $P=0.0031$). Additionally, carriers of the TT genotype showed increased risk of AF rehospitalization or CV death/heart transplant in unmatched (HR=1.01, $P=0.0043$) and matched participants (HR=1.081, $P=0.0012$). The impact of ADRB1 on cardiovascular outcomes before and after matching can be seen in Figure 4. Results of covariate balancing can be seen in Supplemental Figure 2.

ACE rs4968782

The ACE gene encodes the angiotensin converting enzyme expressed in the lungs which is a key step in the renin/aldosterone-angiotensin system regulating vascular tone and blood pressure. The variant rs4968782 likely affects transcription factor binding and is a strong eQTL affecting ACE levels. Inhibitors of ACE such as lisinopril are a mainstay of treatment for hypertension, HF, and AF. Carriers of at least one A allele exhibited decreased risk of CV death/heart transplant from first AF diagnosis in unmatched (HR=0.94, $P=0.029$) and matched (HR=0.85, $P=0.0014$) participants. When the composite outcome – HF rehospitalization or CV death/heart transplant – was considered, the impact of the A allele was only seen in matched data (HR=0.84, $P=0.017$ vs HR=0.93, $P=0.14$ in unmatched data). It is important to note that the time-to-event analysis of CV death/heart transplant from first AF diagnosis was performed in the Caucasians as the SNP only showed effects in this subset, while time to composite outcome was performed in the entire population regardless of ethnicity. These results underscore the impact of propensity score matching in elucidating the significance of the ACE SNP in cardiovascular outcomes, revealing previously undetected associations, and enhancing the precision of risk assessments (Figure 5). Results of covariate balancing are displayed in Supplemental Figure 3.

ACE rs4363

For ACE rs4363, carriers of the A allele did not exhibit decreased risk of CV death/heart transplant from first AF diagnosis (HR=0.97, $P=0.197$ in unmatched, HR=0.95, $P=0.17$ in matched). When the composite outcome – HF rehospitalization or CV death/heart transplant was considered, the impact of the A allele was not seen in either unmatched (HR=1.01, $P=0.89$) or matched data (HR=1.04, $P=0.45$). It is important to note that the time-to-event analysis of CV death/heart transplant from first AF diagnosis was performed in Caucasians, while time to composite outcome was performed in the entire population, regardless of ethnicity in order to mirror the analysis of ACE rs4968782. Survival analysis results can be seen in Supplemental Figure 7, while the results of covariate balancing can be seen in Supplemental Figure 8.

BAG3 rs2234962 (C151R)

BAG3 plays an important role in the protein quality control pathway within cardiomyocytes, and loss of function mutations can cause a form of dilated cardiomyopathy²⁸. The protective effect of the C allele on the composite outcome of HF rehospitalization or CV death/transplant within 10 years was observed in all HF patients (HR=0.96, $P=0.033$) and remained essentially unchanged after matching (HR=0.95, $P=0.011$). Since a recent study has shown that delivery of additional Bag3 to the heart using AAV9 may result in improved functional parameters in the myocardial infarction model of mouse HF²⁹, we tested effect of the SNP in the subset of HF patients of ischemic origin. We observed a protective HR against HF rehospitalization or CV death/transplant composite outcome 10 years in matched data (HR=0.95); however, this effect did not reach statistical significance ($P=0.059$ and HR=0.97, $P=0.202$ in ischemic HF patients before matching, Figure 6). Results of covariate balancing are presented in Supplemental Figure 4. There was no observed effect of rs2234962 on ischemic HF or all cause HF when CV death/heart transplant was considered independently (HR=0.99, $P=0.703$ and HR=1.01, $P=0.803$ respectively in unmatched data).

PheWAS was conducted for all the SNPs mentioned above, with the aim of finding associations mentioned in previous publications. PheWAS results for each individual SNP can be seen in Supplemental Figures 9-14.

DISCUSSION

Here we describe a genetic survival analysis which employs a well-characterized genetic instrumental variable as a proxy for potential treatment with a specific drug or target. This approach has the potential to identify treatment effects using real-world outcome data through time-to-event analyses. In three examples of targets with approved therapies (PCSK9, ADRB1, ACE), this analytical framework identifies beneficial treatment effects seen in clinical trials, and a potential benefit in a clinical scenario for a novel target (BAG3) suggested by mouse models of ischemic cardiomyopathy. The approach described here is a natural extension of GWAS of clinical outcomes, Mendelian Randomization, and similar time-to-event GWAS^{30–32}. Genetic survival analysis offers a critical advantage over standard applications of GWAS or Mendelian Randomization which typically only consider non-genetic covariates such as age and sex. Our proposed approach not only estimates the direction of effect for a specific drug or target in a real-world setting, but it also incorporates clinical covariates, causal factors, and standard of care therapies commonly encountered in clinical trials.

Matching is a well-accepted statistical technique to mitigate the influence of measured confounders and constitutes an essential part of clinical trial design often implemented in the form of ‘randomization’. Although matching on the propensity score is often effective at eliminating differences between the allele carrier groups to achieve covariate balance, the performance of this method must be carefully assessed. In addition to covariate balance, the quality of the match is determined by how many units remain after matching^{33,34}. Matching often involves discarding units that are not paired with others. Moreover, some matching options, such as setting restrictions for common support or calipers, may further reduce the number of units available for analysis. In our research, successful matching often led to a notable reduction in sample size but also a more pronounced effect of the SNP. This suggests that our matching procedure effectively mitigated the influence of confounding factors and enabled us to discern a true effect of the SNP. Put differently, the improvement in power to detect known effects of genetic instruments for PCSK9, ADRB1, and ACE following matching underscores the methodological robustness of our approach and highlights its compatibility with complexities inherent in real-world longitudinal observational data. It is important to note that matching did not resolve all the complexities of real-world data, as the ACE variant rs4363 did not appear to

modify the risk of adverse outcomes in AF or HF, which contradicts findings from previous clinical trials.

Cardiovascular diseases such as HF are frequently complex, where a broad diagnosis may arise from a combination of one and often more environmental and genetic causes. Findings from genetic data such as GWAS hold great potential in guiding drug discovery, as these studies have the unique capability for identifying risk loci for a disease without an *a priori* hypotheses. While GWAS is useful in identifying novel genetic loci that are beyond our current understanding of the disease, cell-based or animal models are unable to fully mimic the human condition, limiting the success rate of translating preclinical findings into clinical practice. Additionally, the genetic factors identified from GWAS for disease susceptibility may not always be the same factors that govern disease progression or long-term outcomes—which are often the endpoints of clinical trials and the most meaningful outcomes to affected individuals and physicians. Therefore, for drug target prioritization, more studies based on clinical samples of affected individuals with real world phenotype data may more accurately reflect the genetic basis of the condition under study. To ensure precise investigation of specific patient populations, our study utilized only Hospital Episode Statistics (HES) data and ICD codes, excluding self-reported cardiovascular complaints without direct medical diagnoses. Further, aligning with best practices in clinical trial enrollment, our analysis did not limit participants to Caucasian individuals —except in noted cases— and included all participants, irrespective of ethnicity.

Like other techniques based upon human genetics, the application and interpretability of genetic survival analysis has important limitations. Coding and non-coding common genetic variants that may be selected as instrumental variables typically have small effect sizes, which may be obscured by other clinical factors unless the number of individuals and events are relatively large. Conversely, the power to detect survival benefit may be compromised even with large effect sizes seen with rare variants which may suffer from a small number of individuals and events. Additionally, while our findings illustrate that matching on comorbidities and stratification appear to be useful tools for ‘focusing’ the technique to enhance the power, they do not eliminate the possibility of reverse causality or collider bias. Such biases can occur when individuals that carry a particular genetic variant have a measurably different risk of disease

susceptibility —often observable in case-control GWAS — which impacts long-term disease outcome^{35,36}. The analytical framework may only be applied when a genetic instrumental variable is present in the population of interest. An additional limitation of our current application is its inability to provide a meaningful estimate of treatment effect size for the target, which is among the most critical information derived from clinical trials for regulatory bodies, commercial entities, and payors^{37,38}.

A significant strength of the genetic survival analysis work presented here lies in the diversity of genetic variants examined and the broad range of cardiovascular outcomes assessed, including IHD, HF, and AF. This comprehensive approach helps identify specific patient populations that would benefit most from potential treatments.

A logical extension of the analytical framework presented here would be to conduct *in silico* trials to select and optimize specific clinical variables and co-morbidities, with the goal of maximizing effect size (and minimizing the cost and duration of a clinical trial) for a new treatment. This approach of stratification might not be limited to clinical variables or co-morbidities but might also logically include other genetic factors such as polygenic scoring or stratification by Mendelian forms of cardiovascular disease. This leads to an important limitation of the analyses presented here; though we chose genetic instruments largely based on efficacious and widely accepted therapies, our findings are derived from a single large cohort study (UKB) and would benefit greatly from validation with an independent dataset.

Conclusions

Our study's findings suggest that genetic survival analysis could be instrumental in generating clinically meaningful hypotheses. This method identifies individuals and diseases that are likely to benefit from treatment of a specific target, and therefore could help guide the design of future clinical trials.

Acknowledgments:

This research was conducted using the UK Biobank Resource under Application number 84103.

Sources of Funding: All authors are employees of Tenaya Therapeutics.

Disclosures: All other authors declare no competing financial interests.

Code and Examples:

<https://github.com/tenayatherapeutics/Genetic-Survival-Analysis-in-UKB/tree/main>

Supplementary Material

Supplemental Methods

Tables S1-S3

Figures S1-S14

1. Wouters OJ, McKee M, Luyten J. Errors in Source Data for Study of Drug Development Costs. *JAMA*. 2022;328.
2. Teerlink JR, Diaz R, Felker GM, McMurray JJV, Metra M, Solomon SD, Adams KF, Anand I, Arias-Mendoza A, Biering-Sørensen T, et al. Cardiac Myosin Activation with Omecamtiv Mecarbil in Systolic Heart Failure. *N Engl J Med*. 2021;384.
3. Felker GM, Solomon SD, Claggett B, Diaz R, McMurray JJV, Metra M, Anand I, Crespo-Leiro MG, Dahlström U, Goncalvesova E, et al. Assessment of Omecamtiv Mecarbil for the Treatment of Patients With Severe Heart Failure: A Post Hoc Analysis of Data From the GALACTIC-HF Randomized Clinical Trial. *JAMA Cardiol* [Internet]. 2022 [cited 2024 Apr 10];7:26–34. Available from: <https://pubmed.ncbi.nlm.nih.gov/34643642/>
4. DOCHERTY KF, MCMURRAY JJV, DIAZ R, FELKER GM, METRA M, SOLOMON SD, ADAMS KF, BÖHM M, BRINKLEY DM, ECHEVERRIA LE, et al. The Effect of Omecamtiv Mecarbil in Hospitalized Patients as Compared With Outpatients With HFrEF: An Analysis of GALACTIC-HF. *J Card Fail*. 2024;30.
5. Lincoff AM, Brown-Frandsen K, Colhoun HM, Deanfield J, Emerson SS, Esbjerg S, Hardt-Lindberg S, Hovingh GK, Kahn SE, Kushner RF, et al. Semaglutide and Cardiovascular Outcomes in Obesity without Diabetes. *N Engl J Med* [Internet]. 2023 [cited 2024 Apr 10];389:2221–2232. Available from: <https://pubmed.ncbi.nlm.nih.gov/37952131/>
6. Bernhard B, Heydari B, Abdullah S, Francis SA, Lumish H, Wang W, Jerosch-Herold M, Harris WS, Kwong RY. Effect of six month's treatment with omega-3 acid ethyl esters on long-term outcomes after acute myocardial infarction: The OMEGA-REMODEL randomized clinical trial. *Int J Cardiol* [Internet]. 2024 [cited 2024 Apr 10];399. Available from: <https://pubmed.ncbi.nlm.nih.gov/38184150/>
7. Fda, Cder, Clampet, Jennifer. Enrichment Strategies for Clinical Trials to Support

- Determination of Effectiveness of Human Drugs and Biological Products Guidance for Industry. 2019 [cited 2024 Apr 10]; Available from: <https://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm> and <https://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/default.htm>
8. Druker BJ, Lydon NB. Lessons learned from the development of an abl tyrosine kinase inhibitor for chronic myelogenous leukemia. *J Clin Invest* [Internet]. 2000 [cited 2024 Apr 10];105:3–7. Available from: <https://pubmed.ncbi.nlm.nih.gov/10619854/>
 9. Ryan CJ, Devakumar LPS, Pettitt SJ, Lord CJ. Complex synthetic lethality in cancer. *Nat Genet* [Internet]. 2023 [cited 2024 Apr 10];55:2039–2048. Available from: <https://pubmed.ncbi.nlm.nih.gov/38036785/>
 10. Millard LAC, Davies NM, Gaunt TR, Smith GD, Tilling K. Software Application Profile: PHESANT: a tool for performing automated phenome scans in UK Biobank. *Int J Epidemiol* [Internet]. 2018 [cited 2024 Apr 10];47:29–35. Available from: <https://pubmed.ncbi.nlm.nih.gov/29040602/>
 11. Rao AS, Lindholm D, Rivas MA, Knowles JW, Montgomery SB, Ingelsson E. Large-Scale Phenome-Wide Association Study of PCSK9 Variants Demonstrates Protection Against Ischemic Stroke. *Circ Genomic Precis Med* [Internet]. 2018 [cited 2019 Sep 9];11:e002162. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/29997226>
 12. Gai MT, Adi D, Chen XC, Liu F, Xie X, Yang YN, Gao XM, Ma X, Fu ZY, Ma YT, et al. Polymorphisms of rs2483205 and rs562556 in the PCSK9 gene are associated with coronary artery disease and cardiovascular risk factors. *Sci Rep* [Internet]. 2021 [cited 2024 Apr 10];11. Available from: <https://pubmed.ncbi.nlm.nih.gov/34075144/>
 13. Chung CM, Wang RY, Chen JW, Fann CSJ, Leu HB, Ho HY, Ting CT, Lin TH, Sheu SH, Tsai WC, et al. A genome-wide association study identifies new loci for ACE activity: potential implications for response to ACE inhibitor. *Pharmacogenomics J* [Internet]. 2010 [cited 2024 Apr 10];10:537–544. Available from: <https://pubmed.ncbi.nlm.nih.gov/20066004/>
 14. Gudjonsson A, Gudmundsdottir V, Axelsson GT, Gudmundsson EF, Jonsson BG, Launer LJ, Lamb JR, Jennings LL, Aspelund T, Emilsson V, et al. A genome-wide association study of serum proteins reveals shared loci with common diseases. *Nat Commun* [Internet]. 2022 [cited 2024 Apr 10];13. Available from: <https://pubmed.ncbi.nlm.nih.gov/35078996/>
 15. Chung CM, Wang RY, Chen JW, Fann CSJ, Leu HB, Ho HY, Ting CT, Lin TH, Sheu SH, Tsai WC, et al. A genome-wide association study identifies new loci for ACE activity: Potential implications for response to ACE inhibitor. *Pharmacogenomics J*. 2010;10:537–544.
 16. Perez-Bermejo JA, Judge LM, Jensen CL, Wu K, Watry HL, Truong A, Ho JJ, Carter M, Runyon W V., Kaake RM, et al. Functional analysis of a common BAG3 allele associated with protection from heart failure. *Nat Cardiovasc Res* 2023 27 [Internet]. 2023 [cited 2024 Apr 10];2:615–628. Available from: <https://www.nature.com/articles/s44161-023-00288-w>
 17. Cameron J, Holla ØL, Ranheim T, Kulseth MA, Berge KE, Leren TP. Effect of mutations in the PCSK9 gene on the cell surface LDL receptors. *Hum Mol Genet* [Internet]. 2006 [cited 2024 Apr 10];15:1551–1558. Available from: <https://pubmed.ncbi.nlm.nih.gov/16571601/>

18. Berge KE, Ose L, Leren TP. Missense mutations in the PCSK9 gene are associated with hypocholesterolemia and possibly increased response to statin therapy. *Arterioscler Thromb Vasc Biol* [Internet]. 2006 [cited 2024 Apr 10];26:1094–1100. Available from: <https://pubmed.ncbi.nlm.nih.gov/16424354/>
19. Correction to: 2023 ACC/AHA/ACCP/HRS Guideline for the Diagnosis and Management of Atrial Fibrillation: A Report of the American College of Cardiology/American Heart Association Joint Committee on Clinical Practice Guidelines. *Circulation* [Internet]. 2024 [cited 2024 Apr 10];149:E936. Available from: <https://pubmed.ncbi.nlm.nih.gov/38408149/>
20. Mason DA, Moore JD, Green SA, Liggett SB. A gain-of-function polymorphism in a G-protein coupling domain of the human beta1-adrenergic receptor. *J Biol Chem* [Internet]. 1999 [cited 2024 Apr 10];274:12670–12674. Available from: <https://pubmed.ncbi.nlm.nih.gov/10212248/>
21. Sandilands AJ, O’Shaughnessy KM. The functional significance of genetic variation within the beta-adrenoceptor. *Br J Clin Pharmacol* [Internet]. 2005 [cited 2024 Apr 10];60:235–243. Available from: <https://pubmed.ncbi.nlm.nih.gov/16120061/>
22. Zhang F, Steinberg SF. S49G and R389G polymorphisms of the β_1 -adrenergic receptor influence signaling via the cAMP-PKA and ERK pathways. *Physiol Genomics* [Internet]. 2013 [cited 2024 Apr 10];45:1186–1192. Available from: <https://pubmed.ncbi.nlm.nih.gov/24151242/>
23. Bengtsson K, Melander O, Orho-Melander M, Lindblad U, Ranstam J, Råstam L, Groop L. Polymorphism in the beta(1)-adrenergic receptor gene and hypertension. *Circulation* [Internet]. 2001 [cited 2024 Apr 10];104:187–190. Available from: <https://pubmed.ncbi.nlm.nih.gov/11447084/>
24. Surendran P, Feofanova E V., Lahrouchi N, Ntalla I, Karthikeyan S, Cook J, Chen L, Mifsud B, Yao C, Kraja AT, et al. Discovery of rare variants associated with blood pressure regulation through meta-analysis of 1.3 million individuals. *Nat Genet* [Internet]. 2020 [cited 2024 Apr 10];52:1314–1332. Available from: <https://pubmed.ncbi.nlm.nih.gov/33230300/>
25. Hyman MC, Levin MG, Gill D, Walker VM, Georgakis MK, Davies NM, Marchlinski FE, Damrauer SM. Genetically Predicted Blood Pressure and Risk of Atrial Fibrillation. *Hypertens (Dallas, Tex 1979)* [Internet]. 2021 [cited 2024 Apr 10];77:376–382. Available from: <https://pubmed.ncbi.nlm.nih.gov/33390040/>
26. Steeds RP, Birchall AS, Smith M, Channer KS. An open label, randomised, crossover study comparing sotalol and atenolol in the treatment of symptomatic paroxysmal atrial fibrillation. *Heart* [Internet]. 1999 [cited 2024 Apr 10];82:170–175. Available from: <https://pubmed.ncbi.nlm.nih.gov/10409530/>
27. Packer M, Poole-Wilson PA, Armstrong PW, Cleland JGF, Horowitz JD, Massie BM, Rydén L, Thygesen K, Uretsky BF. Comparative effects of low and high doses of the angiotensin-converting enzyme inhibitor, lisinopril, on morbidity and mortality in chronic heart failure. ATLAS Study Group. *Circulation* [Internet]. 1999 [cited 2024 Apr 10];100:2312–2318. Available from: <https://pubmed.ncbi.nlm.nih.gov/10587334/>
28. Judge LM, Perez-Bermejo JA, Truong A, Ribeiro AJS, Yoo JC, Jensen CL, Mandegar MA, Huebsch N, Kaake RM, So PL, et al. A BAG3 chaperone complex maintains cardiomyocyte function during proteotoxic stress. *JCI insight* [Internet]. 2017 [cited 2024 Apr 10];2. Available from: <https://pubmed.ncbi.nlm.nih.gov/28724793/>

29. Knezevic T, Myers VD, Su F, Wang JF, Song J, Zhang XQ, Gao E, Gao G, Madesh M, Gupta MK, et al. Adeno-associated Virus Serotype 9 - Driven Expression of BAG3 Improves Left Ventricular Function in Murine Hearts with Left Ventricular Dysfunction Secondary to a Myocardial Infarction. *JACC Basic to Transl Sci* [Internet]. 2016 [cited 2024 Apr 10];1:647–656. Available from: <https://pubmed.ncbi.nlm.nih.gov/28164169/>
30. Hukerikar N, Hingorani AD, Asselbergs FW, Finan C, Schmidt AF. Prioritising genetic findings for drug target identification and validation. *Atherosclerosis* [Internet]. 2024 [cited 2024 Apr 10];390. Available from: <https://pubmed.ncbi.nlm.nih.gov/38325120/>
31. Bi W, Fritsche LG, Mukherjee B, Kim S, Lee S. A Fast and Accurate Method for Genome-Wide Time-to-Event Data Analysis and Its Application to UK Biobank. *Am J Hum Genet* [Internet]. 2020 [cited 2024 Apr 10];107:222–233. Available from: <https://pubmed.ncbi.nlm.nih.gov/32589924/>
32. Pedersen EM, Agerbo E, Plana-Ripoll O, Steinbach J, Krebs MD, Hougaard DM, Werge T, Nordentoft M, Børglum AD, Musliner KL, et al. ADuLT: An efficient and robust time-to-event GWAS. *Nat Commun* [Internet]. 2023 [cited 2024 Apr 10];14. Available from: <https://pubmed.ncbi.nlm.nih.gov/37689771/>
33. Walsh MC, Trentham-Dietz A, Newcomb PA, Gangnon R, Palta M. Use of Propensity Scores to Reduce Case-Control Selection Bias. *Epidemiology*. 2012;23:772–735.
34. Fu EL, Groenwold RHH, Zoccali C, Jager KJ, Van Diepen M, Dekker FW. Merits and caveats of propensity scores to adjust for confounding. *Nephrol Dial Transplant* [Internet]. 2019 [cited 2024 Apr 10];34:1629–1635. Available from: <https://pubmed.ncbi.nlm.nih.gov/30215791/>
35. Karim ME, Petkau J, Gustafson P, Platt RW, Tremlett H, BeAMS Study Group. Comparison of statistical approaches dealing with time-dependent confounding in drug effectiveness studies. *Stat Methods Med Res* [Internet]. 2018 [cited 2018 Jun 27];27:1709–1722. Available from: <http://www.ncbi.nlm.nih.gov/pubmed/27659168>
36. VanderWeele TJ, Tchetgen Tchetgen EJ, Cornelis M, Kraft P. Methodological Challenges in Mendelian Randomization. *Epidemiology* [Internet]. 2014 [cited 2019 Aug 23];25:427–435. Available from: <http://content.wkhealth.com/linkback/openurl?sid=WKPTLP:landingpage&an=00001648-201405000-00014>
37. Estrada K, Froelich S, Wuster A, Bauer CR, Sterling T, Clark WT, Ru Y, Trinidad M, Nguyen HP, Luu AR, et al. Identifying therapeutic drug targets using bidirectional effect genes. *Nat Commun* [Internet]. 2021 [cited 2024 Apr 10];12. Available from: <https://pubmed.ncbi.nlm.nih.gov/33850126/>
38. Sun BB, Kurki MI, Foley CN, Mechakra A, Chen CY, Marshall E, Wilk JB, Ghen CY, Wilk JB, Runz H, et al. Genetic associations of protein-coding variants in human disease. *Nature* [Internet]. 2022 [cited 2024 Apr 10];603:95–102. Available from: <https://pubmed.ncbi.nlm.nih.gov/35197637/>

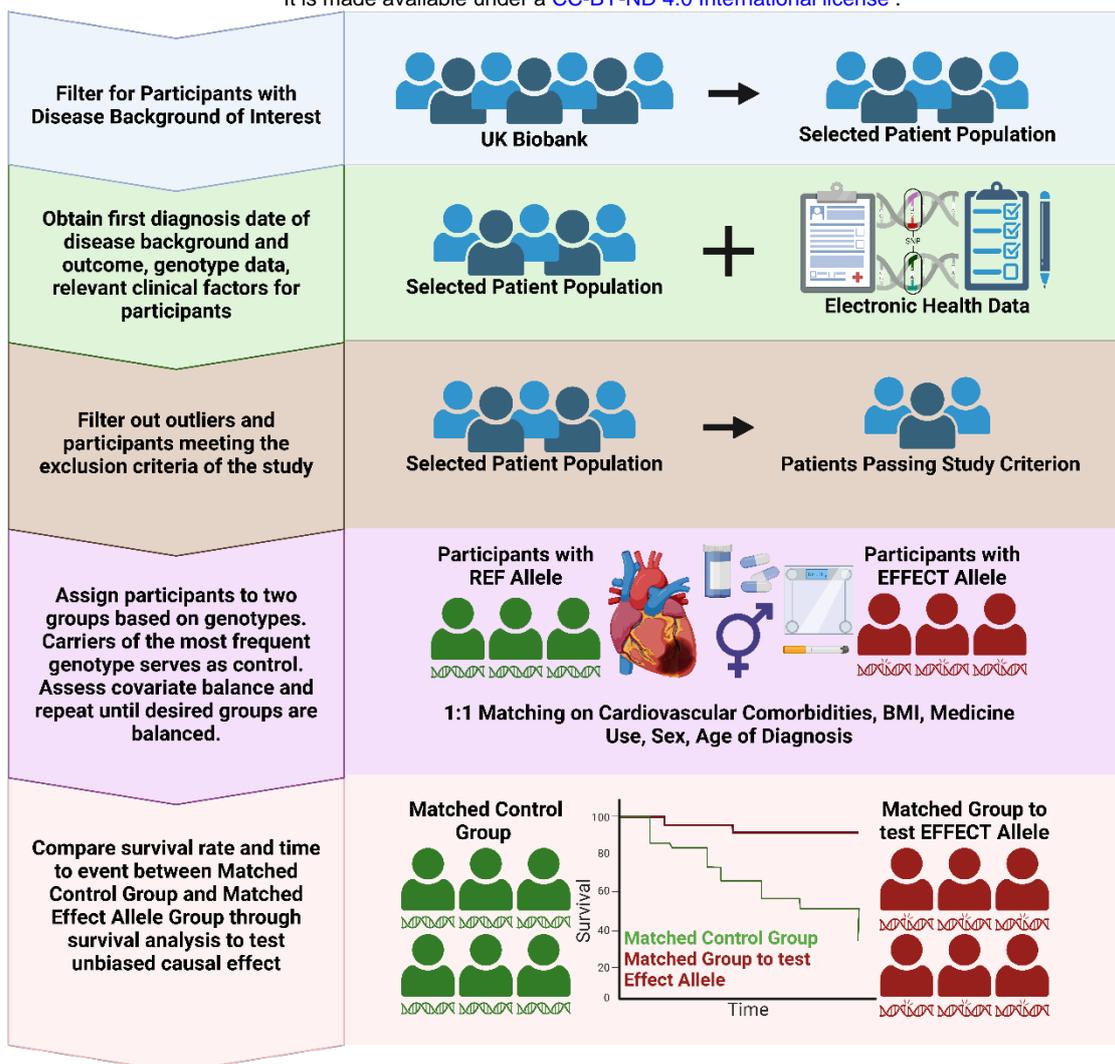


Figure 1. Analytical framework. Participants diagnosed with a disease of interest were included in the study. Information on co-morbidities and medication use was obtained and used for propensity score matching procedure. Survival analysis was performed for matched allele carriers.

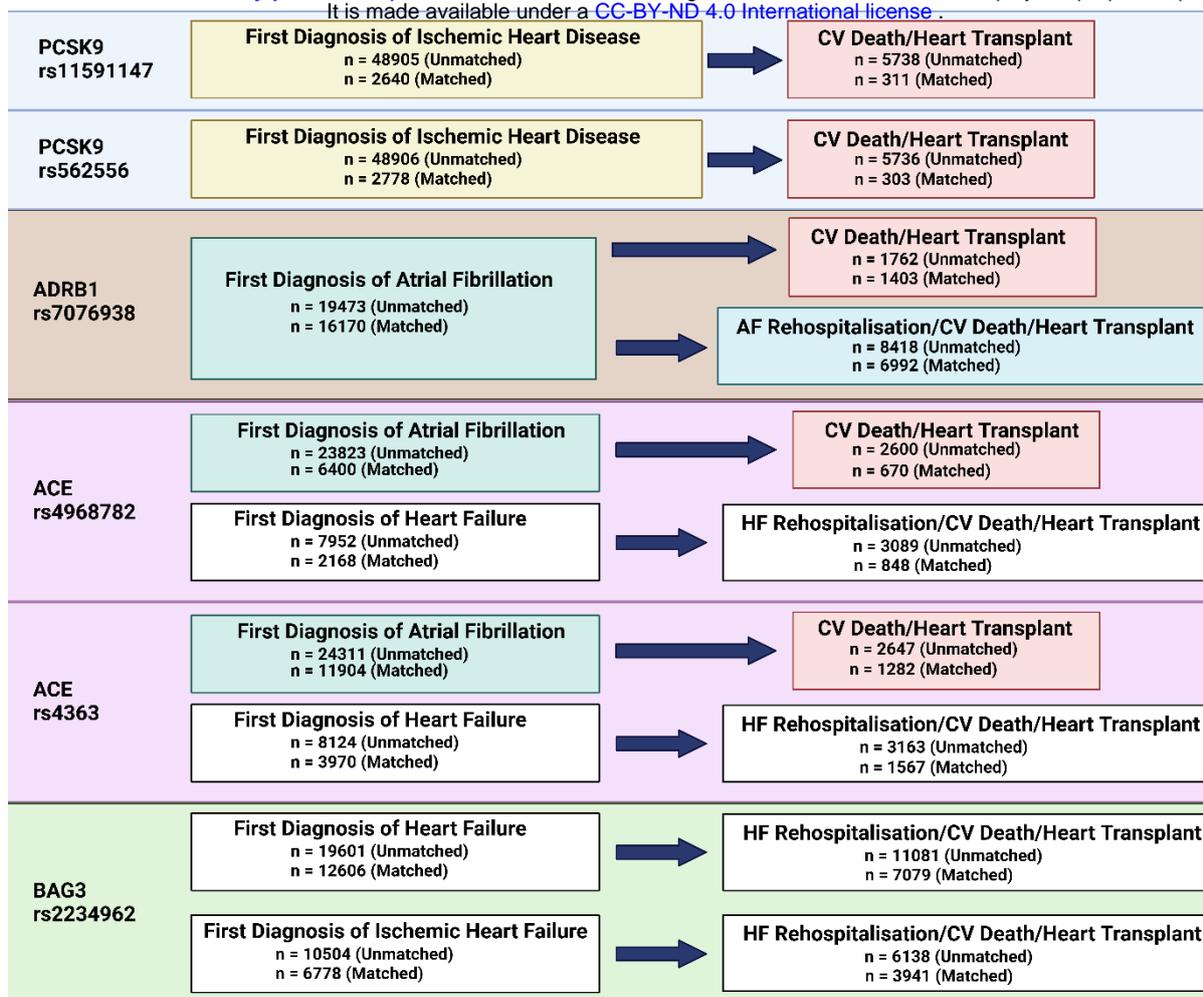


Figure 2. Summary of all survival analysis performed for each SNP in the study, indicating disease of interest and considered outcomes.

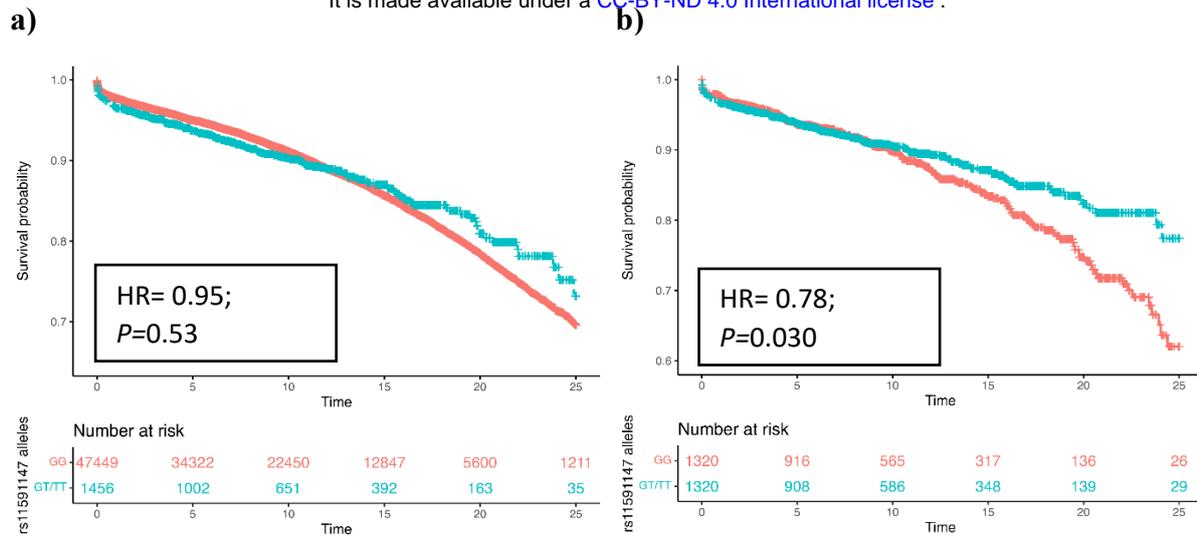


Figure 3. Effect of PCSK9 rs11591147 on time to cardiovascular (CV) death/heart transplant from first ischemic heart disease diagnosis (IHD) in a) unmatched and b) matched participants. Only Caucasians were included. Individuals with GT/TT genotypes (proxy for PCSK9 inhibitor treatment) showed better survival for CV death/heart transplant compared to individuals with GG genotype (controls).

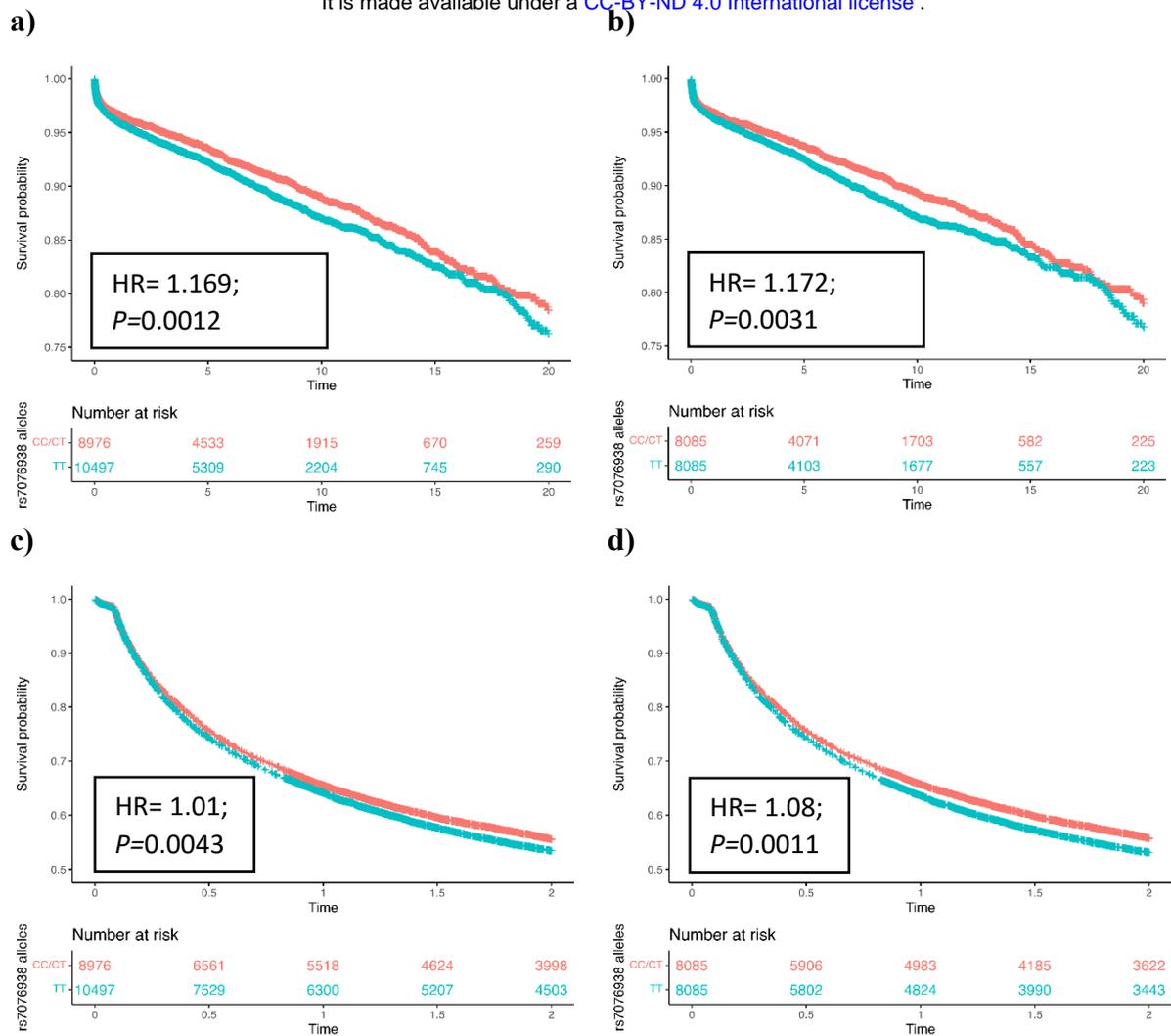


Figure 4. Effect of ADRB1 rs7076938 on time to cardiovascular (CV) death/heart transplant from first atrial fibrillation (AF) diagnosis in a) unmatched and b) matched participants, and time to AF rehospitalization or CV death/heart transplant in c) unmatched and d) matched participants. Individuals with TT genotype showed increased risk of AF rehospitalization of CV death/heart transplant composite outcome compared to individuals with CC/CT genotypes.

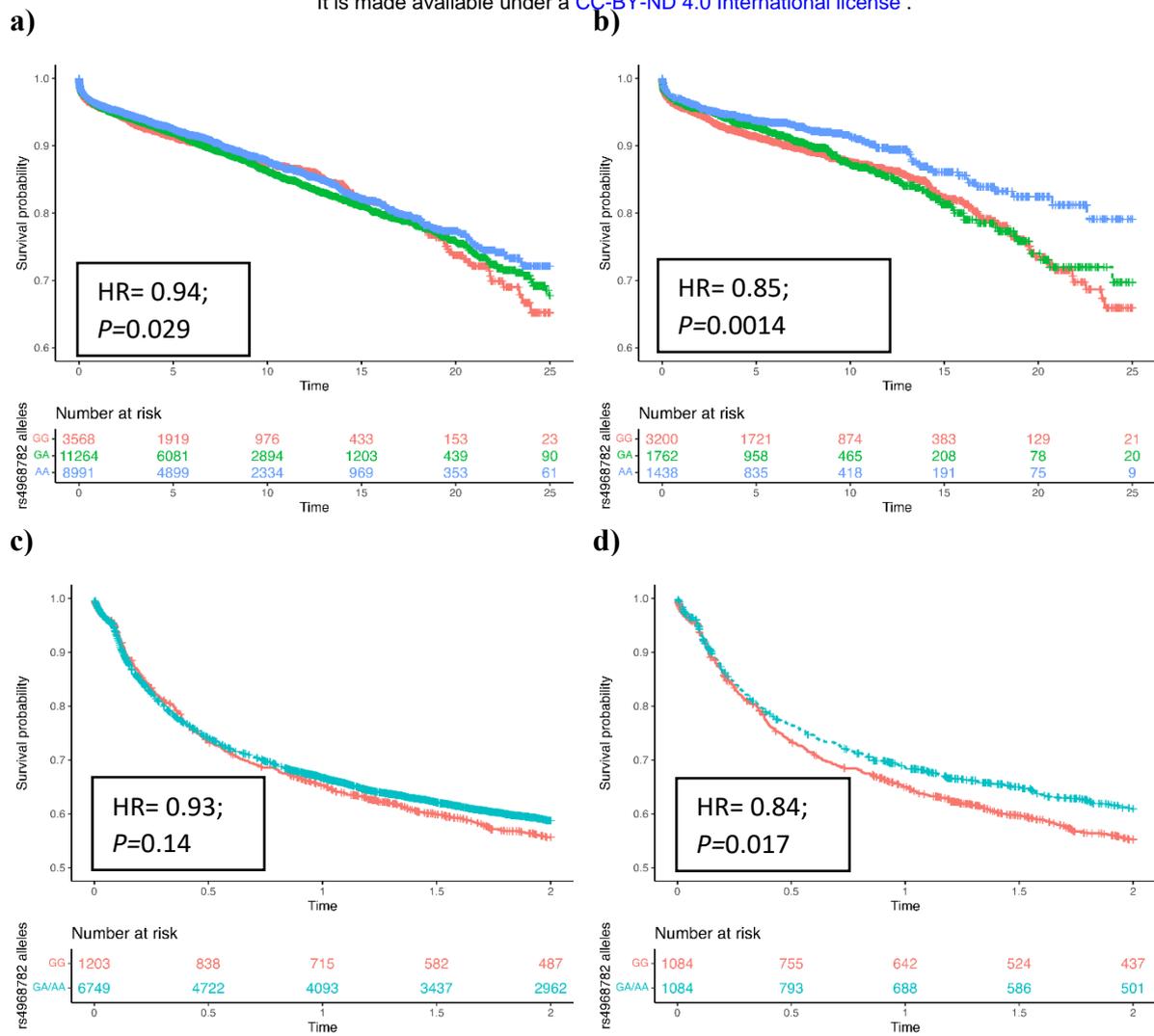


Figure 5. Effect of ACE rs4968782 on time to cardiovascular (CV) death/heart transplant from first atrial fibrillation (AF) diagnosis in a) unmatched and b) matched Caucasians, and time to heart failure (HF) rehospitalization or CV death/heart transplant in c) unmatched and d) matched participants of any ethnicity. Carriers of GA or AA genotype showed better survival for rehospitalization and severe cardiovascular outcome compared to individuals with GG genotype.

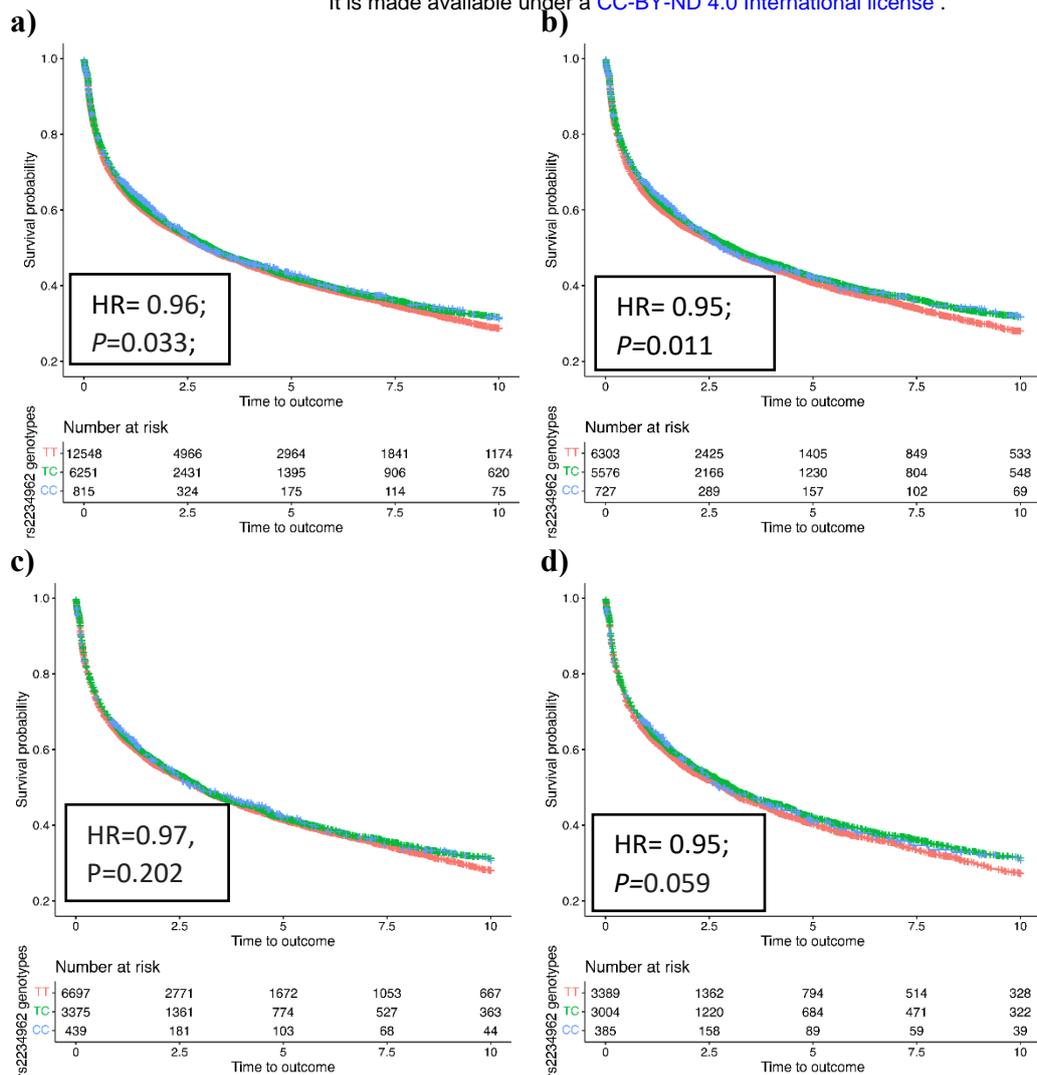


Figure 6. Effect of BAG3 rs2234962 on time to composite outcome – heart failure (HF) rehospitalization or cardiovascular (CV) death/heart transplant from first HF diagnosis in a) unmatched and b) matched participants, and on time to composite outcome – HF rehospitalization or CV death/heart transplant from first HF diagnosis due to ischemic origin in c) unmatched and d) matched participants. Carriers of the C allele showed decreased risk of HF rehospitalization or CV death/heart transplant composite outcome in HF patients compared to those with the T allele.

SUPPLEMENTAL MATERIAL

Study samples

The UK Biobank (UKB) is a prospective cohort study with genotype and detailed phenotype information for up to 502,269 individuals aged 40-69 years when recruited between 2006 and 2010. UKB participant data, linked to NHS central registers for information on death, cancer registrations, and hospital admissions in England, Scotland, and Wales up to September 1st, 2023, are enriched with demographic characteristics, lifestyle details, and self-reported health outcomes. This comprehensive dataset, which includes hospital admission records from 1985 for Scotland and 1997 for England and Wales, undergoes record linkage based on NHS number, postcode, sex, and age. Collected baseline information encompasses various facets, such as age, gender, ethnicity, country of birth, education, socioeconomic status, BMI, and smoking status. Extracted from linked HES data are dates, duration of hospitalization, ICD-10 diagnostic codes, and OPCS-4 procedural codes. Genotype data is also available for most participants, which includes positions that are either directly genotyped on one of two arrays (UKB Axiom and UK BiLEVE arrays) or imputed using a reference panel from the Trans-Omics for Precision Medicine (TOPMed). Participants who did not have genetic data were excluded from this study. First available diagnosis dates, derived from HESIN (Hospital inpatient data) operations and diagnosis were considered for this study. Rehospitalization was defined as the next instance occurring 30 days after the first diagnosis, with endpoints including rehospitalization, cardiovascular death, study end (September 1st, 2023). Events of non-cardiovascular death were censored at the time of death. The starting dates correspond to when a patient was initially diagnosed with a specific disease of interest. UK Biobank was approved by the North West Multi-Centre Research Ethics Committee and all participants provided written informed consent to participate.

Propensity score matching

Propensity score matching, implemented through nearest neighbor stratification using the *MatchIt* R package (Ho D, Imai K, King G, Stuart E (2011). “MatchIt: Nonparametric Preprocessing for Parametric Causal Inference.” *Journal of Statistical Software*, 42(8), 1–28. doi:10.18637/jss.v042.i08.), was a pivotal method employed to enhance comparability between treatment and control populations for the subsequent assessment of survival outcomes. The technique involved assigning individuals with available genetic data (either genotyped or TOPMed imputed) into control and treatment groups based on their reported SNP dosage, typically categorized as rounded dosages of 0 and those having at least one of the effect alleles (combining 1 and 2 dosages). Propensity scores were generated to represent the probability of receiving the genetic treatment, considering various baseline covariates.

To ensure a well-balanced comparison, participants were meticulously matched on critical factors listed for each SNP below. The evaluation of balance before and after matching was conducted using the standardized mean difference, with an absolute value less than 0.1 for total distance serving as a criterion for well-balanced groups. Propensity score matching not only facilitated the control of potential confounding variables but also enhanced the validity of subsequent survival analyses by creating more comparable cohorts based on clinical characteristics.

The study focused on testing single nucleotide polymorphisms (SNPs) associated with known cardiovascular effects, specifically examining rs11591147 and rs562556 from PCSK9, rs7076938 from ADRB1, and rs4968782 as well as rs4363 from the ACE gene. Regression analyses were conducted using glm function of the *glmnet* package in R, employing both Gaussian and binomial models linear and binary outcomes. The *glmnet* package (Version 4.1-8; Friedman et al., 2010) facilitated data analysis. Genotypic information was available for 464,428 participants for rs4968782, 473,376 participants for rs7076938,

and 487,227 participants for rs11591147. This robust dataset enables a comprehensive exploration of the genetic associations between these selected SNPs and cardiovascular outcomes, providing valuable insights into the potential role of these genetic variations in cardiovascular health.

The rs11591147 and rs562556 polymorphisms within the PCSK9 gene were investigated, revealing an overall mean standardized difference of less than 0.01 in the examined cohort. The control group consisted of participants homozygous for the G allele while the treatment group consisted of participants with at least one copy of the T allele for rs11591147, or at least one copy of the A allele for rs562556. Covariate adjustments were implemented for sex, C-Reactive Protein, smoking, body mass index (BMI), low density lipoprotein (LDL) levels, first systolic blood pressure reading, diuretics use, angiotensin II receptor blocker (ARB) use, beta-blockers use, ACE inhibitor (ACEi) use, calcium channel blocker use, digoxin use, statins use, anti-coagulants use, age of ischemic heart disease (IHD) diagnosis, and previous disease diagnoses encompassing atrial fibrillation (AF), aortic valve stenosis, chronic renal failure, chronic obstructive pulmonary disease (COPD), heart failure (HF), dilated cardiomyopathy (DCM), hypertrophic cardiomyopathy (HCM), mitral valve disorder, type 1 diabetes (T1D), type 2 diabetes (T2D), unspecified stroke, and ventricular arrhythmias. The exclusion criteria for this study encompassed patients undergoing treatment with ezetimibe and centrally acting anti-hypertensive and vasodilator anti-hypertensive medications. These exclusions were instituted due to the pharmacological actions of these agents, which target LDL cholesterol, akin to the mechanism of PCSK9 inhibitors. In this analysis of PCSK9, only genetically determined Caucasian patients were included due to the observed heightened effect of the SNP within this demographic, enhancing detection sensitivity and statistical power for assessing its influence on the phenotype. Analysis in unmatched data was adjusted for sex and age of disease diagnosis, while analysis in matched data was unadjusted (Supplemental Figure 1).

The rs7076938 polymorphism within the ADRB1 gene was investigated, revealing an overall mean standardized difference of less than 0.01 in the studied population. The control group consisted of participants with at least one copy of the C allele, while the treatment group consisted of participants homozygous for the T allele. Covariate adjustments were made for various factors including sex, C-Reactive Protein level, smoking status, BMI, LDL levels, diuretics use, calcium channel blocker, alpha adrenoceptor blocking drug use, statins use, digoxin use, ARB use, anti-coagulant medication use, previous disease diagnosis for aortic valve stenosis, chronic renal failure, (COPD, HF, T2D, T1D, ischemic heart disease (IHD), mitral valve disorder, unspecified stroke, and age of AF diagnosis. We excluded individuals using ACEi or beta-blockers to mitigate confounding effects associated with the excluded medications (Supplemental Figure 2). This matching criterion was used for both the assessment of severe cardiovascular outcomes (CV death/heart transplant) after AF diagnosis and time to rehospitalization or CV death/heart transplant composite outcome from AF. Analysis in unmatched data was adjusted for sex and age of disease diagnosis, while analysis in matched data was unadjusted.

The rs4968782 and rs4363 polymorphisms within the ACE gene was scrutinized, revealing an overall mean standardized difference of less than 0.01 in the studied population. For each SNP, the control group consisted of participants homozygous for the G allele while the treatment group consisted of participants with at least one copy of the A allele. When investigating time to rehospitalization due to all-cause HF or CV death/heart transplant in patients diagnosed with HF, we balanced for sex, C-Reactive Protein level, smoking status, BMI, LDL, diuretics use, calcium channel blocker use, statins use, digoxin use, ARB use, anti-coagulant medication use, aortic valve stenosis status, chronic renal failure status, COPD status, HF status, T2D status, unspecified stroke status, IHD status, and age of HF diagnosis. We excluded individuals using ACEi, beta-blockers, alpha-adrenoceptor blocking drugs, vasodilator anti-hypertensives, and centrally acting anti-hypertensives to mitigate potential confounding effects associated with these medications (Supplemental Figure 3).

Furthermore, in the assessment of severe cardiovascular outcomes (CV death/heart transplant) after first available AF diagnosis, we balanced the data for sex, C-Reactive Protein, smoking, BMI, LDL, diuretics, beta blockers, calcium channel blockers, alpha-adrenoceptor blocking drugs, statins, anticoagulants, diabetic medications, age of AF disease diagnosis, and previous disease diagnosis for aortic valve stenosis, chronic renal failure, COPD, HF, unspecific stroke, IHD, T2D, unspecific stroke, ventricular arrhythmias, and cardiac arrhythmias (Supplemental Figure 3). We excluded individuals using ACEi, vasodilator anti-hypertensives, and centrally acting anti-hypertensives to mitigate potential confounding effects associated with these medications. Analysis in unmatched data was adjusted for sex and age of disease diagnosis, while analysis in matched data was unadjusted.

In this study, we employed the SurvSNP package (PMID: 22685040) to conduct power calculations, leveraging asymptotic estimates derived from genetic hazard ratios, event rates, study population size, and allele frequencies. These calculations were based on Cox proportional hazards model results obtained before and after genetic matching for rs11591147 and rs562556 polymorphisms within the PCSK9 gene. Our objective was to assess the extent of power enhancement achieved through our methodology when examining the time to severe cardiovascular outcomes from the first available diagnosis of IHD. Power calculations for all SNPs can be seen in Supplemental Table 3. Notably, for rs11591147, the power within the unmatched subset was 9.5, demonstrating a substantial increase to 73.4 after matching. Similarly, for rs562556, the power within the unmatched subset was 6.9, exhibiting an increase to 11.16 after matching. These findings underscore the enhancements in statistical power attainable through matched subset analysis, underscoring the pivotal role of selecting appropriate SNPs with substantial effect sizes for robust survival analysis outcomes. (Supplemental Figure 6)

To comprehensively elucidate the diverse array of phenotypic manifestations associated with each of the SNPs, we performed phenome-wide association study (PHEWAS) adjusted for age, sex and 10 genetic principal components, and applied additive model. The PHEWAS analyses were performed using the PHESANT (PHENome Scan ANalysis Tool) software application (PMID: 29040602). This tool facilitates automated phenome scans within the UK Biobank dataset, thereby enabling comprehensive exploration of phenotypic associations. The main dataset from the UK Biobank downloaded in March 2023 was supplemented with custom cardiovascular phenotypes created using HES update version September 2023. For logistic regression we only included phenotypes with cases > 50.

Supplemental Table 1: Phenotype definitions for cardiovascular endpoints and comorbidities used for propensity score matching.

Phenotype	ICD-9 codes	ICD-10 codes	OPCS-4 codes/other definitions
All-Cause Heart Failure	4254, 4280, 4281, 4289	I110, I130, I132, I255, I420, I425, I428, I429, I500, I501, I509	
Atrial Fibrillation	4273	I48, I480, I481, I482, I483, I484, I489	K571, K621, K622, K623, K624, X501, X502
Type II Diabetes		E110, E111, E112, E113, E114, E115, E116, E117, E118, E119	
Type I Diabetes		E100, E101, E102, E103, E104, E105, E106, E107, E108, E109	
Chronic Renal Failure	5850-5859	N180-N189, I120, I131, I132	M010-M019
Myocardial Infarction	4100-4109	I210-I219	
Mitral Valve Disorder	4240, 3940, 3942, 3949	I050-1059, I080, I081, I340-I349, I390	K250-K259, K301, K341, K351
Aortic Valve Stenosis	4241, 3951, 3959	I060-I069, I080, I082, I083, I350-I359, I391	K260-K269, K302
Ischemic Heart Disease	4100-4109, 4110-4119, 4120-4129, 4130-4139, 4140-4149	I200-I209, I210-I219, I220-I229, I230-I239, I240-I249, I250-I259	
Chronic Obstructive Pulmonary Disease	4910-4919, 4920-4929, 4960-4969	J410-J419, J420-J429, J430-J439, J440-J449	
Hypertrophic Cardiomyopathy	4251	I421, I422	
Coronary Artery Disease		I210-I219, I220-I229, I230-I239, I252	K490-K499, K500-I509, K750-K759, K400-K409, K410-K419, K420-K429, K430-

			K439, K440-K449, K450-K459, K460-K469
Dilated Cardiomyopathy	4254	I429, I420	
Cardiac Arrhythmias		I490-I499	
Ventricular Arrhythmias		I470-I479	
Unspecific Stroke	4300-4309, 4310-4319, 4320-4329, 4330-4339, 4340-4349, 4350-4359, 4360-4369	G450-G459, I600-I609, I610-I619, I620-I629, I630-I639	
Ischemic Stroke	4330-4339, 4340-4349	I630-I639	
Hemorrhagic Stroke	4310-4319, 4320-4329	I610-I619, I620-I629	
Heart Transplant (Endpoint)			K011, K012, K018, K019, K021, K022, K023, K024, K025, K026, K028, K029

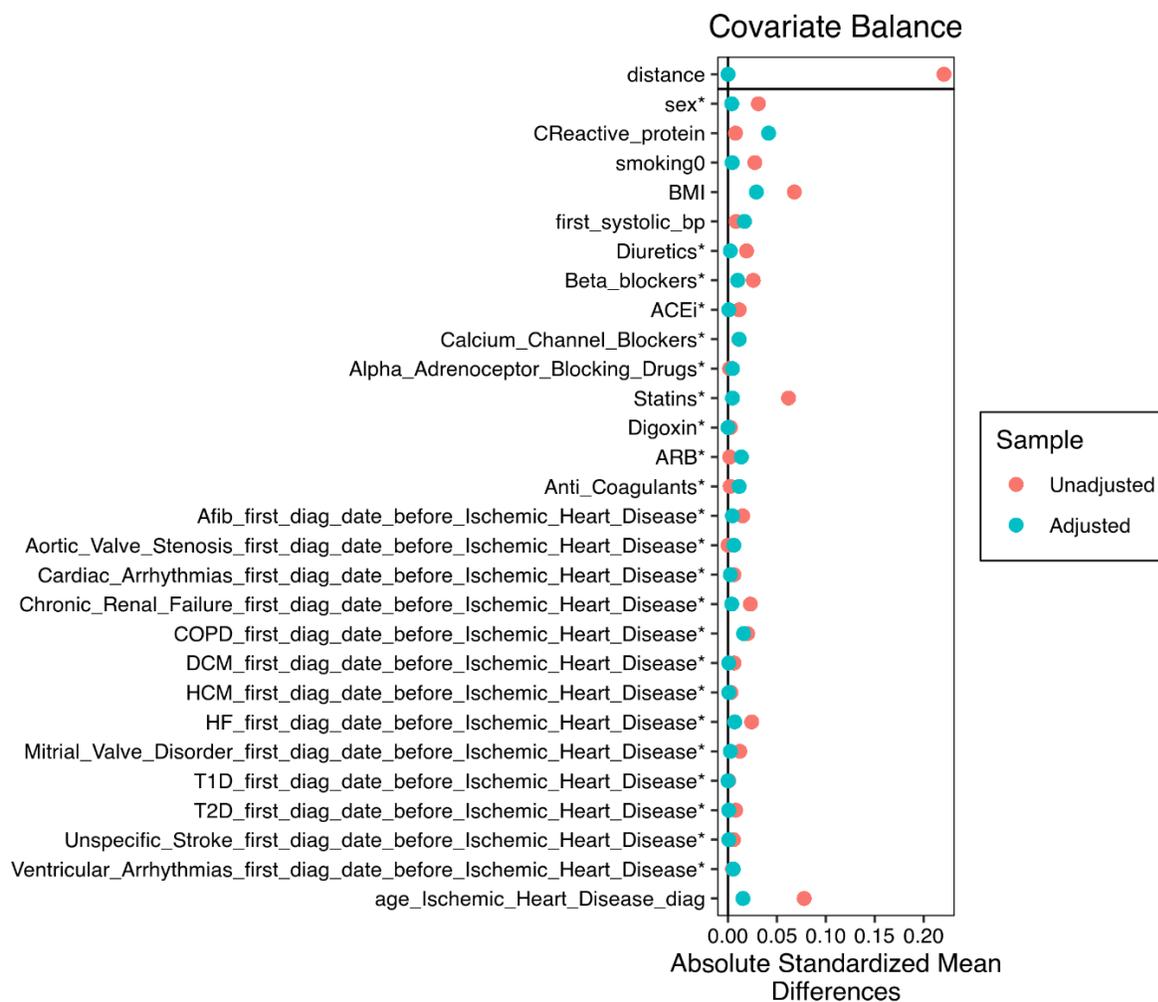
Supplemental Table 2: Previously known associations for the tested SNPs

Gene	SNP	Phenotype	N (cases/controls or total N)	Effect allele	Other allele	Beta	P
PCSK9	rs11591147	IHD	59775/427336	G	T	-0.1965	1.460E-13
		LDL levels	463346			-0.3055	< 2E-16
		LDL levels adj. for lipid medication	463324			-0.3548	< 2E-16
PCSK9	rs562556	IHD	59767/427295	A	G	0.0117	1.617E-01
		LDL levels	463299			0.0210	< 2E-16
		LDL levels adj. for lipid medication	463177			0.0452	1.55E-11
ADRB1	rs7076938	HF	19857/453403	T	C	0.0145	2.191E-01
		Afib	39706/433554			0.0217	1.231E-02
		SBP	449353			0.3598	< 2E-16
		SBP no hypertensive medication	325934			0.3905	7.050E-15
ACE	rs4968782	HF	19530/444790	A	G	-0.0210	5.100E-02
		Afib	38900/425420			-0.0179	2.382E-02
		SBP	440786			-0.1852	4.480E-06
		SBP no hypertensive medication	319628			-0.2214	1.810E-06
ACE	rs4363	HF	19956/455158	A	G	-0.0017	8.668E-01
		Afib	39807/435307			-0.0061	4.240E-01
		SBP	451093			-0.0392	3.143E-01
		SBP no hypertensive med	327167			-0.0629	1.589E-01
BAG3	rs2234962	HF	19684/450092	C	T	-0.0028	1.610E-08
		Ischemic HF	10553/459223			-0.0015	9.800E-05
		DCM	2441/467335			-0.0011	1.540E-09

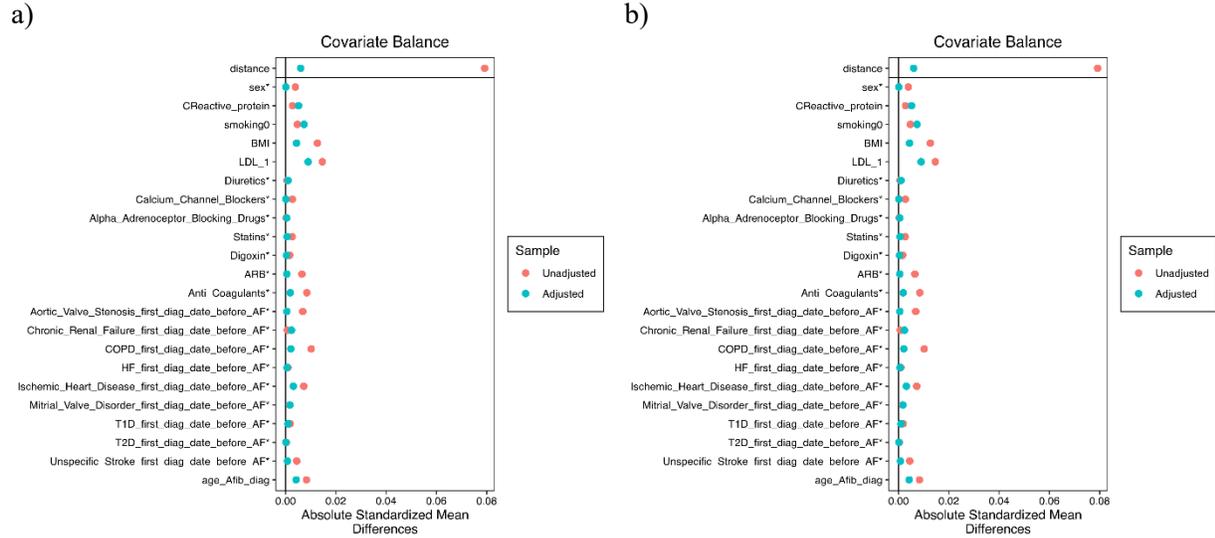
SNPs were tested in additive model, adjusted for age at recruitment, sex and 10 PCs. For SBP analysis all individuals taking ACE inhibitors, Beta-blockers, Centrally acting antihypertensive medication, vasodilator antihypertensive drugs, Alpha-Adrenoceptor Blocking Drugs, ARBs or Calcium-Channel Blockers were excluded from the analysis.

Supplemental Table 3: Power calculations for survival analysis

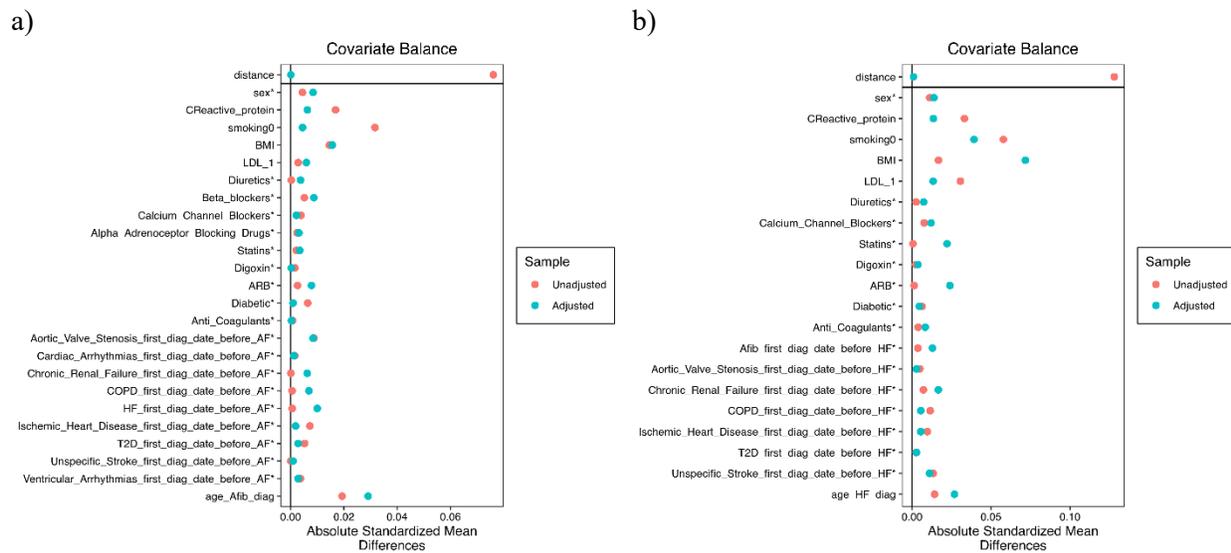
Gene	RSNUM	Diagnosis	Endpoint	Population	Matched	Power (%)
PCSK9	rs11591147	Ischemic Heart Disease	CVD/Heart Transplant	Caucasians	No	9.5
PCSK9	rs11591147	Ischemic Heart Disease	CVD/Heart Transplant	Caucasians	Yes	73.4
PCSK9	rs562556	Ischemic Heart Disease	CVD/Heart Transplant	Caucasians	No	6.9
PCSK9	rs562556	Ischemic Heart Disease	CVD/Heart Transplant	Caucasians	Yes	11.16
ACE	rs4968782	AF	CVD/Heart Transplant	Caucasians	No	59.23
ACE	rs4968782	AF	CVD/Heart Transplant	Caucasians	Yes	80.28
ACE	rs4968782	HF	HF Composite	All	No	79.81
ACE	rs4968782	HF	HF Composite	All	Yes	92.23
ACE	rs4363	AF	CVD/Heart Transplant	Caucasians	No	19.93
ACE	rs4363	AF	CVD/Heart Transplant	Caucasians	Yes	22.97
ACE	rs4363	HF	HF Composite	All	No	5.65
ACE	rs4363	HF	HF Composite	All	Yes	17.96
ADRB1	rs7076938	AF	CVD/Heart Transplant	All	No	98.09
ADRB1	rs7076938	AF	CVD/Heart Transplant	All	Yes	96.15
ADRB1	rs7076938	AF	AF Composite	All	No	8.75
ADRB1	rs7076938	AF	AF Composite	All	Yes	98.4
BAG3	rs2234962	HF	HF Composite	All	No	68.11
BAG3	rs2234962	HF	HF Composite	All	Yes	77.88
BAG3	rs2234962	Ischemic HF	HF Composite	All	No	26.58
BAG3	rs2234962	Ischemic HF	HF Composite	All	Yes	51.89



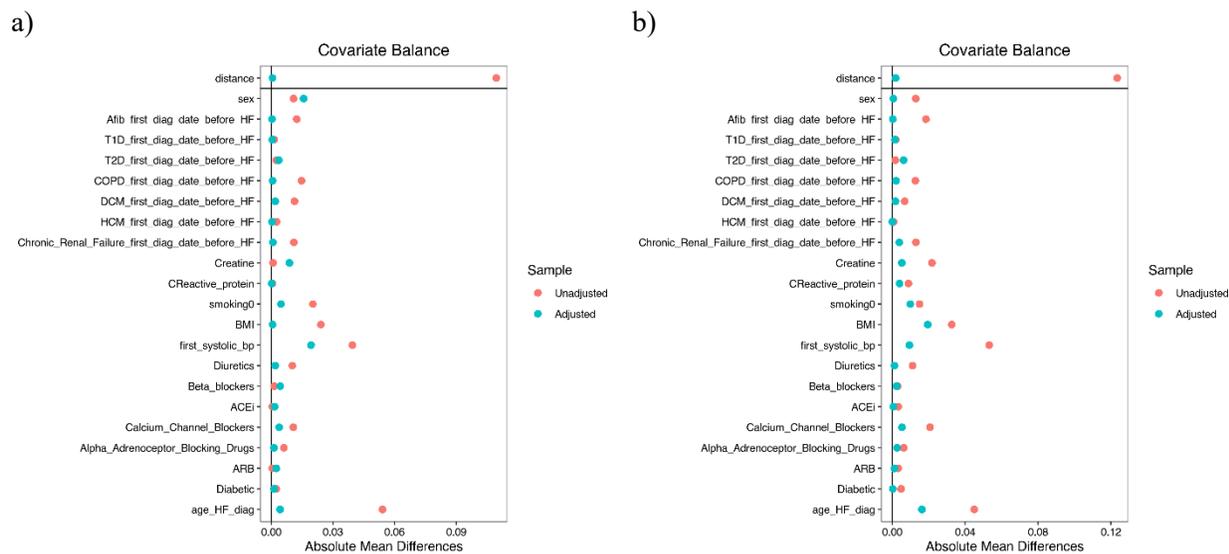
Supplemental Figure 1. Matching results for PCSK9 11591147 in Caucasians. Love plot indicates shift towards 0 in absolute standardized mean difference, representing successful matching. Asterisk (*) indicates variables for which the displayed value is the raw (unstandardized) difference in means.



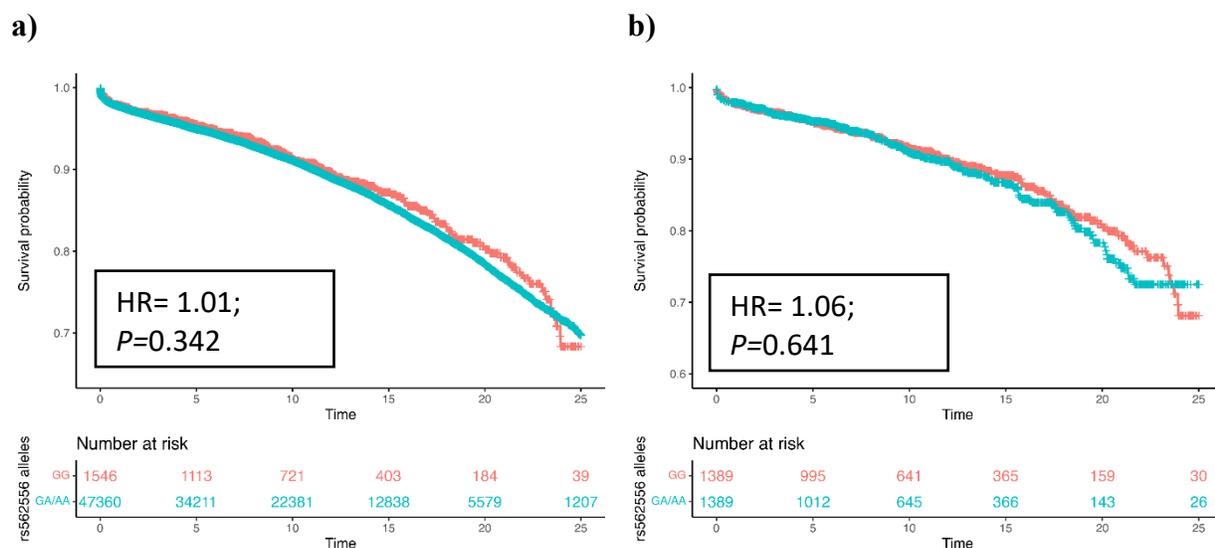
Supplemental Figure 2. Propensity score matching results for ADRB1 rs7076938 for a) time to cardiovascular (CV) death/heart transplant from first atrial fibrillation diagnosis (AF), and b) time to rehospitalization or CV death/heart transplant due to AF. A shift towards 0 for absolute standardized mean difference indicates improvement in differences between populations carrying specific alleles.



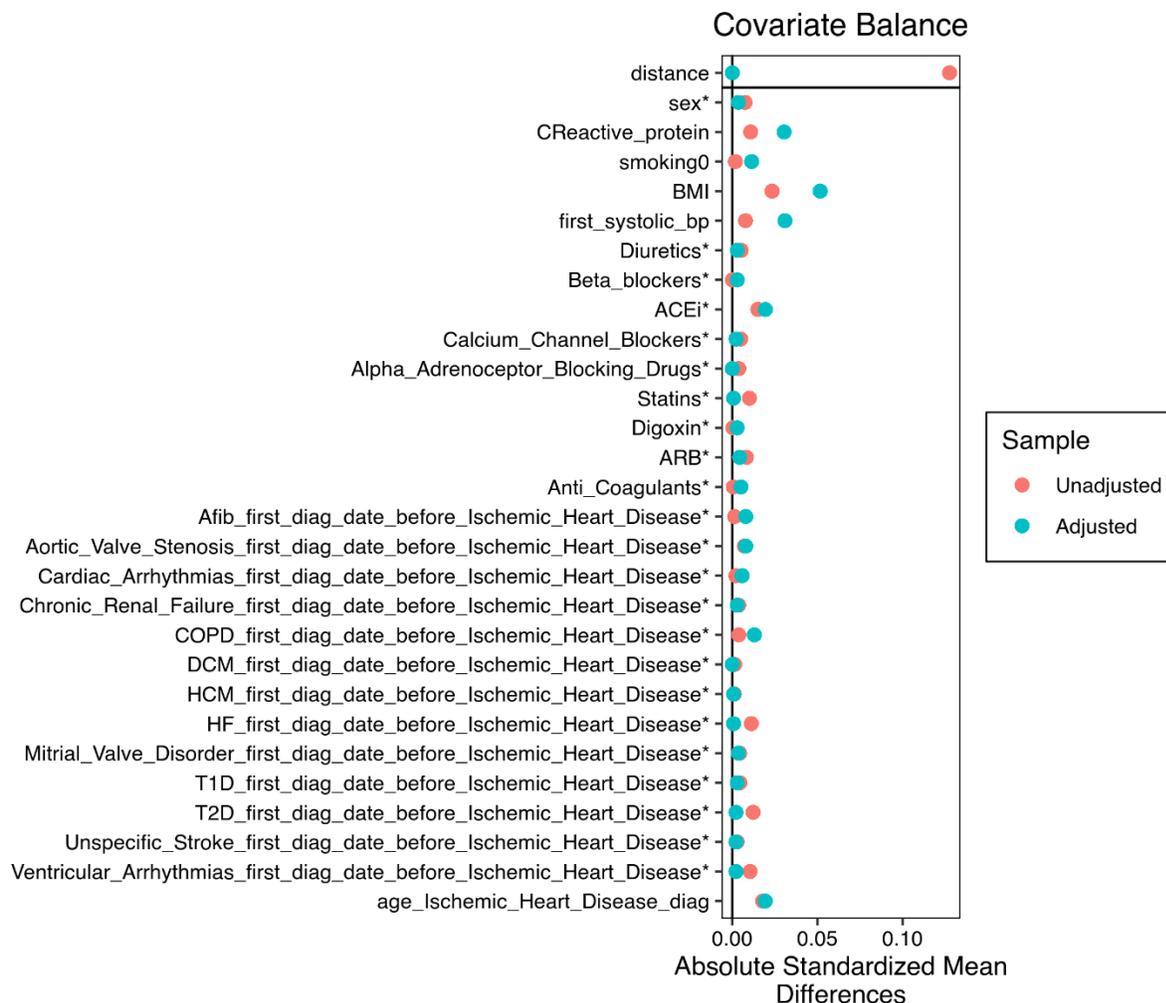
Supplemental Figure 3. Propensity score matching results for ACE rs4968782 for a) time to cardiovascular (CV) death/heart transplant from first atrial fibrillation (AF) diagnosis in Caucasians, and b) time to rehospitalization or CV death/heart transplant due to all-cause heart failure (HF) in all ethnicities. A shift towards 0 for absolute standardized mean difference represents successful matching.



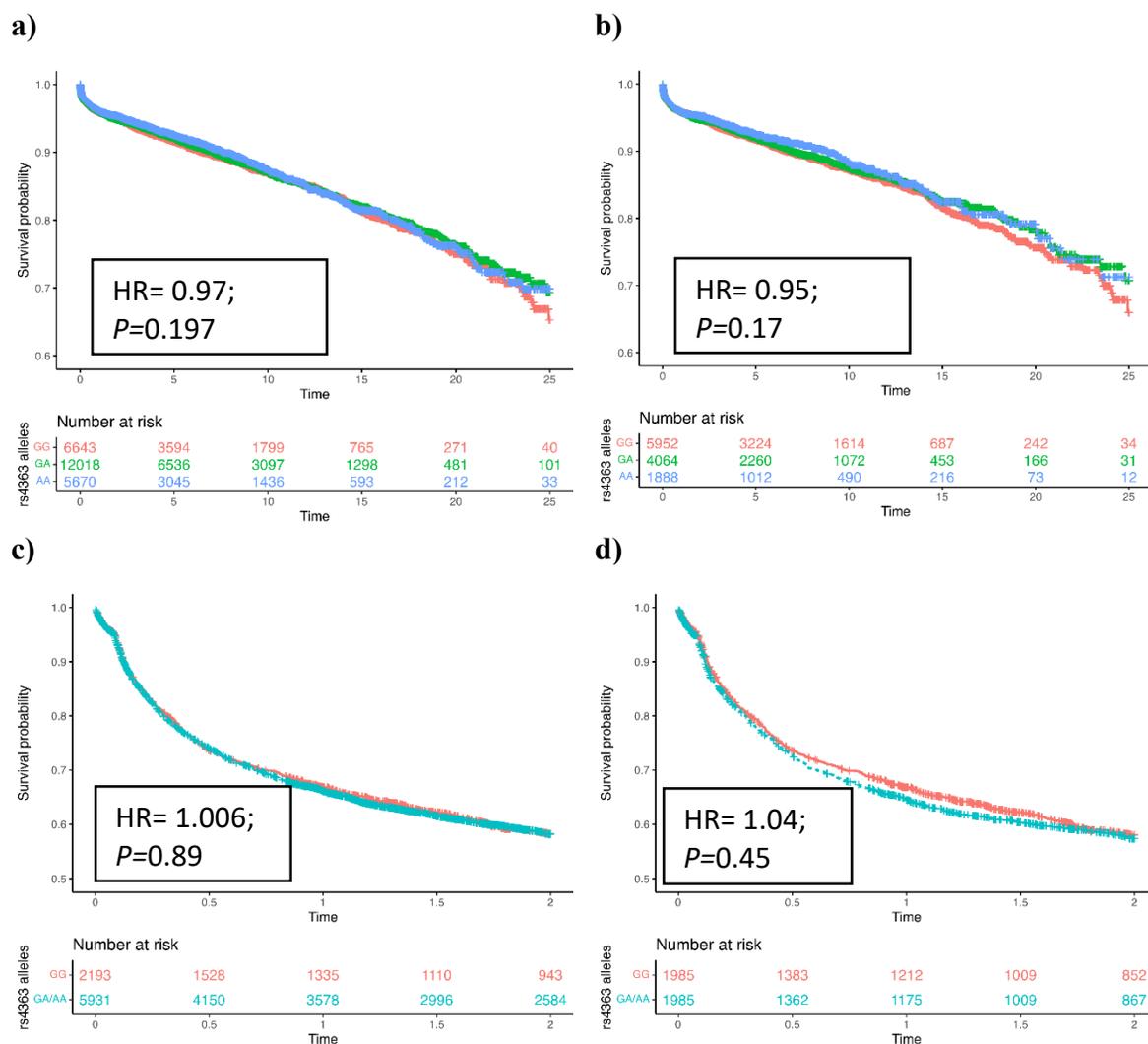
Supplemental Figure 4. Propensity score matching results for BAG3 for a) all-cause HF patients, and b) Ischemic HF patients. A shift towards 0 for absolute standardized mean difference represents successful matching.



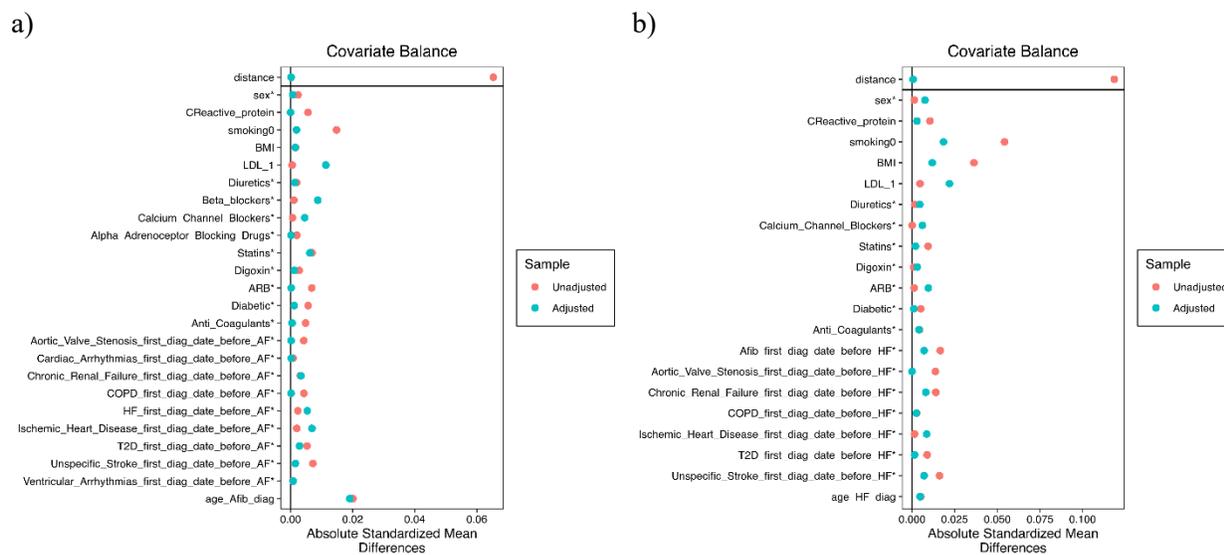
Supplemental Figure 5. Effect of PCSK9 rs562556 on time to cardiovascular (CV) death/heart transplant from first ischemic heart disease (IHD) diagnosis in a) unmatched and b) matched participants. Participants were filtered for Caucasian ethnicity. The time variable represents years to outcome while the allele variable represents control and treatment groups (those with GG genotype versus those with GA/AA genotype).



Supplemental Figure 6. Matching results for PCSK9 rs562556 in Caucasians. Shift towards 0 in absolute standardized mean difference indicates successful matching. Asterisk (*) indicates variables for which the displayed value is the raw (unstandardized) difference in means.



Supplemental Figure 7. Effect of ACE rs4363 on time to cardiovascular (CV) death/heart transplant from first atrial fibrillation (AF) diagnosis in a) unmatched and b) matched Caucasian participants, and time to rehospitalization or CV death/heart transplant due to all cause heart failure (HF) in c) unmatched and d) matched (all ethnicities) participants. The time variable represents years to severe cardiovascular outcome (death or transplant) while the allele variable represents control and treatment groups (GG genotype versus GA/AA genotype).



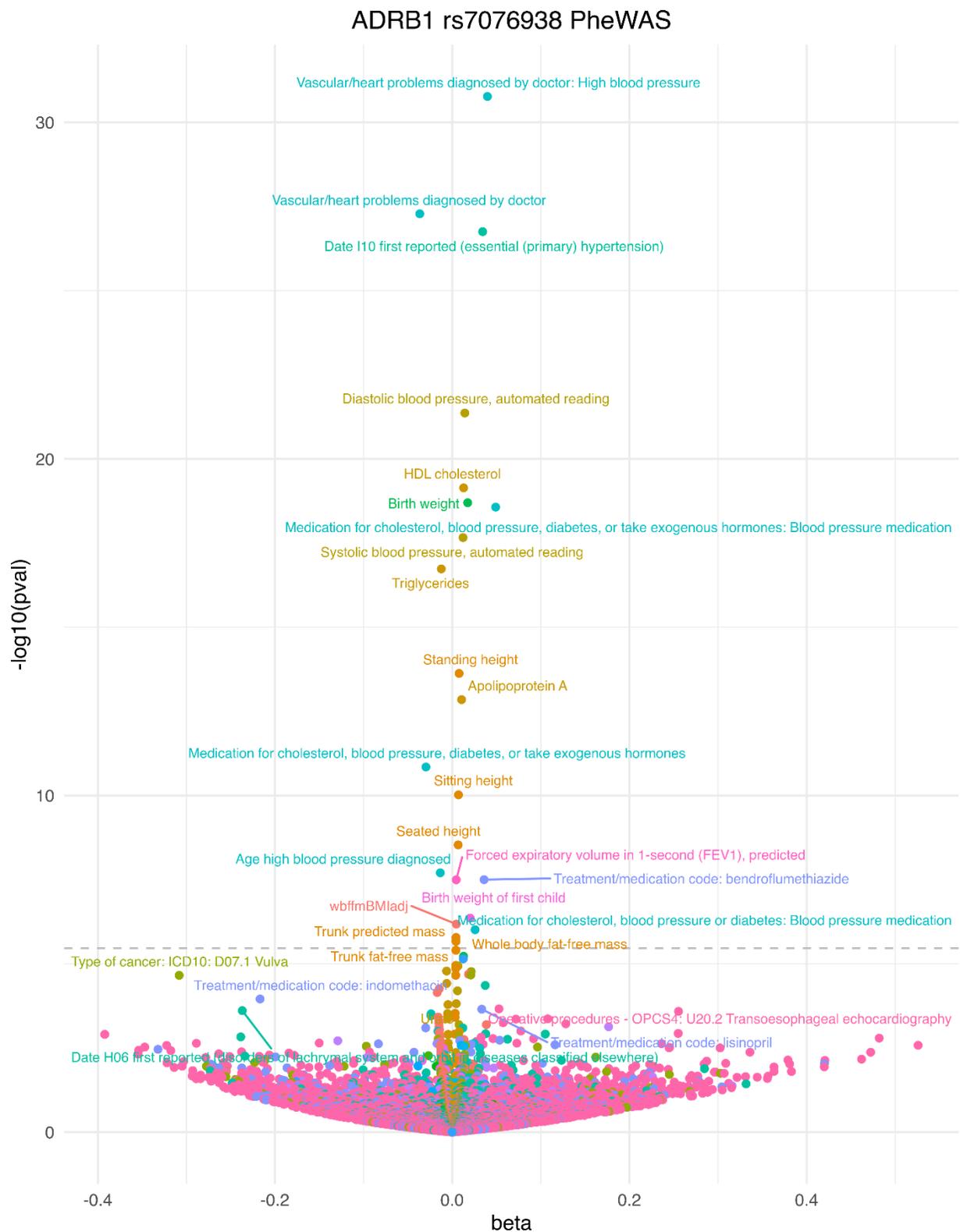
Supplemental Figure 8. Propensity score matching results for ACE rs4363 for a) time to cardiovascular (CV) death/heart transplant from first atrial fibrillation (AF) diagnosis in Caucasians, and b) time to rehospitalization or CV death/heart transplant due to all-cause heart failure (HF) in all ethnicities. A shift towards 0 for absolute standardized mean difference represents successful matching.



Supplemental Figure 9. PheWAS results for PCSK9 rs11591147 which includes all participants with available genotype data and adjusted for sex, age of recruitment, and ten genetic principal components.



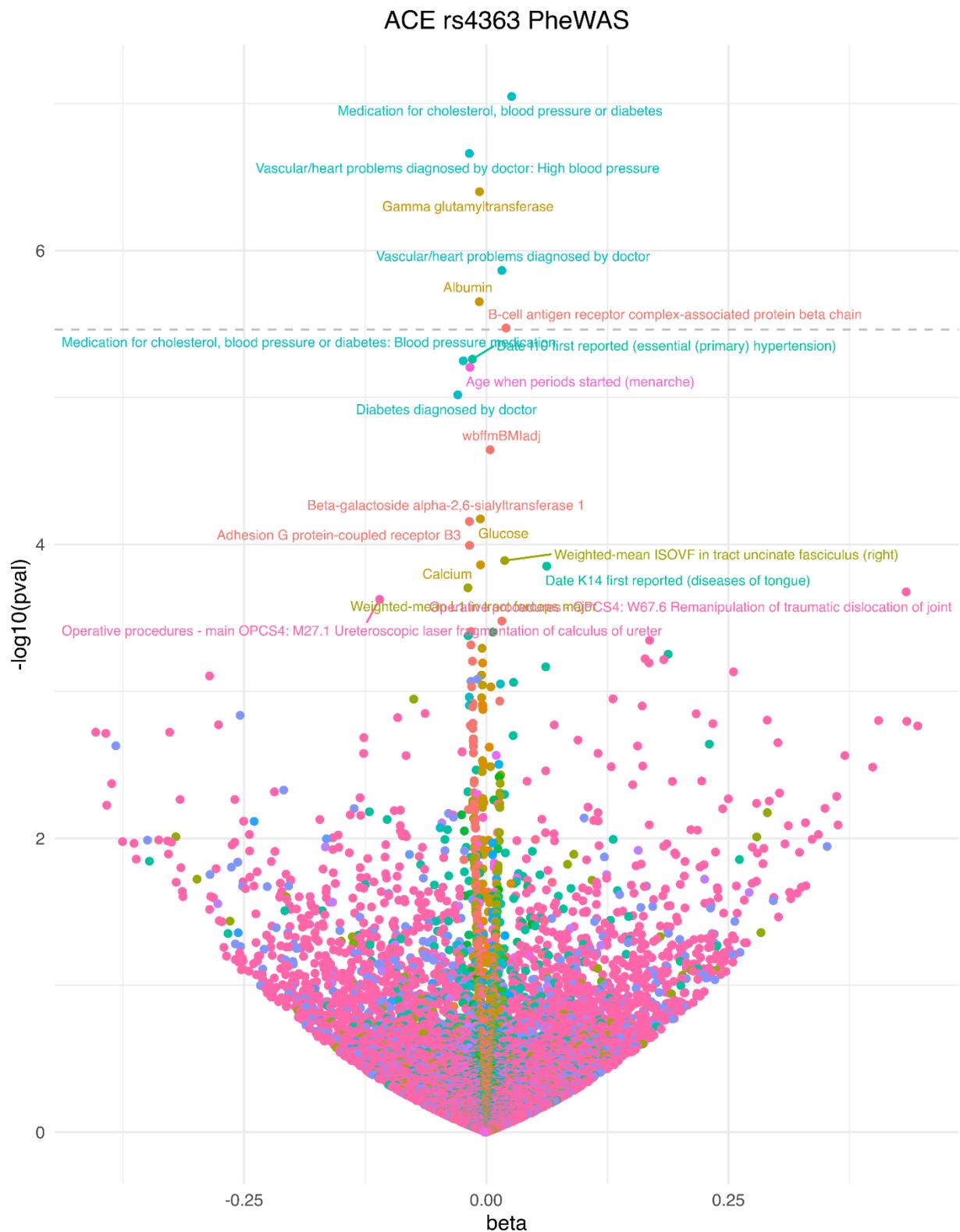
Supplemental Figure 10. PheWAS results for PCSK9 rs562556. Analysis for all participants with available genotype data, adjusted for sex, age of recruitment, and 10 genetic principal components.



Supplemental Figure 11. PheWAS results for ADRB1 rs7076938 (X allele). Analysis for all participants with available genotype data, adjusted for sex, age of recruitment, and 10 genetic principal components.



Supplemental Figure 12. PheWAS results for ACE rs4968782. Analysis for all participants with available genotype data, adjusted for sex, age of recruitment, and 10 genetic principal components.



Supplemental Figure 13. PheWAS results for ACE rs4363. Analysis for all participants with available genotype data, adjusted for sex, age of recruitment, and 10 genetic principal components.



Supplemental Figure 14 PheWAS results for BAG3 rs2234962. Analysis for all participants with available genotype data, adjusted for sex, age of recruitment, and 10 genetic principal components.