

1 **Clinical conditions and their impact on utility of genetic scores for prediction of acute
2 coronary syndrome**

3 Jiwoo Lee^{1,2,3,4}, Tuomo Kiiskinen⁴, Nina Mars⁴, Sakari Jukarainen⁴, FinnGen Project, Erik
4 Ingelsson¹, Benjamin Neale^{2,3}, Samuli Ripatti^{2,3,4}, Pradeep Natarajan^{2*}, Andrea Ganna^{2,3,4*}

5 [1] Department of Biomedical Data Science, Stanford University, Stanford, CA, USA

6 [2] Broad Institute of MIT and Harvard, Cambridge, MA, USA

7 [3] Analytical and Translational Genetics Unit, Massachusetts General Hospital, Boston, MA,
8 USA

9 [4] Finnish Institute for Molecular Medicine, HiLIFE, University of Helsinki, Helsinki, Finland

10 * Corresponding authors

11

1 Abstract

2 Aims

3 Early prediction of acute coronary syndrome (ACS) is a major goal for prevention of coronary
4 heart disease (CHD). Genetic information has been proposed to improve prediction beyond well-
5 established clinical risk factors. While polygenic scores (PS) can capture an individual's genetic
6 risk for ACS, its prediction performance may vary in the context of diverse correlated clinical
7 conditions. Here, we aimed to test whether clinical conditions impact the association between PS
8 and ACS.

9 Methods and Results

10 We explored the association between 405 clinical conditions diagnosed before baseline and
11 9,080 incident cases of ACS in 387,832 individuals from the UK Biobank. We identified 80
12 conventional (e.g., stable angina pectoris (SAP), type 2 diabetes mellitus) and unconventional
13 (e.g., diaphragmatic hernia, inguinal hernia) associations with ACS. Results were replicated in
14 6,430 incident cases of ACS in 177,876 individuals from FinnGen. The association between PS
15 and ACS was consistent in individuals with and without most clinical conditions. However, a
16 diagnosis of SAP yielded a differential association between PS and ACS. PS was associated with
17 a significantly reduced (interaction p-value=2.87×10⁻⁸) risk for ACS in individuals with SAP
18 (HR=1.163 [95% CI: 1.082–1.251]) compared to individuals without SAP (HR=1.531 [95% CI:
19 1.497–1.565]). These findings were replicated in FinnGen (interaction p-value=1.38×10⁻⁶).

20 Conclusion

21 In summary, while most clinical conditions did not impact utility of PS for prediction of ACS,
22 we found that PS was substantially less predictive of ACS in individuals with prevalent stable

1 CHD. PS for ACS may be more appropriate for asymptomatic individuals than symptomatic
2 individuals with clinical suspicion for CHD.

3 **Introduction**

4 Acute coronary syndrome (ACS), including myocardial infarction, is an unstable
5 consequence of coronary heart disease (CHD), the leading cause of death worldwide. In Europe,
6 CHD accounts for 1.8 million deaths per year, making up almost 20% of all cardiovascular
7 disease-related deaths (1). Thus, prediction of ACS remains a major public health issue for
8 prevention of CHD.

9 Well-established clinical risk factors, such as type 2 diabetes mellitus and
10 hypercholesterolemia, are highly predictive of future health outcomes, including ACS (5, 6).
11 Individuals with an accumulation of such risk factors are candidates for preventative statin
12 interventions per American and European guidelines (5-7). Recently, other risk-enhancing
13 factors independently associated with ACS have been proposed to improve prediction of ACS
14 and inform statin intervention decisions (8). An unbiased, comprehensive evaluation of diverse
15 correlated clinical conditions in large-scale biobanks may better inform prediction of ACS (9-
16 11).

17 Furthermore, genetic information may inform prediction of ACS beyond well-established
18 clinical risk factors (12-16). Currently, guidelines on evaluating risk for ACS and initiating statin
19 interventions rely on several clinical risk factors but do not support the use of genetic
20 information. However, genetic information has the advantage of remaining stable throughout life
21 and therefore could be used to improve earlier prediction of ACS (17-19). Results from large
22 genome-wide association studies (GWAS) can be used to derive polygenic scores (PS) based on

1 millions of single nucleotide polymorphisms (SNPs) that are robustly associated with CHD (20).
2 Several studies have investigated the use of PS for prediction of ACS in addition to well-
3 established clinical risk factors (16, 21-23). However, there are no studies to date that examine
4 the association between PS and ACS in the context of diverse correlated clinical conditions.

5 In this study, we have two main goals that we assessed in two large-scale biobanks. First,
6 we aimed to comprehensively explore the association of clinical conditions with ACS. Second,
7 we aimed to assess whether clinical conditions associated with ACS impact the association
8 between PS and ACS to provide context for the use of PS in a clinical setting for prediction of
9 ACS.

10 Methods

11 *Study population*

12 This study utilized data from the UK Biobank and replicated findings in FinnGen. The
13 UK Biobank is a prospective cohort study that includes approximately 500,000 individuals
14 between ages 40- to 69-years old enrolled between 2006 and 2010 (24, 25). The baseline visit,
15 defined as the date of assessment center visit, included questionnaires and interviews that
16 captured sociodemographic and health status, physical measurements, and phlebotomy
17 measurements. Detailed information and measurements for each individual were recorded.
18 Methods for follow-up included linking individuals to national hospital inpatient and outpatient
19 records, death registries, and primary care diagnoses (25). We excluded individuals with a
20 history of CHD at baseline. Detailed exclusion criteria are summarized in **Supplementary Table**
21 **1.** In total, 387,832 individuals were considered. Age was calculated from the date of assessment
22 center visit (data field 53) and the date of birth (data field 33). We adjusted our analysis for sex

1 (data field 31). Measured risk factors considered in our analysis included body mass index (BMI,
2 data field 21001), systolic blood pressure (SBP, data field 4080), smoking status (data field
3 20116), high-density lipoprotein cholesterol (HDL cholesterol, data field 30760), and total
4 cholesterol (data field 30690). Analyses included all UK Biobank participants with relevant
5 information and did not exclude individuals of non-European ancestry. All UK Biobank
6 participants were included because the goal of the study was not to assess the prediction
7 performance of PS per se, which was expected to be lower in individuals of non-European
8 ancestry than in individuals of European ancestry. Rather, the goal of the study was to examine
9 the association between PS and ACS across different clinical conditions. In a sensitivity analysis,
10 we examined this association separately for each ancestry group defined using the same
11 approach as used in the Pan-UK Biobank project (<https://pan.ukbb.broadinstitute.org/>).

12 FinnGen is a prospective cohort study that includes approximately 224,000 individuals
13 representing about 3% of the Finnish adult population (21). Individuals were linked to national
14 hospital discharge, death, and medication reimbursement registries. The baseline visit was
15 defined as the date of DNA sample collection. Similar exclusion criteria were applied to
16 FinnGen, in which 177,876 individuals were considered.

17 ***Clinical conditions and outcomes***

18 In the UK Biobank, clinical conditions were defined by the International Classification of
19 Diseases, Tenth Revision (ICD-10) obtained from the Hospital Episode Statistics (HES) database
20 which covers inpatient (1997–present) and outpatient (2003–present) admissions to hospitals. In
21 individuals who developed ACS, we only considered clinical conditions that were diagnosed
22 before baseline and at least 30 days prior to an ACS event to remove clinical conditions that may
23 be captured or confounded by an ACS event. We excluded clinical conditions diagnosed after

1 baseline to reflect a clinical setting in which prediction of ACS is assessed using past medical
2 history. We excluded clinical conditions involving external causes of morbidity and mortality
3 and factors influencing health status and contact with health services (ICD-10 chapters XX and
4 XXI). In the case where an individual received multiple diagnoses for the same clinical
5 condition, only the date of the first diagnosis was considered. We only considered clinical
6 conditions with at least 10 co-occurrences with ACS to improve statistical power. Detailed
7 exclusion criteria are summarized in **Supplementary Table 1**. ACS was algorithmically defined
8 with ICD-10 codes capturing unstable angina (I20.0), acute myocardial infarction (I21), and
9 subsequent ST elevation (STEMI) and non-ST elevation (NSTEMI) myocardial infarction (I22).

10 In FinnGen, clinical conditions were based on inpatient (1995-present) and outpatient
11 (1998-present) admissions to hospitals. The same definition of ACS and exclusion criteria from
12 the UK Biobank were used in FinnGen. From the UK Biobank, 9,080 cases of ACS and 405
13 ICD-10 codes were considered after exclusion. In FinnGen, 6,430 cases of ACS and 645 ICD-10
14 codes were considered after exclusion. Baseline characteristics of each cohort are summarized in
15 **Table 1**.

16 **Polygenic scores**

17 PS was constructed using PRS-CS. Briefly, PRS-CS is a polygenic prediction method
18 that uses Bayesian regression and infers posterior SNP effect sizes under continuous shrinkage
19 priors using only GWAS summary statistics and an external linkage disequilibrium reference
20 panel (26). PRS-CS was validated in several traits and common complex diseases, including
21 CHD, and was robust and comparable to other well-established methods. To construct PS for
22 individuals in the UK Biobank and FinnGen, we used summary statistics from the largest GWAS
23 of CHD to date (20). This GWAS excluded individuals from the UK Biobank and FinnGen and

1 included 60,801 cases and 123,504 controls. Case status was defined by an inclusive diagnosis of
2 CHD. Summary statistics and the 1000 Genomes EUR reference panel were used to output
3 weights using standard PRS-CS settings. Only HapMap3 variants were included. PS was
4 calculated using PLINK 2.0 (27).

5 ***Statistical analysis***

6 First, we associated 405 clinical conditions diagnosed before baseline with ACS in the
7 UK Biobank using Cox regression survival models, modeling clinical conditions as time-varying
8 covariates. Baseline was defined as the date of assessment center visit. The end of follow-up date
9 was defined as an ACS event, death, or our pre-defined end of follow-up date, March 31, 2020.
10 Models were adjusted for age at baseline, sex, and principal components to control for ancestry.
11 We adjusted subsequent models for measured risk factors, including BMI, SBP, smoking status,
12 HDL cholesterol, and total cholesterol. This analysis was replicated in FinnGen with 645 clinical
13 conditions. All models were evaluated for statistical significance after multiple testing
14 correction.

15 Next, we looked at the interaction between clinical conditions and PS in models testing
16 for the association with ACS. Only clinical conditions that were associated with ACS after
17 multiple testing correction were considered (UK Biobank p-value $<1.23\times10^{-4}$, FinnGen p-
18 value $<7.75\times10^{-5}$). Among clinical conditions with significant interactions, we explored whether
19 there were differences in the prediction performance between individuals with and without these
20 clinical conditions. We looked at the effect sizes (hazard ratio), discrimination (Harrell's C-
21 index), and net reclassification index (NRI) to evaluate the association and the prediction
22 performance (28, 29). To calculate NRI, we used four risk categories of 0-5%, 5-7.5%, 7.5-20%,
23 and >20% based on the 10-year risk for ACS including measured risk factors (12, 30, 31).

1 Results

2 *Several clinical conditions are associated with acute coronary syndrome*

3 In the UK Biobank, we explored the association between 405 clinical conditions
4 diagnosed before baseline and 9,080 cases of ACS and identified 80 clinical conditions that were
5 significantly ($p\text{-value} < 1.23 \times 10^{-4}$) associated with ACS after multiple testing correction (**Figure 1**). The 80 associated clinical conditions had an average hazard ratio of 2.4 for ACS. Some of
6 these associations were conventional (e.g., stable angina pectoris (SAP), type 2 diabetes mellitus)
7 and others were unconventional (e.g., diaphragmatic hernia, inguinal hernia). Adjustment for
8 measured risk factors modestly reduced the association, but statistical significance was retained
9 for all 80 clinical conditions. **Supplementary Table 2** includes all 80 clinical conditions that
10 were associated with ACS before and after adjustment for measured risk factors. This analysis
11 was replicated in FinnGen, in which we explored the association between 645 clinical conditions
12 and 6,430 cases of ACS and identified 71 clinical conditions that were significantly ($p\text{-}$
13 value $< 7.75 \times 10^{-5}$) associated with ACS after multiple testing correction (**Supplementary Figure 1**). The 71 associated clinical conditions had an average hazard ratio of 2.1 for ACS.
14 **Supplementary Table 3** includes all 71 clinical conditions that were associated with ACS.
15 Thirty-three clinical conditions were associated with ACS in both the UK Biobank and FinnGen.
16 For these clinical conditions, we observed good consistency ($R^2 = 0.65$) between the association
17 observed in the UK Biobank and FinnGen, suggesting the generalizability of these associations
18 across two large-scale biobanks (**Supplementary Figure 3**).
19
20
21 **Most clinical conditions do not impact association between polygenic scores and acute**
22 **coronary syndrome except for stable angina pectoris**

1 First, we confirmed that PS was associated with ACS in the UK Biobank (HR for 1 SD
2 increase in PS=1.49 [95% CI: 1.46–1.53], p-value= 1.36×10^{-294}) and FinnGen (HR for 1 SD
3 increase in PS=1.44 [95% CI: 1.40–1.47], p-value= 1.71×10^{-166}). Next, we aimed to understand
4 the value of measuring PS in individuals with or without clinical conditions associated with ACS
5 and how the association between PS and ACS changed given the diagnosis of certain clinical
6 conditions. We tested for the interaction between PS and 80 clinical conditions that were
7 associated with ACS in the UK Biobank and identified a significant interaction (p-
8 value= 2.87×10^{-8}) between SAP and PS (**Figure 2**). Here, SAP was defined as angina pectoris
9 with documented spasm, other forms of angina pectoris, and unspecified angina pectoris (ICD-10
10 codes I20.1, I20.8, I20.9). This was the only interaction (p-value= 1.38×10^{-6}) that was replicated
11 in FinnGen (**Supplementary Figure 2**). In the UK Biobank, there were 5,477 cases of SAP
12 (prevalence=1.41%), and in FinnGen, there were 8,326 cases of SAP (prevalence=4.68%). There
13 was also a significant interaction between PS and type 2 diabetes mellitus (p-value= 6.01×10^{-6}) in
14 the UK Biobank, but this interaction was not significant after multiple testing correction in
15 FinnGen (p-value= 5.00×10^{-2}).

16 In the UK Biobank, individuals with SAP had a significantly reduced risk for ACS (HR
17 for 1 SD=1.163 [95% CI: 1.082–1.251]) compared to individuals without SAP (HR for 1
18 SD=1.531 [95% CI: 1.497–1.565]) (**Figure 3**). In a model including measured risk factors in the
19 UK Biobank, results remained consistent (SAP HR=1.159 [95% CI: 1.078–1.246], no SAP
20 HR=1.488 [95% CI: 1.455–1.522], interaction p-value= 2.05×10^{-7}). This observation was
21 replicated in FinnGen, and individuals with SAP had a significantly reduced risk for ACS (HR
22 for 1 SD=1.247 [95% CI: 1.173–1.326]) compared to individuals without SAP (HR for 1
23 SD=1.448 [95% CI: 1.407–1.490]).

1 We also noticed that the association between measured risk factors and ACS were
2 generally weaker among individuals with SAP than individuals without SAP (**Supplementary**
3 **Table 4**). For example, there was no association between total cholesterol and ACS in
4 individuals with SAP (HR=0.96 [95% CI: 0.89–1.03], p-value=0.247), but there was a
5 significant association in individuals without SAP (HR=1.11 [95% CI: 1.09–1.13], p-
6 value= 6.43×10^{-28})

7 ***Sensitivity analyses***

8 We performed several sensitivity analyses. First, because PS has been shown to have age-
9 dependent effects (21) and individuals with SAP tended to be older and include more males than
10 individuals without SAP, we performed an age- and sex-matched analysis to test whether
11 differences in age or sex explained the differential association of PS and ACS. After matching
12 age and sex distributions in individuals with and without angina, the interaction between SAP
13 and PS remained significant in both the UK Biobank (p-value= 1.51×10^{-3}) and FinnGen (p-
14 value= 1.16×10^{-2}).

15 Second, because many well-established guidelines suggest statin interventions for
16 individuals with SAP, we adjusted our analysis for statin use to explore whether statin use in
17 individuals with SAP lowered the association with ACS. Statin use in the UK Biobank was self-
18 reported and statin use in FinnGen was defined using high-quality prescription registries. In the
19 UK Biobank, 82.1% of individuals with SAP used statins and 15.9% of individuals without SAP
20 used statins. In FinnGen, 81.7% of individuals with SAP used statins and 27.8% of individuals
21 without SAP used statins. The interaction between SAP and PS remained significant after
22 adjusting for statin use in both the UK Biobank (p-value= 1.20×10^{-10}) and FinnGen (p-

1 value= 3.77×10^{-4}). In the UK Biobank, the interaction between SAP and PS remained significant
2 (p-value= 7.39×10^{-4}) after removing statin users.

3 Third, we increased the time window between a diagnosis of SAP and an ACS event to
4 reduce the chances of capturing the same underlying cardiovascular event. Increasing the time
5 window from 30 days to 300 and 500 days only modestly reduced the interaction between SAP
6 and PS in the UK Biobank (300 day p-value= 5.14×10^{-11} , 500 day p-value= 3.76×10^{-11}) and
7 FinnGen (300 day p-value= 5.83×10^{-4} , 500 day p-value= 1.86×10^{-3}). Average and median time
8 between an SAP diagnosis and an ACS event was 9.1 and 8.9 years, respectively, in the UK
9 Biobank. Similarly, average and median time was 6.7 and 5.7 years, respectively, in FinnGen.
10 This suggested that the cases of SAP captured in our study were diagnosed earlier and
11 independent of ACS.

12 Fourth, we examined the interaction between SAP and ACS in different ancestry groups.
13 Our cohort from the UK Biobank included individuals of African (AFR, n = 5,543), Ad Mixed
14 American (AMR, n = 820), Central and South Asian (CSA, n = 6,976), East Asian (EAS, n =
15 2,196), European (EUR, n = 339,771), and Middle Eastern (MID, n = 1,258) ancestry. In CSA
16 (ACS cases = 306, SAP diagnoses = 194, p-value = 1.14×10^{-2}) and EUR (ACS cases = 7,969,
17 SAP diagnoses = 4,810, p-value = 2.36×10^{-11}), the interaction between SAP and PS remained
18 significant. Other ancestry groups, which included less than 100 diagnoses of SAP and 100 cases
19 of ACS, were not analyzed due to the lack of statistical power.

20 ***Prediction performance of polygenic scores differs between individuals with and without***
21 ***stable angina pectoris***

1 We examined Harrell's C-index to evaluate the prediction performance of PS (**Figure 4**).
2 In the UK Biobank, individuals with SAP had a C-index for PS of 0.606 [95% CI: 0.595–0.616]
3 and individuals without SAP had a C-index for PS of 0.747 [95% CI: 0.743–0.750] when
4 considering models adjusted for age, sex, principal components, BMI, SBP, smoking status,
5 HDL cholesterol, and total cholesterol. These findings were replicated in FinnGen, and
6 individuals with SAP had a C-index of 0.665 [95% CI: 0.655–0.674] and individuals without
7 SAP had a C-index of 0.825 [95% CI: 0.821–0.830]. The lower C-index in individuals with SAP
8 was explained by the lower association between established cardiovascular risk factors, including
9 age, and ACS in these individuals.

10 Because it is easier to obtain a higher improvement in predictive performance (Δ C-index)
11 when the baseline model has a lower prediction performance (32, 33), we expected the
12 improvement in C-index due to PS to be higher in individuals with SAP. Nevertheless, the
13 addition of PS to measured risk factors increased the C-index more in individuals without SAP
14 (Δ C-index=0.021) than in individuals with SAP (Δ C-index=0.010). In FinnGen, PS increased the
15 C-index by 0.012 in individuals without SAP and by 0.013 in individuals with SAP. Lastly, we
16 calculated the NRI to evaluate the improvement in reclassification when adding PS to the
17 baseline prediction performance of measured risk factors (**Supplementary Table 5**). Addition of
18 PS significantly improved the prediction performance in individuals without SAP (categorical
19 NRI=0.116, p-value<1.00×10⁻³; continuous NRI=0.315, p-value<1.00×10⁻³) but not in
20 individuals with SAP (NRI=0.038, p-value=0.187; continuous NRI=0.113, p-value=3.41×10⁻³).
21 Overall, the NRI was higher among individuals without SAP compared to individuals with SAP.

22 **Discussion**

1 In this study, we performed an unbiased, comprehensive evaluation of the association
2 between diverse correlated clinical conditions and ACS across two large-scale biobanks.
3 Moreover, by exploring the impact of clinical conditions on the association between PS and
4 ACS, we provided context for the use of PS in a clinical setting for prediction of ACS.

5 First, we found that a large number of clinical conditions were associated with ACS,
6 independent of measured risk factors. Using clinical conditions to predict risk for ACS may be
7 more convenient than using measured risk factors because this information is easily accessed and
8 obtained from electronic health records and does not require in-person visits. In addition to well-
9 established clinical risk factors, such as type 2 diabetes mellitus, that are routinely considered,
10 additional clinical conditions that are associated with ACS may be taken into consideration when
11 evaluating risk for ACS. Identification of such clinical conditions is critical to provide a
12 comprehensive, well-informed assessment of risk for ACS.

13 Second, by using a hypothesis-free approach across 80 clinical conditions associated with
14 ACS, we found that the association between PS and ACS was consistent across individuals with
15 or without most of these clinical conditions. Diagnosis of a clinical condition did not change the
16 association between PS and ACS, suggesting that clinical conditions do not impact the utility of
17 PS. Previous work has shown that different groups of individuals may benefit more from
18 measuring PS than other groups. For example, previous work has shown that PS was more
19 predictive of ACS among individuals who had never smoked (34). However, in our study, we
20 generally found that the diagnosis of a clinical condition did not affect the association between
21 PS and ACS. The only differential association that was found in both the UK Biobank and
22 FinnGen was a diagnosis of SAP, where SAP was significantly more associated with ACS in

1 individuals without a previous diagnosis of SAP than in individuals with a previous diagnosis of
2 SAP.

3 There are at least two possible explanations for this phenomenon. First, progression of
4 ACS occurs with or without the presentation of SAP. It is possible that individuals who develop
5 ACS with or without prior SAP are clinically and genetically distinct. It is widely accepted that
6 ACS varies in clinical presentation, and other studies have explored the significance of a
7 diagnosis of SAP before ACS (35-38). For example, other studies have found that individuals
8 with prior SAP had more extensive atherosclerosis and a greater prevalence of clinical risk
9 factors, such as hypertension, than individuals without prior SAP who developed ACS (39, 40).

10 In another study, myocardial infarction was the most common first clinical presentation in
11 individuals without prior SAP who developed ACS (35). Other studies have also found that
12 individuals with prior SAP had a more favorable prognosis compared to individuals without
13 prior SAP who developed ACS (39, 41). In our study, we found a differential association
14 between PS and ACS in individuals with and without SAP, further suggesting that these
15 individuals may be clinically and genetically distinct. Taken together, there may be different
16 pathways of progression of ACS that may warrant different assessments and interventions.

17 Second, a diagnosis of SAP may be synonymous to a diagnosis of CHD. In a clinical
18 setting, individuals with SAP often have already started developing CHD. Subsequently,
19 individuals with SAP may already have a higher risk for ACS, and PS may not be informative.
20 As a result, PS may lose its prediction performance in individuals who have already been
21 diagnosed with SAP. This observation is supported by other studies that have shown that
22 prediction of secondary CHD events by PS is attenuated compared to that of primary CHD
23 events (42).

1 An open matter that may explain this phenomenon is index bias. Index bias may occur in
2 studies that select patients based on the occurrence of an index event, and as a result, measured
3 risk factors may have less of an effect on recurrent events (43-45). In our study, SAP may be an
4 index event that decreased the prediction performance of PS in individuals with SAP compared
5 to individuals without SAP. In other words, because ACS is dependent on multiple factors, such
6 as underlying chronic conditions and measured risk factors, conditioning on a diagnosis of SAP
7 creates a dependence between such factors. As a result, individuals with SAP do not require the
8 same burden of increased measured risk factors and PS to develop ACS. As mentioned
9 previously, individuals with SAP are more likely to have underlying chronic conditions, such as
10 extensive atherosclerosis. This may entail that measured risk factors and PS do not need to have
11 a large association with ACS to confer the same risk, due to the presence of such underlying
12 chronic conditions. Indeed, our study showed that measured risk factors and PS were less
13 associated with ACS in individuals with SAP, reflecting other studies that have also shown that
14 individuals with SAP often have less associated measured risk factors (35). Thus, PS becomes
15 less predictive of ACS in individuals with SAP than in individuals without SAP.

16 A few limitations of this study must be considered. First, definition of ACS and clinical
17 conditions were based on diagnostic codes derived from national health registries.
18 Underreporting and incorrect use of diagnostic codes in a clinical setting are inherent to these
19 types of data. Nevertheless, we obtained similar results in two large-scale biobanks from the
20 United Kingdom and Finland. Additional evidence from other studies also support the quality of
21 this type of data (46-48). Future studies are needed to fully assess the validity of using registry-
22 based clinical conditions across large-scale biobanks from different countries, but our study
23 supports the feasibility of cross-nation registry-based analyses. Second, our PS represents a

1 “snapshot” of current research in genetic scores. PS is inherently tied to the study used to
2 construct the genetic scores, and as larger GWAS are conducted, these genetic scores will
3 necessarily improve and become more predictive. Therefore, our results represent a lower bound
4 of prediction performance of PS, and it is likely that additional studies that improve genetic
5 scores will improve the prediction performance of PS.

6 **Conclusion**

7 Using an unbiased, comprehensive approach, we identified clinical conditions that were
8 associated with ACS, independent of measured risk factors. Overall, we found that the
9 association between PS and ACS remained consistent across many clinical conditions,
10 suggesting that the utility of PS for prediction of ACS is largely independent of previous clinical
11 conditions experienced by the patient. Nevertheless, we found one exception: individuals without
12 SAP had a greater association between PS and ACS than individuals with SAP. In conclusion,
13 we contextualized the use of PS in a clinical setting and showed that a PS for ACS may be more
14 appropriate among asymptomatic individual than symptomatic individuals with clinical
15 suspicion for CHD.

1 Tables

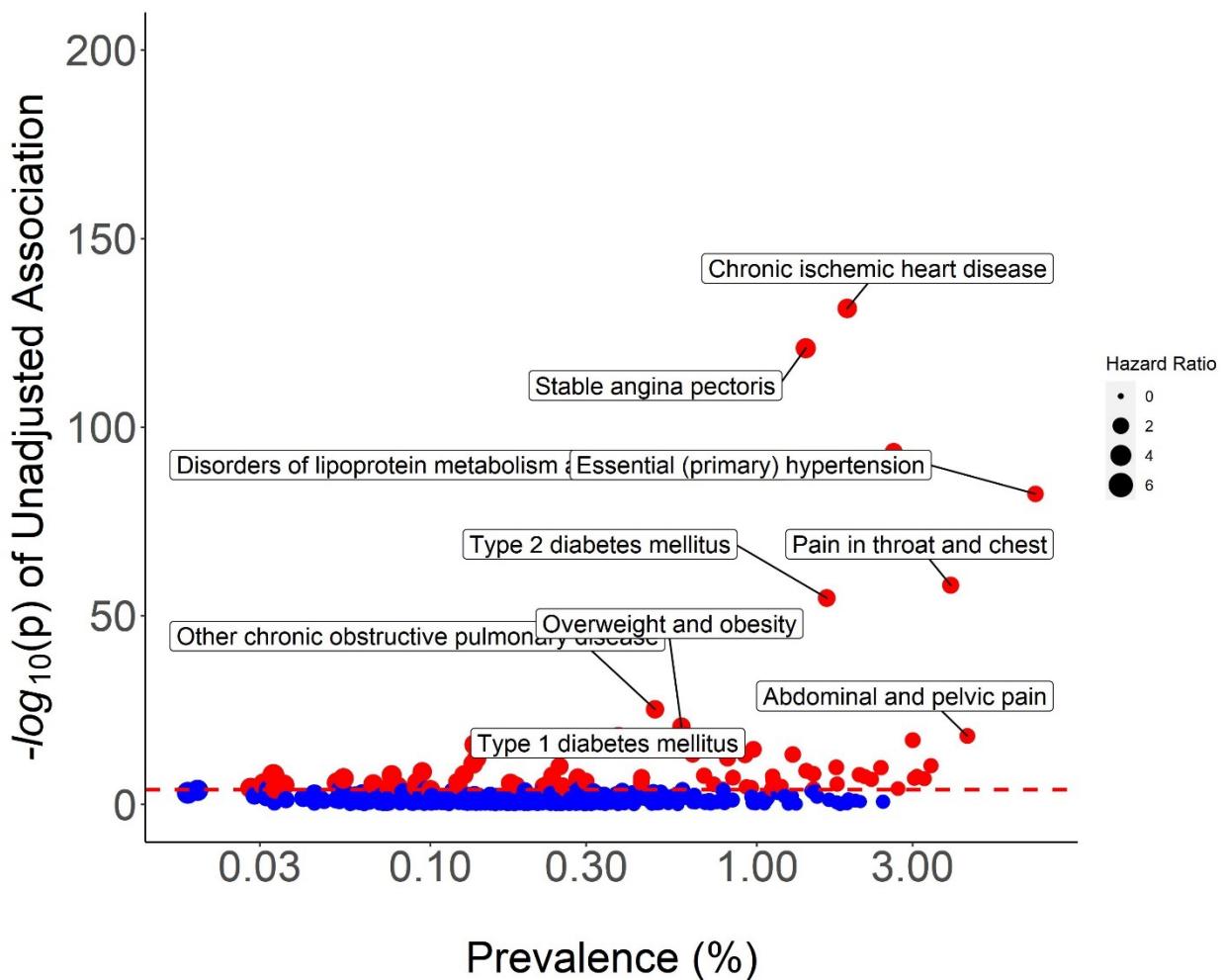
2 **Table 1:** Summary of baseline characteristics of cohorts from UK Biobank and FinnGen.

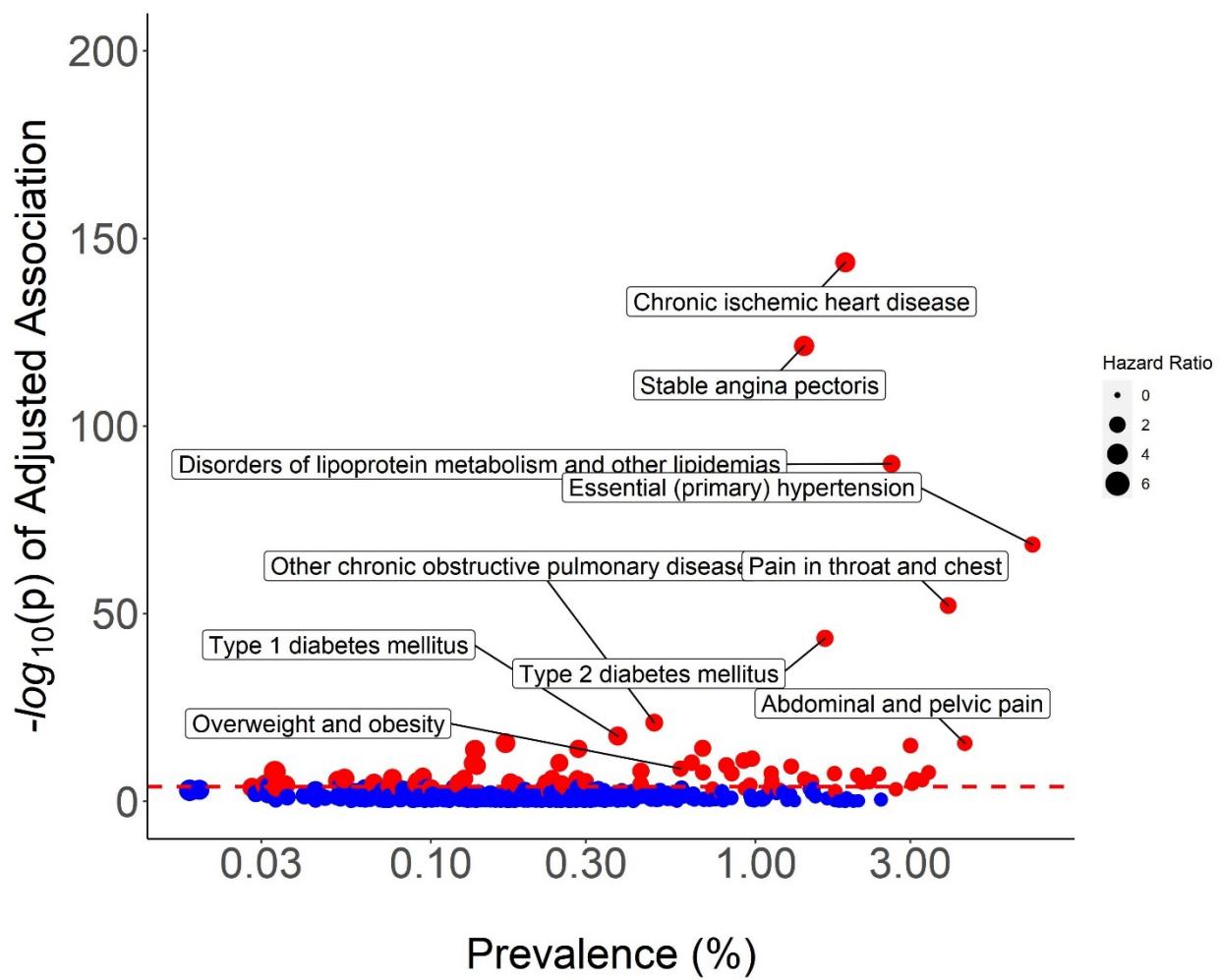
| | UK Biobank | FinnGen |
|---|----------------|----------------|
| Number of individuals without ACS at baseline | 387832 | 177876 |
| Number of ACS cases (%) | 9080 (2.34) | 6430 (3.61) |
| Number angina cases (%) | 5477 (1.41) | 8326 (4.68) |
| Average follow-up time in years (SD) | 11.97 (1.36) | 7.18 (8.29) |
| Average age (SD) | 56.47 (8.09) | 51.29 (17.33) |
| Number of females (%) | 210340 (54.23) | 100929 (56.74) |
| Number of statin users (%) | 65419 (16.87) | 53886 (30.29) |
| Number of smokers (%) | 174763 (45.06) | - |
| Average body mass index (SD) | 27.38 (4.75) | - |
| Average systolic blood pressure (SD) | 139.8 (19.64) | - |
| Average high-density lipoprotein cholesterol (SD) | 1.45 (0.38) | - |
| Average total cholesterol (SD) | 5.71 (1.13) | - |

3

1 Figures

2 **Figure 1:** Association between 405 clinical conditions and 9,080 cases of ACS unadjusted (top)
3 and adjusted (bottom) for measured risk factors from the UK Biobank. We used time-varying
4 covariate Cox survival models adjusted for age, sex, and principal components. Measured risk
5 factors included BMI, SBP, smoking status, HDL cholesterol, and total cholesterol. Red dotted
6 line is the significance threshold after multiple testing correction ($p\text{-value} < 1.23 \times 10^{-4}$). Top 10
7 clinical conditions are labeled.

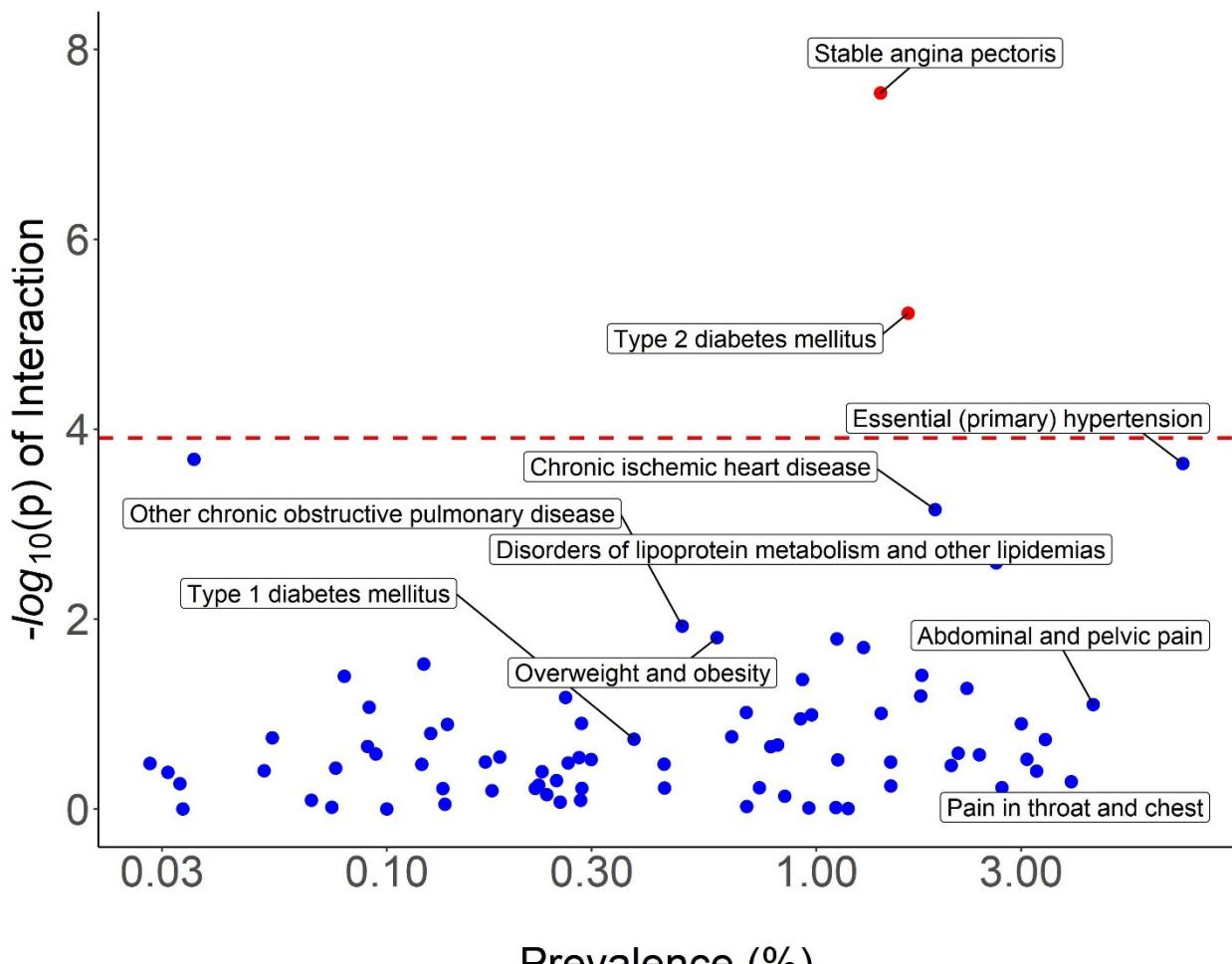




1

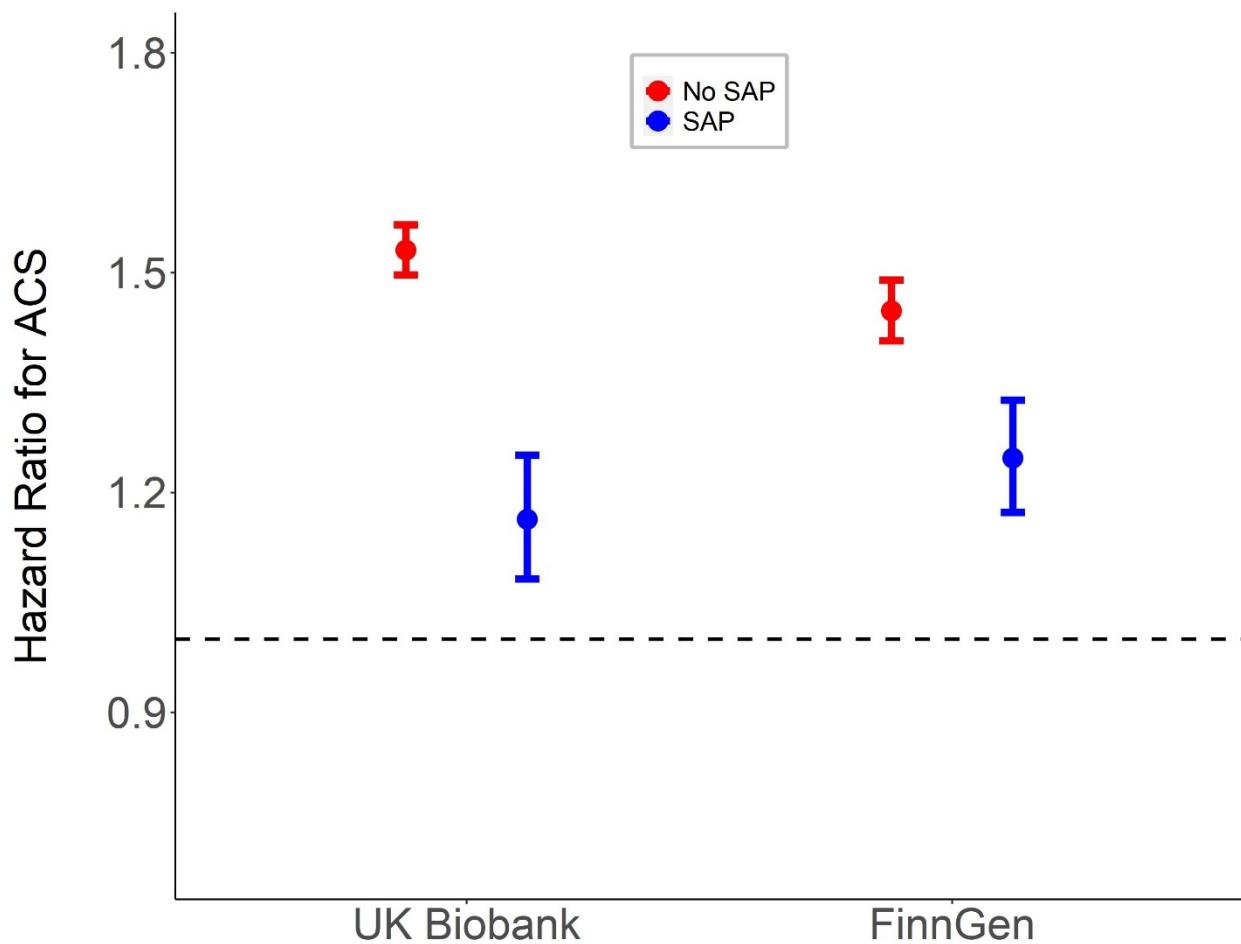
2

1 **Figure 2:** Interaction between PS and 80 clinical conditions significantly ($p\text{-value}<1.23\times10^{-4}$)
2 associated with ACS after multiple testing correction from the UK Biobank. Red dotted line is
3 the significance threshold ($p\text{-value}<1.23\times10^{-4}$). Top 10 clinical conditions are labeled. There was
4 a significant interaction between SAP and PS in association with ACS ($p\text{-value}=2.87\times10^{-8}$) that
5 was replicated in FinnGen ($p\text{-value}=1.38\times10^{-6}$). There was also a significant interaction between
6 SAP and type 2 diabetes mellitus ($p\text{-value}=6.01\times10^{-6}$) in the UK Biobank, but this interaction
7 was not significant after multiple testing correction in FinnGen ($p\text{-value}=5.00\times10^{-2}$).

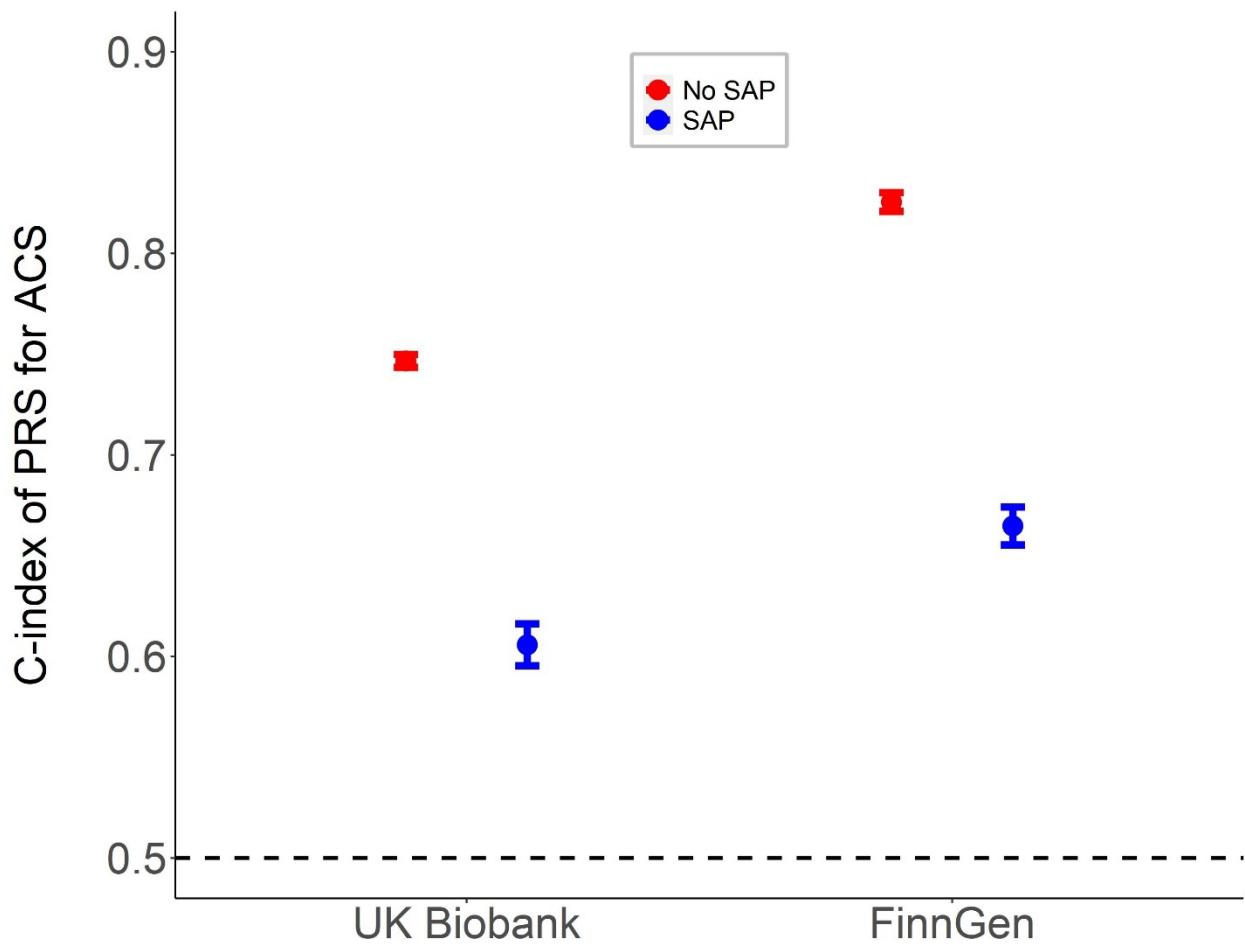


8
9

1 **Figure 3:** Hazard ratios for association between PS and ASC in individuals with and without
2 SAP. In the UK Biobank, individuals with SAP had a hazard ratio of 1.163 [95% CI: 1.082–
3 1.251] and individuals without SAP had a hazard ratio of 1.531 [95% CI: 1.497–1.565]. In
4 FinnGen, individuals with SAP had a hazard ratio of 1.247 [95% CI: 1.173–1.326] and
5 individuals without SAP had a hazard ratio of 1.448 [95% CI: 1.407–1.490].



1 **Figure 4:** Difference in discrimination for ACS, measured by Harrell's C-index. In the UK
2 Biobank, individuals with SAP had a C-index of 0.606 [95% CI: 0.595–0.616] and individuals
3 without SAP had a C-index of 0.747 [95% CI: 0.743–0.750]. In FinnGen, individuals with SAP
4 had a C-index of 0.665 [95% CI: 0.655–0.674] and individuals without SAP had a C-index of
5 0.825 [95% CI: 0.821–0.830]. The lower C-index in individuals with SAP is explained by the
6 lower association between established cardiovascular risk factors, including age, and ACS in
7 these individuals.



1 Acknowledgements

2 UK Biobank analyses were conducted under application 31063. We acknowledge all study
3 participants for their generous participation in the UK Biobank and FinnGen. This work was
4 supported by the Stanford University Undergraduate Research Student Grant.

5

1 References

- 2 1. Townsend N, Wilson L, Bhatnagar P, Wickramasinghe K, Rayner M, Nichols M.
- 3 Cardiovascular disease in Europe: epidemiological update 2016. Eur Heart J. 2016;37(42):3232-
- 4 45.
- 5 2. World Health O. The atlas of heart disease and stroke / Judith Mackay and George
- 6 Mensah ; with Shanthi Mendis and Kurt Greenland. Geneva: World Health Organization; 2004.
- 7 3. BHF Statistics Factsheet. British Heart Foundation: British Heart Foundation; 2020.
- 8 4. Mozaffarian D, Benjamin EJ, Go AS, Arnett DK, Blaha MJ, Cushman M, et al.
- 9 Executive Summary: Heart Disease and Stroke Statistics--2016 Update: A Report From the
- 10 American Heart Association. Circulation. 2016;133(4):447-54.
- 11 5. Mach F, Baigent C, Catapano AL, Koskinas KC, Casula M, Badimon L, et al. 2019
- 12 ESC/EAS Guidelines for the management of dyslipidaemias: lipid modification to reduce
- 13 cardiovascular risk: The Task Force for the management of dyslipidaemias of the European
- 14 Society of Cardiology (ESC) and European Atherosclerosis Society (EAS). European Heart
- 15 Journal. 2019;41(1):111-88.
- 16 6. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, et al.
- 17 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease. A Report of
- 18 the American College of Cardiology/American Heart Association Task Force on Clinical
- 19 Practice Guidelines. 2019;74(10):e177-e232.
- 20 7. Aronson D, Edelman ER. Coronary artery disease and diabetes mellitus. Cardiol Clin.
- 21 2014;32(3):439-55.
- 22 8. Grundy SM, Stone NJ, Bailey AL, Beam C, Birtcher KK, Blumenthal RS, et al. 2018
- 23 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APhA/ASPC/NLA/PCNA Guideline on

- 1 the Management of Blood Cholesterol: A Report of the American College of
- 2 Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. Circulation.
- 3 2019;139(25):e1082-e143.
- 4 9. Tušek-Bunc K, Petek D. Comorbidities and characteristics of coronary heart disease
- 5 patients: their impact on health-related quality of life. Health Qual Life Outcomes.
- 6 2016;14(1):159.
- 7 10. Liu M, Li XC, Lu L, Cao Y, Sun RR, Chen S, et al. Cardiovascular disease and its
- 8 relationship with chronic kidney disease. Eur Rev Med Pharmacol Sci. 2014;18(19):2918-26.
- 9 11. Rahman MM, Kopec JA, Cibere J, Goldsmith CH, Anis AH. The relationship between
- 10 osteoarthritis and cardiovascular disease in a population health survey: a cross-sectional study.
- 11 BMJ Open. 2013;3(5).
- 12 12. Abraham G, Havulinna AS, Bhalala OG, Byars SG, De Livera AM, Yetukuri L, et al.
- 13 Genomic prediction of coronary heart disease. Eur Heart J. 2016;37(43):3267-78.
- 14 13. Knowles JW, Ashley EA. Cardiovascular disease: The rise of the genetic risk score.
- 15 PLoS Med. 2018;15(3):e1002546.
- 16 14. Thanassoulis G, Vasan RS. Genetic cardiovascular risk prediction: will we get there?
- 17 Circulation. 2010;122(22):2323-34.
- 18 15. Inouye M, Abraham G, Nelson CP, Wood AM, Sweeting MJ, Dudbridge F, et al.
- 19 Genomic Risk Prediction of Coronary Artery Disease in 480,000 Adults: Implications for
- 20 Primary Prevention. J Am Coll Cardiol. 2018;72(16):1883-93.
- 21 16. Ganna A, Magnusson PK, Pedersen NL, de Faire U, Reilly M, Arnlöv J, et al. Multilocus
- 22 genetic risk scores for coronary heart disease prediction. Arterioscler Thromb Vasc Biol.
- 23 2013;33(9):2267-72.

- 1 17. Lambert SA, Abraham G, Inouye M. Towards clinical utility of polygenic risk scores.
- 2 Hum Mol Genet. 2019;28(R2):R133-R42.
- 3 18. Natarajan P. Polygenic Risk Scoring for Coronary Heart Disease: The First Risk Factor. J
- 4 Am Coll Cardiol. 2018;72(16):1894-7.
- 5 19. Aragam KG, Dobbyn A, Judy R, Chaffin M, Chaudhary K, Hindy G, et al. Limitations of
- 6 Contemporary Guidelines for Managing Patients at High Genetic Risk of Coronary Artery
- 7 Disease. J Am Coll Cardiol. 2020;75(22):2769-80.
- 8 20. Nikpay M, Goel A, Won HH, Hall LM, Willenborg C, Kanoni S, et al. A comprehensive
- 9 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. Nat
- 10 Genet. 2015;47(10):1121-30.
- 11 21. Mars N, Koskela JT, Ripatti P, Kiiskinen TTJ, Havulinna AS, Lindbohm JV, et al.
- 12 Polygenic and clinical risk scores and their impact on age at onset and prediction of
- 13 cardiometabolic diseases and common cancers. Nat Med. 2020;26(4):549-57.
- 14 22. Rao AS, Knowles JW. Polygenic risk scores in coronary artery disease. Curr Opin
- 15 Cardiol. 2019;34(4):435-40.
- 16 23. Khera AV, Chaffin M, Aragam KG, Haas ME, Roselli C, Choi SH, et al. Genome-wide
- 17 polygenic scores for common diseases identify individuals with risk equivalent to monogenic
- 18 mutations. Nat Genet. 2018;50(9):1219-24.
- 19 24. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank
- 20 resource with deep phenotyping and genomic data. Nature. 2018;562(7726):203-9.
- 21 25. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J, et al. UK biobank: an open
- 22 access resource for identifying the causes of a wide range of complex diseases of middle and old
- 23 age. PLoS Med. 2015;12(3):e1001779.

- 1 26. Ge T, Chen CY, Ni Y, Feng YA, Smoller JW. Polygenic prediction via Bayesian
- 2 regression and continuous shrinkage priors. *Nat Commun.* 2019;10(1):1776.
- 3 27. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation
- 4 PLINK: rising to the challenge of larger and richer datasets. *Gigascience.* 2015;4:7.
- 5 28. Kerr KF, Wang Z, Janes H, McClelland RL, Psaty BM, Pepe MS. Net reclassification
- 6 indices for evaluating risk prediction instruments: a critical review. *Epidemiology.*
- 7 2014;25(1):114-21.
- 8 29. Kundu S, Aulchenko YS, van Duijn CM, Janssens AC. PredictABEL: an R package for
- 9 the assessment of risk prediction models. *Eur J Epidemiol.* 2011;26(4):261-4.
- 10 30. Hoffmann U, Massaro JM, D'Agostino RB, Kathiresan S, Fox CS, O'Donnell CJ.
- 11 Cardiovascular Event Prediction and Risk Reclassification by Coronary, Aortic, and Valvular
- 12 Calcification in the Framingham Heart Study. *J Am Heart Assoc.* 2016;5(2).
- 13 31. Ripatti S, Tikkanen E, Orho-Melander M, Havulinna AS, Silander K, Sharma A, et al. A
- 14 multilocus genetic risk score for coronary heart disease: case-control and prospective cohort
- 15 analyses. *Lancet.* 2010;376(9750):1393-400.
- 16 32. Austin PC, Pencina MJ, Steyerberg EW. Predictive accuracy of novel risk factors and
- 17 markers: A simulation study of the sensitivity of different performance measures for the Cox
- 18 proportional hazards regression model. *Stat Methods Med Res.* 2017;26(3):1053-77.
- 19 33. Tzoulaki I, Liberopoulos G, Ioannidis JP. Assessment of claims of improved prediction
- 20 beyond the Framingham risk score. *JAMA.* 2009;302(21):2345-52.
- 21 34. Hindy G, Wiberg F, Almgren P, Melander O, Orho-Melander M. Polygenic Risk Score
- 22 for Coronary Heart Disease Modifies the Elevated Risk by Cigarette Smoking for Disease
- 23 Incidence. *Circ Genom Precis Med.* 2018;11(1):e001856.

- 1 35. Dunder K, Lind L, Lagerqvist B, Zethelius B, Vessby B, Lithell H. Cardiovascular risk
- 2 factors for stable angina pectoris versus unheralded myocardial infarction. Am Heart J.
- 3 2004;147(3):502-8.
- 4 36. Cassar A, Holmes DR, Rihal CS, Gersh BJ. Chronic coronary artery disease: diagnosis
- 5 and management. Mayo Clin Proc. 2009;84(12):1130-46.
- 6 37. Eisen A, Bhatt DL, Steg PG, Eagle KA, Goto S, Guo J, et al. Angina and Future
- 7 Cardiovascular Events in Stable Patients With Coronary Artery Disease: Insights From the
- 8 Reduction of Atherothrombosis for Continued Health (REACH) Registry. J Am Heart Assoc.
- 9 2016;5(10).
- 10 38. Doll JA, Tang F, Cresci S, Ho PM, Maddox TM, Spertus JA, et al. Change in Angina
- 11 Symptom Status After Acute Myocardial Infarction and Its Association With Readmission Risk:
- 12 An Analysis of the Translational Research Investigating Underlying Disparities in Acute
- 13 Myocardial Infarction Patients' Health Status (TRIUMPH) Registry. J Am Heart Assoc.
- 14 2016;5(6).
- 15 39. Carpeggiani C, Michelassi C, Landi P, L'Abbate A. Long-term prognosis of unheralded
- 16 myocardial infarction vs chronic angina; role of sex and coronary atherosclerosis burden. BMC
- 17 Cardiovasc Disord. 2018;18(1):156.
- 18 40. Crea F, Gasparedone A, Tomai F, Shoulders C, De Fazio A, Versaci F, et al. Risk factors
- 19 in schoolchildren associated with a family history of unheralded myocardial infarction or
- 20 uncomplicated stable angina in male relatives. J Am Coll Cardiol. 1994;23(6):1472-8.
- 21 41. Hjemdal P, Eriksson SV, Held C, Forslund L, Näsman P, Rehnqvist N. Favourable long
- 22 term prognosis in stable angina pectoris: an extended follow up of the angina prognosis study in
- 23 Stockholm (APSIS). Heart. 2006;92(2):177-82.

- 1 42. Levin MG, Rader DJ. Polygenic Risk Scores and Coronary Artery Disease: Ready for
 - 2 Prime Time? *Circulation*. 2020;141(8):637-40.
 - 3 43. Hu YJ, Schmidt AF, Dudbridge F, Holmes MV, Brophy JM, Tragante V, et al. Impact of
 - 4 Selection Bias on Estimation of Subsequent Event Risk. *Circ Cardiovasc Genet*. 2017;10(5).
 - 5 44. Dahabreh IJ, Kent DM. Index event bias as an explanation for the paradoxes of
 - 6 recurrence risk research. *JAMA*. 2011;305(8):822-3.
 - 7 45. Smits LJ, van Kuijk SM, Leffers P, Peeters LL, Prins MH, Sep SJ. Index event bias-a
 - 8 numerical example. *J Clin Epidemiol*. 2013;66(2):192-6.
 - 9 46. Conroy M, Sellors J, Effingham M, Littlejohns TJ, Boultwood C, Gillions L, et al. The
 - 10 advantages of UK Biobank's open-access strategy for health research. *J Intern Med*.
 - 11 2019;286(4):389-97.
 - 12 47. Sund R. Quality of the Finnish Hospital Discharge Register: a systematic review. *Scand J*
 - 13 *Public Health*. 2012;40(6):505-15.
 - 14 48. Tolonen H, Salomaa V, Torppa J, Sivenius J, Immonen-Räihä P, Lehtonen A, et al. The
 - 15 validation of the Finnish Hospital Discharge Register and Causes of Death Register data on
 - 16 stroke diagnoses. *Eur J Cardiovasc Prev Rehabil*. 2007;14(3):380-5.
- 17