

Ancestry-specific polygenic risk scores are risk enhancers for clinical cardiovascular disease assessments

Received: 26 April 2023

Accepted: 25 October 2023

Published online: 04 November 2023

 Check for updates

George B. Busby¹✉, Scott Kulm¹, Alessandro Bolli¹, Jen Kintzle¹, Paolo Di Domenico¹ & Giordano Bottà¹✉

Clinical implementation of new prediction models requires evaluation of their utility in a broad range of intended use populations. Here we develop and validate ancestry-specific Polygenic Risk Scores (PRSs) for Coronary Artery Disease (CAD) using 29,389 individuals from diverse cohorts and genetic ancestry groups. The CAD PRSs outperform published scores with an average Odds Ratio per Standard Deviation of 1.57 (SD = 0.14) and identify between 12% and 24% of individuals with high genetic risk. Using this risk factor to reclassify borderline or intermediate 10 year Atherosclerotic Cardiovascular Disease (ASCVD) risk improves assessments for both CAD (Net Reclassification Improvement (NRI) = 13.14% (95% CI 9.23–17.06%)) and ASCVD (NRI = 10.70 (95% CI 7.35–14.05)) in an independent cohort of 9,691 individuals. Our analyses demonstrate that using PRSs as Risk Enhancers improves ASCVD risk assessments outlining an approach for guiding ASCVD prevention with genetic information.

Atherosclerotic Cardiovascular Diseases (ASCVD) contribute an increasing burden on healthcare systems and are responsible for 1 in 3 deaths worldwide¹. Whether an individual will suffer from ASCVD depends on the complex interplay between environmental and lifestyle exposures such as poor diet and lack of physical activity, physiological components such as high LDL cholesterol and high systolic blood pressure² and their genes. Given the high heritability of ASCVD (40–60%)³ it is unsurprising that many studies have successfully uncovered genetic contributions from both rare pathogenic mutations in genes such as *APOB*, *LDLR* and *PCSK9*, that can cause hypercholesterolemia, and more common genetic variants spread throughout the genome^{4–8}. Utilizing common genetic variation, Polygenic Risk Scores (PRSs), have shown promise as tools to quantify genetic risk and identify significant proportions of the population at high genetic risk of cardiovascular disease^{9–12}.

However, questions remain about the utility of PRSs in clinical care. These include doubts and misconceptions about what PRSs are and how they should be used¹³, as well as important considerations about the lack of transferability across populations with divergent ethnicities and genetic ancestries from those used in their

development¹⁴. Historical biases in the collection and analyses of genomic data have led to an over-representation of individuals of European genetic ancestry in clinical datasets¹⁵, which in turn has led to a lack of sufficiently diverse data with which to build PRSs. While the attenuation of PRSs in different populations is expected based on known differences in allele frequencies and linkage disequilibrium patterns across global populations¹⁶, for PRSs to be utilized clinically they need to deliver meaningful risk stratification regardless of the range of genetic ancestries present in the populations in which they are intended to be used.

Here we describe the development and validation of multi-ancestry PRSs for CAD using an approach that utilizes multiple ancestry-specific GWASs¹⁷ and diverse PRS discovery, validation, and testing datasets. The aim of this work is to provide evidence of the performance of CAD PRSs across diverse ancestries and to demonstrate how they can be used to enhance clinical ASCVD prevention tools. Throughout, we use the term *genetic ancestry* and *ancestry* interchangeably to refer to groupings of individuals who are similar to each other genetically and label these groups using current-day geographical names that relate to the continents on which most

¹Allelica Inc, 447 Broadway, New York, NY 10013, USA. ✉e-mail: george@allelica.com; giordano@allelica.com

individuals belonging to that group currently reside. These labels are intended to assign individuals to groups based on their genetics alone and are not social or ethnicity identifiers. We acknowledge that these labels are imperfect, both because modern human populations rarely contain ancestry from a single region¹⁸, but also because they enforce the artificial discretisation of continuous human genetic diversity¹⁹. Moreover, when grouping individuals into continental-level ancestries, as we do here, a large amount of within-group diversity is concealed. Nevertheless, because genetic variation is shared amongst individuals with similar genetic ancestry, these approximations help to assess genetic effects in groups of similar individuals, and are a necessary tradeoff between optimizing PRSs for a diverse range of populations and having sufficient data from such groups to validate and test resulting scores.

Results

We developed 152 PRSs for CAD by applying 38 different combinations of seven separate ancestry-specific GWASs, including a range of allele frequency filters and finemapped variations (Supplementary Table 1, Supplementary Figs. 1 and 2 and “Methods”), to the PRS-CSx tool¹⁷ implemented within Allelica’s *DISCOVER* software. Finemapping is particularly appropriate for this study because it increases the portability of GWAS signals across populations by identifying putatively causal as opposed to linked variants, and has been shown to increase the transferability of PRSs across ancestries²⁰. We used three publicly available prospective cohorts comprising clinical ASCVD related covariates and matched genetic data to validate and test these PRSs: the MultiEthnic Study of Atherosclerosis (MESA),²¹ Atherosclerosis Risk in the Community (ARIC),²² and UK Biobank (UKB)²³. These data were harmonized into a combined dataset of 30,809 cases and 466,860 controls of which 27,158 cases and 440,404 controls were employed in genome wide association studies and 2931 cases and 26,456 controls were employed in PRS development (Supplementary Table 2). We inferred continent-level genetic ancestry proportions for each individual and assigned them to genetic ancestry groups on the basis of this inference (“Methods”). We applied the 152 PRSs to five genetic ancestry-specific PRS validation datasets, both individually and in combination as metaPRSs, so that the best performing score in each genetic ancestry group could be identified (Supplementary Table 3). To the best of our knowledge, the output of PRS-CSx has never been previously used to create a metaPRS in this way. These best scores were then evaluated in independent ancestry-specific testing datasets to define a final assessment of their predictive performance (Fig. 1). When possible, we focused on fixed effects meta-analyses of the results of two independent cohorts of a single ancestry group. The predictive performance measured by the score’s Odds Ratio per Standard Deviation (ORxSD) ranged from 1.47 (95% CI 1.08–2.01) in the American genetic ancestry group to 1.81 (95% CI 1.31–2.50) in the South Asian genetic ancestry group (Fig. 1a).

To contextualize the performance of the novel multi-ancestry scores we benchmarked them against three previously published scores^{9,10,12}. Our multi-ancestry PRSs had a greater ORxSD than previously published panels across all ancestry groups (Fig. 2 and Supplementary Tables 4 and 5). As an illustrative example, models fit to predict incident and prevalent cases of CAD in individuals of African ancestry in the UK Biobank resulted in an ORxSD of 1.37 (95% CI 0.935–2.01) and for individuals in MESA resulted in an ORxSD of 1.79 (95% CI 1.14–2.81). The ORxSD from a meta-analysis of these two results was 1.53 (95% CI 1.15–2.05), greater than ORxSDs of 1.04 (95% CI 0.79–1.39) from the GPS_CAD PRS, 1.24 (95% CI 0.92–1.67) from the Allelica_CAD_EUR_2020 PRS and 1.30 (95% CI 0.96–1.76) from the multiGRS_CAD PRS (Supplementary Table 6). These differences suggest that not utilizing the novel PRSs for future use would likely lead to relatively inferior predictions.

As a major aim of this study is to build accurate models across ancestries it is necessary to test calibration as well as model performance. Our multi-ancestry PRSs were similarly better calibrated (lower Brier score) than competing scores for all single ancestry groups (Supplementary Table 6) as well as in a group of individuals who were of admixed ancestry and could not be assigned a single ancestry label. In this group, we used the ancestry weighted sum of the single ancestry model predictions to compute a Brier score of 0.06085 for all of the Allelica_CAD_vj PRSs (“Methods”). The Brier scores computed in an equivalent process for the GPS_CAD score was 0.06089, Allelica_CAD_EUR_2020 PRS was 0.06095 and multiGRS_CAD PRS was 0.06107 (Supplementary Table 6). We hypothesize that this increase in calibration performance compared to all alternative PRSs was due to the additional information gained from the ancestry-specific effect sizes present in the multiple GWASs and the use of finemapping to focus effects towards causal variants. Additional statistics, such as Nagelkerke R^2 and the Area Under the receiver operator Curve, and models containing additional sets of covariates, such as age and/or sex alone, further confirmed that our novel multi-ancestry score performed better than the alternative scores (Fig. 2, Supplementary Figs. 3 and 4, and Supplementary Tables 4 and 8).

The Pooled Cohort Equations (PCE)²⁴ are a validated clinical tool that take multiple risk factors associated with ASCVD to calculate an individual’s 10-year risk of disease²⁵. The PCE and other related ASCVD risk prediction algorithms do not currently take genetics into account despite growing evidence demonstrating both that individuals at heightened risk of disease can be identified by PRSs, and that these individuals are largely invisible to clinical risk assessments^{9–11,26,27}. Current American Heart Association and American Academy of Cardiology blood lipid management guidelines advocate that for individuals with borderline (5–7.5%) or intermediate (7.5–20%) 10-year risk (BIR), additional Risk Enhancers can be used to further refine risk mitigation discussions between physician and patient, and PRSs have recently been proposed to aid clinical management of ASCVD in such discussions^{25,28} while reducing healthcare costs from a payer’s perspective²⁹. We therefore sought to assess the potential for the multi-ancestry PRSs to be used as a discriminatory risk enhancing factor for CAD by using a twofold increased risk threshold to identify individuals at high risk of CAD. Similar to the ORxSD, this threshold carries a range of error that is exacerbated by the size of the sample analyzed. Nevertheless its point estimate is equivalent to that of well established risk factors such as family history and some Mendelian-inherited genetic variants³⁰ as well other risk enhancing factors such as diabetes and ethnicity which are currently used to upclassify BIR PCE individuals^{11,24,31}.

In a PCE testing dataset comprising 9691 individuals of diverse ancestry, we identified ancestry-specific twofold risk thresholds as the percentile of the PRS distribution at which the risk in the upper tail is twice the risk of the lower tail (“Methods”). With meta-analyses where possible, we identified the top 12–24% of individuals of each ancestry group with twofold increased risk compared to the remainder (Fig. 1c, d and Fig. 2c). We tested the accuracy of these thresholds by applying them to a unused subset of 6480 admixed individuals from the testing dataset by constructing unique individual-level thresholds as an average of the ancestry-specific percentiles weighted by each admixed individual’s ancestry proportions (“Methods”). This approach allowed us to classify admixed individuals whose PRS scores placed them at twofold increased genetic risk. We then tested whether this binary risk factor recovered twofold increased risk for CAD for those carrying it with logistic regression, and confirmed that the thresholds were able to accurately assign individuals at twofold risk in this group (ORxSD of 1.96; 95% CI 1.42–2.70). In addition to accurately assigning risk in the independent admixed group, the percentile thresholds also identified a greater share of individuals than any of the competing PRSs. For example, the average difference between the twofold threshold found

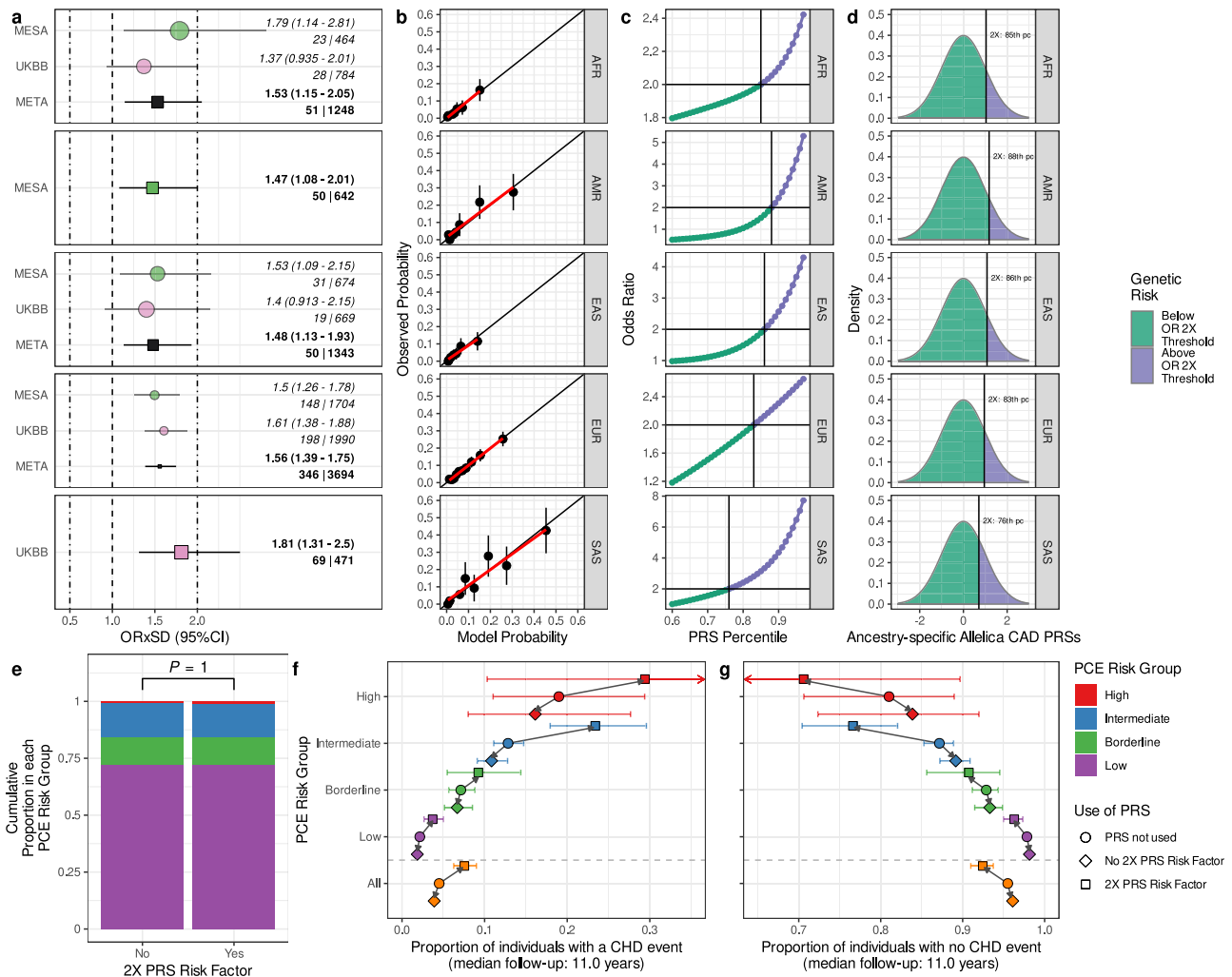


Fig. 1 | Multi-ancestry PRSs for CAD predict risk in diverse ancestries and add to clinical risk assessments. For five different genetic ancestry groups we show **a** the point estimate and 95% CI of the Odds Ratio per Standard Deviation (ORxSD) of the ancestry-specific polygenic risk scores in the UK Biobank and MESA Testing populations. META describes the results of a meta analysis of the cohort-specific ORxSDs. Squares denote the reported ORxSD for each ancestry and point size is proportional to the standard error of the ORxSD estimate. Values on the right of the panel are the reported ORxSD with 95% CI together with the numbers of cases and controls in the Testing populations. **b** The calibration curves of the PRS models that compare each model's response predictions to the actual prevalence of disease. The points show the mean value of each quantity across deciles of the predictions, vertical error bars show the 95% CI around the mean probability in each decile, the red line follows the trend of these decile values and the black line shows perfect calibration; **c** the Odds Ratio based on the upper tail compared to the remainder of the distribution and the percentile threshold at which 2x increased risk is met; **d** the proportion of the PRS distribution at at least twofold increased risk

compared with the remainder; **e** comparison showing no significant difference (calculated via one-sided Fisher's Exact Test) between the proportions of individuals in each of the PCE risk groups in those with and without PRS as a risk enhancing factor; **f** the proportion of individuals in each of the PCE strata who had a CHD event during the follow-up time of the prospective cohort that they were sampled from. We show the proportions for individuals classified into each PCE group and the total dataset (All) split into individuals with and without PRS as a Risk Enhancing Factor. Horizontal error bars denote the 95% CI around the proportion of individuals in each risk group; **f** is the same as **g** except showing the proportion of individuals in each of the PCE strata who did not have a CHD event during the follow-up time of the prospective cohort that they were sampled from. Horizontal error bars denote the 95% CI around the proportion of individuals in each risk group. Total sample sizes for each group are (cases/controls): AFR / African (51/1248); AMR / American (50/642); EAS / East Asian (50/1343); EUR / European: (346/3694); SAS / South Asian (69/471).

with all of the Allelic_{CAD_v} scores and the lowest percentile threshold of each of the three competing scores (Supplementary Tables 3 and 8) averaged 7.0% (SD 3.74%). This difference would imply that for every 1000 individuals screened the Allelic scores would identify 70 more individuals at high risk who would have otherwise been missed.

We used the PCE to classify individuals with BIR of ASCVD and performed two analyses using the CAD PRS as a risk factor to reclassify individuals using separate disease endpoints of Coronary Heart Disease (CHD; $N = 9691$) and ASCVD ($N = 9569$). There was no significant difference in the proportion of individuals in each PCE risk strata with and without the twofold risk factor (Fig. 1e) indicating that the PCE alone does not capture the increased risk

attributable to the PRS risk factor. Across BIR individuals, 16% had the twofold PRS risk factor. Among BIR individuals with the twofold risk factor, there were 64% more CHD events and 47% more ASCVD events than in those BIR individuals without the PRS risk factor (Supplementary Tables 10 and 11). Upclassifying BIR individuals on the basis of the PRS risk factor resulted in a Net Reclassification Improvement (NRI) of 13.14% (95% CI 9.23–17.06%) when applying the CAD PRSs as a risk enhancing factor with CHD as an endpoint (Supplementary Table 9) and 10.70% (7.35–14.05%) when using ASCVD as the primary endpoint of the analysis (Supplementary Table 10). We confirmed that the twofold increased risk effect of the CAD PRSs was present in the PCE testing dataset (OR = 2.03 95% CI

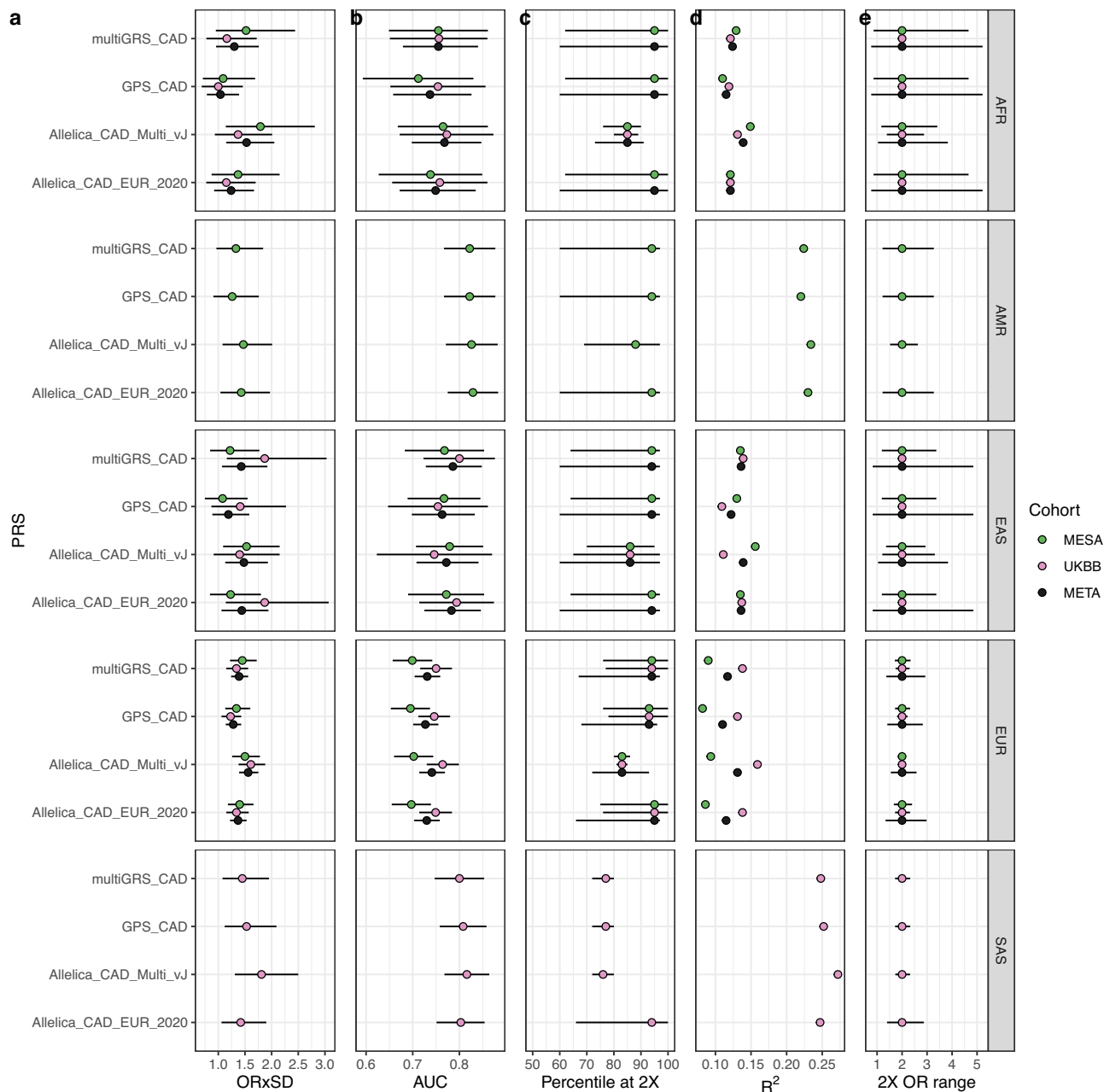


Fig. 2 | Evaluation of the performance of the Allelica CAD multi-ancestry PRS against three published PRSs. For each PRS/genetic ancestry combination we show benchmarking results for the Allelica_CAD_vj PRSs and three published PRSs applied to exactly the same testing populations. Total sample sizes of the testing populations are (cases/controls): AFR / African (51/1248); AMR / American (50/642); EAS / East Asian (50/1343); EUR / European (346/3694); SAS / South Asian (69/471) and a breakdown per sub-cohort shown in Supplementary Table 15. We assessed

the statistics on the sub-cohorts and, where possible, a meta analysis of the sub-cohort values. The statistics shown, with horizontal error bars detailing the 95% CI of these values, are **a** the point estimate of the Odds Ratio per Standard Deviation of the PRSs; **b** the Area Under the Receiver Operator Curve with 95% CI estimated; **c** the threshold percentile of the PRS distribution where 2× risk is achieved; **d** Nagelkerke R^2 ; and **e** the range of the odds at the 2× threshold.

1.49–2.76) even after controlling for other PCE covariates with logistic regression (Supplementary Tables 11 and 12). Stratifying CHD cases within each PCE risk strata by the PRS risk factor confirmed that individuals with the PRS risk factor had around twice as many events than those without (Fig. 1f, g). Finally, we assessed the improvement in classification of integrating the CAD PRSs directly into the PCE (Supplementary Table 13) and observed a lower but significant categorical NRI for both CAD (NRI = 4.31%; 95% CI 1.85–6.78) and ASCVD (NRI = 2.90%; 95% CI 0.61–5.18).

Discussion

Here we have demonstrated that validated and calibrated ancestry-specific PRSs for CAD can be used as a risk factor to better classify an individual’s overall risk of ASCVD. Nevertheless, there are several limitations to this study. First, following best practice³² we developed and validated PRSs using independent PRS training, validation and testing cohorts which led to relatively small sample sizes, particularly for the East Asian and American testing groups. Second, these sample sizes necessitated the approach of grouping individuals into continental-level genetic ancestry groups. This is sub-optimal because

it masks genetic diversity and structure within continents which may contribute to differential disease susceptibility within such groups^{33,34}. We also acknowledge that the individuals present in the non-European genetic ancestry groups were sampled from individuals currently residing in Europe and the USA and so the applicability of these multi-ancestry PRSs to individuals living in Asia and Africa remains to be tested³⁵. Third, while differences in the predictive performance across genetic ancestries may be the result of linkage disequilibrium patterns or may reflect true differences in genetic architecture they may also be affected by multiple other factors, such as age and sex distribution, disease definitions, sample ascertainment, as well as variation in environmental risk factors³⁶. Future work that jointly considers the environmental and genetic effects of ancestry differences on risk predictions could also provide a potential avenue for further developing the approaches outlined here. Fourth, we have concentrated on an analysis of using PRS as a Risk Enhancing Factor for the PCE. While the applicability of this approach to other risk algorithms requires further investigation, there is no reason to believe that these results are not generalizable to any ASCVD risk algorithm that uses well established clinical risk factors. Alternative approaches, such as integrating CAD and other PRSs into these risk models are also possible³⁷, but their development is contingent on suitably powered testing datasets. Moreover, if we view PRS as neither a deterministic nor diagnostic test for ASCVD but as a biomarker equivalent to others currently in use, then using such information to guide risk discussions has a precedent in national guidelines with Coronary Artery Calcium scoring, which is currently available to refine risk estimates for individuals at borderline/intermediate 10-year ASCVD risk²⁵. It also captures the essence of how PRS should be used: as one tool among several that can be combined to understand an individual's overall risk of ASCVD.

Recent publications have highlighted the lack of diversity in genetic studies and the poor transferability of polygenic scores across populations^{14,35,38}. These studies rightly advocate for the increased collection and analysis of diverse genomes to aid the clinical application of genomics in healthcare, a mission that we wholeheartedly support. However, unlike the PRSs described here, not all clinical tools currently utilized have been validated across all populations, and this has not precluded their broad use. As an example, most 10-year ASCVD risk models, including the PCE, have not yet been clinically validated in individuals of South Asian descent (and have been shown to underpredict risk in such groups³⁹). While the ultimate aim should always be to develop clinical tools that work equally well across groups, pragmatism should lead us to use those tools that can be considered good enough to significantly enhance preventative efforts to ensure that their utility can be realized. In the context of this study, this means utilizing independently validated ancestry-specific PRSs that are able to identify individuals with higher genetic risk. These PRSs provide actionable information that can not be gleaned from other tests available in the current standard of care, even if there may be marginal differences in model performance across groups. We show here that multi-ancestry CAD PRSs can play a role in risk assessments by reclassifying significant proportions of individuals at increased risk of disease because of their genetics, but who are invisible and therefore missed by current ASCVD risk assessments and may nevertheless go on to have disease. These PRSs are validated and calibrated on US populations and given their demonstrated cost-effectiveness^{29,40} it is increasingly clear that they satisfy the criteria to be considered a Risk Enhancing Factor in the management of ASCVD risk.

Methods

Association data

GWAS. We identified three large GWAS from the literature that did not include any individuals in the downstream Validation and Testing datasets that we used to develop multi-ancestry PRS. These were the

CARDioGRAMplusC4D GWAS⁴¹, the Japanese 161k and the Japanese 52k studies^{42,43}. The summary statistics for all of these studies were downloaded from the GWAS Catalog (Supplementary Table 2).

In the absence of publicly available South Asian and African specific GWAS summary statistics, we performed GWAS on separate subsets of individuals from both ancestries in the UK Biobank using *fastGWA*⁴⁴. The South Asian GWAS comprised 932 cases and 4043 controls and the African GWAS comprised 91 cases and 3230 controls. *fastGWA* was used due to the computational efficiency of the generalized linear mixed model (GLMM)-based method that it implements, which has improved statistical properties when applied to binary traits. In addition we also performed a secondary GWAS on a subset of individuals from the UK Biobank with European ancestry (26,135 cases and 433,131 controls) to generate an additional set of European-based variant associations.

Finemapping. GWAS provides an estimate of the size and significance of the associations of alleles and disease at millions of sites across the genome. Even so, a variant that surpasses a significance threshold may not directly cause the disease of interest but rather only tag the causal variant by being in linkage disequilibrium (LD) with it. To identify putatively causal variants, we applied *POLYFUN*^{20,45}, a method that uses ancestry-specific LD scores and additional data on the putative function of variants across the genome to generate a set of finemapped summary statistics. We applied *POLYFUN* to the CARDioGRAMplusC4D summary statistics together with functional information from Gazal et al.⁴⁶ and LD maps provided with the software, with a causal parameter of 1 and maintaining all other parameters at their default value. We only applied finemapping to the CARDioGRAMplusC4D summary statistics because it allowed us to introduce functional information into our investigation with the best chance of utility, owing to the CARDioGRAMplusC4D's great statistical power, without creating too many, unwieldy modified summary statistics.

The combination of three publicly available GWAS (CARDioGRAMplusC4D, Japanese 161k, Japanese 52k), three self-computed GWAS (UK Biobank African, UK Biobank South Asian, UK Biobank European) and one finemapped GWAS (CARDioGRAMplusC4D, finemapped) resulted in a final list of 7 different sets of summary statistics for downstream polygenic risk score panel creation (Supplementary Table 1).

Variant filtering. A series of quality control steps performed by Allelica's *DISCOVER* software were applied to each of the GWAS summary statistics prior to their use in generating PRS. Specifically, we removed all non-biallelic variants, removed variants with a minor allele frequency (MAF) less than one percent (according to the ancestry-specific 1000 Genomes reported allele frequencies⁴⁷), and removed variants with duplicated IDs. For variants that underwent finemapping we also removed all ambiguous variants and those with non-canonical allele identifiers. For the European GWAS, we generated additional filtered datasets that only contained the 2 million variants with the lowest *P*-values, as well as a filtered set that only contained variants with $MAF > 0.00001$ (Supplementary Table 1). We implemented these filters to remove variants whose imbalanced allele counts lead to biased statistics.

Developing PRS with *DISCOVER*

Polygenic risk scores are computed as a sum of single variant risk effects. The variance explained by the resultant score can theoretically reach the SNP-level of heritability. However, in practice factors, such as limited sample size, prevent the score from reaching even close to the heritability limit^{14,48,49}. Determining the variants and the corresponding effects that create the most accurate polygenic risk score is an area of active research. We used *PRS-CSx*¹⁷ a Bayesian approach that shrinks the GWAS computed effects of variants to disentangle the impact of

linkage disequilibrium while considering multiple ancestral groups, to build ancestry-specific PRSs for this study.

We implemented *PRS-CSx* within Allelica's *DISCOVER* PRS development software. The purpose of the *DISCOVER* software is to take in summary statistics and output a panel of variants, alongside effect sizes, that can be used to create a polygenic risk score of superior accuracy. This design matches the design of other well known genetics software, like *LDpred* and *PRS-CS*. In this specific investigation, within *DISCOVER* we combined GWAS summary statistics that were generated from different ancestral groups so that the modified variant effects in the output would benefit from a diversity of data. Specifically, *PRS-CSx* generated a separate PRS panel for each set of GWAS summary statistics. These PRS panels were combined into a single panel for downstream validation and testing, by combining the constituent variant effect sizes across panels using an inverse-variance-weighted meta-analysis of the panel specific posterior effect sizes. For example, we may input the EUR-1 (from CARDIoGRAMplusC4D), EUR-3 (from UK Biobank European) and AFR-1 (from UK Biobank African) GWAS summary statistics into the *DISCOVER* software, which are jointly analyzed and output as three separate PRS panels. These are then ultimately joined into a single panel of modified variants effects. We used 38 different combinations of multiple input GWAS in this way and named the resulting panel by the GWAS datasets used to generate it, so in this example the panel was named EUR-1;EUR-3;AFR-1 (Supplementary Table 1).

In addition to running *PRS-CSx* with multiple combinations of GWAS summary statistics, we also ran *PRS-CSx* with a range of global shrinkage parameter values (ϕ). We applied a small range of ϕ values to each combination of GWAS summary statistics, (e.g., $\phi = 1e-6, 1e-4, 1e-2, 1$) to generate multiple panels of variant effect sizes for a single combination of GWAS summary statistics. This resulted in 152 total panels of variant effect sizes by applying 38 different combinations of GWAS summary statistics. Throughout the application of *PRS-CSx* we utilized LD panels that were made available in the GitHub repository <https://github.com/getian107/PRScsx>.

Applying PRSs to cohort genetic data with *PREDICT*

For each of the 152 panels of variant effect sizes generated by *DISCOVER* we used Allelica's *PREDICT* software to compute individual-level PRS scores. The purpose of the *PREDICT* software is to take in a PRS panel (list of variants, effect alleles and effect sizes) and genotypes then output a PRS value for each individual who input a genotype. In other circumstances the *PREDICT* software can also calculate genetic risk percentiles and longitudinal risk projections. In this investigation we applied *PREDICT* to 12,751 individuals in the UK Biobank (all of which were not utilized in a previous GWAS), 5748 individuals within MESA, and 10,888 individuals within ARIC. After each PRS was computed it underwent ancestry adjustment and normalization. Ancestry adjustment was accomplished following an established approach³¹ which subtracts the effect of the first four principal components from the PRS values (as indicated in equations 1 and 2 below). Ancestry adjustment was performed on the ancestry-specific cohorts separately. Normalization resulted in the PRSs having a mean of zero and standard deviation of one.

$$\text{adjust_model} : PRS_{raw} \sim PC1 + PC2 + PC3 + PC4 \quad (1)$$

$$PRS_{adjusted} = PRS_{raw} - \text{adjust_model}(PC1, PC2, PC3, PC4) \quad (2)$$

PRS validation and testing

Datasets. As described above, we compiled a joint dataset of 29,387 individuals for the purpose of assessing the predictive value of the PRSs (Supplementary Tables 2 and 3). This joint dataset comprised the

UK Biobank²³ (specifically the portion not already utilized for the GWAS), the MultiEthnic Study of Atherosclerosis²¹ (MESA) and the Atherosclerosis Risk in Communities Study²² (ARIC). Details of the collection of these datasets are detailed in the original publications. Briefly, the UK Biobank is a large, general purpose, prospective biobank containing approximately 500,000 individuals who were aged 40–69 when they were enrolled from 2006 and 2010. If a UK Biobank individual met any of the conditions listed in Supplementary Table 14, we defined the individual as being a CAD case. MESA is a diverse, moderate sized, atherosclerosis-focused study that contains approximately 6500 individuals who were aged 45–84 years when they were enrolled around 2000. If a MESA individual had a physiological event that was recorded as myocardial infarction, resuscitated cardiac arrest, definite angina, probable angina or coronary heart disease (CHD) death we denoted them as being a CAD case. ARIC is a prospective study of atherosclerosis that contains approximately 15,000 individuals enrolled from 1987 to 1998 who were aged 35–74 years. If an ARIC individual reported CAD at baseline or during a follow-up interview then we defined them as being a CAD case. Additional covariates such as age, sex, and family history of CAD were directly accessed from each cohort's descriptive data.

We applied quality control to each of these three datasets. First, we removed individuals from the UK Biobank dataset who were outliers in heterozygosity, contained putative sex chromosome aneuploidy, corresponded to a genotype missing rate greater than 10% or had excess relatives. All of these steps follow quality control flags provided by the UK Biobank²³, and have been taken by other investigators^{50–52}. Second, in both ARIC and MESA we removed individuals whose (pre-imputation) genotype missing rate was greater than 10%. Third, the following variant level QC was applied to all three datasets: variants that had an allele frequency less than 0.01 or a Hardy–Weinberg Equilibrium p -value less than $1e-50$ were removed. These quality control metrics fall in line with those others have applied^{53–55}.

Genetic ancestry inference. We inferred the ancestry of every individual in the Validation and Testing datasets with *iAdmix*⁵⁶ following default settings and with individuals from the 1000 Genomes⁴⁷ study as a reference dataset. For each individual we estimated the proportion of their genetic ancestry that originated from each of the 5 established One Thousand Genome Project continental-level superpopulations: African, American, East Asian, European and South Asian (Supplementary Table 15). These individual-level inferred proportions, which we also call ancestry components, are used in both the validation and testing of the polygenic risk scores in this investigation. Visualizations showing the first and second genetic principal component colored by cohort, reported ethnicity and the genetic ancestry defined above are provided in Supplementary Fig. 5.

Data organization. We used separate independent Validation and Testing datasets to assess the performance of the PRSs across different genetic ancestries (Supplementary Table 3). We split the entirety of the available data into these independent validation and testing datasets by following three straightforward guidelines: (1) keep cohorts together and (2) achieve at least 50 cases within the testing dataset and 10 cases within the validation dataset for each ancestry group and (3) maintain ancestry groups that are as homogeneous (non-admixed) as possible. Additionally, because of different goals corresponding to the validation and testing datasets we counted cases and analyzed the data within each differently. For the validation dataset we fit models with all individuals regardless of their ancestry, weighting their contribution to the model by their *iAdmix* ancestry component. For the testing dataset we fit models with only the individuals who belonged to the specific ancestry group. We followed this procedure in order to have the greatest statistical power as possible within the validation dataset and

have accuracy estimates that most faithfully matched the ancestry label in the testing dataset. As the models fit with validation datasets considered far more individuals than just those in a single ancestry group we lowered the number of cases needed for a single ancestry group. However, to maintain clarity the counts within Supplementary Table 3 for the validation dataset are based on single ancestry group thresholds.

We applied data splitting guidelines following the flow chart in Supplementary Fig. 2. For example, when considering the African ancestry group we counted 336 cases when placing the ARIC cohort in the validation dataset and 51 cases when placing the UK Biobank and MESA cohorts in the testing dataset. Another example, when considering the South Asian cohort nearly all of the cases were held in a single cohort, the UK Biobank. However, we could split this single cohort evenly and reach our guideline of having 50 cases in the testing dataset and 10 cases in the validation dataset. A final example, when considering the American ancestry group, no single cohort could be split evenly to reach the testing and validation dataset case quotas. Even when putting all of one cohort's cases in the testing dataset and splitting the other cohort cases across testing and validation datasets we were still unable to reach the case quotas. We therefore lowered the *iAdmix* ancestry component threshold used to count cases to 50%, and were then able to satisfy the case quotas.

The limited number of cases and controls within non-European ancestry groups increased the chance of generating an anomalous result. We attempted to prevent this by carefully constructing entirely independent validation and testing datasets, applying cross-validation repeatedly within the validation dataset (as will be described in the next section) and ultimately reporting the confidence intervals of all statistics generated from the testing dataset after appropriate control for multiple testing. The only way for us, and many others who have presented similar sample sizes in comparable genetic-based investigations^{57–59}, to achieve an exceedingly large sample size would be to limit the investigation to individuals of European ancestry, a move that would only increase the current inequity in healthcare.

PRS validation. Within each ancestry group of the validation dataset we aimed to identify a best performing PRS that could be evaluated in the testing dataset. We began with 152 PRS values for each individual. Each PRS value was generated from a unique mix of component GWASs and computational parameters as outlined above. Following previous work^{60,61} we theorized that rather than selecting a single PRS value for an individual it might be advantageous to form a single ensemble or meta PRS value. Our motivation was that each different combination of component GWASs and computational parameters captures a different aspect of the true underlying genetic architecture. For example, with one parameterization we may highlight rare high effect variants, in the same way that the clumping method might, whereas in another parameterization we may evenly weight thousands of marginally relevant variants, in the same way that the LDpred method might. By allowing the data, in a [meta PRS] cross-validation procedure, to determine which constituent PRSs from those generated assuming a range of genetic architectures, or methodological philosophies, best explain the data we may achieve the best possible risk predictions. This approach has been used by multiple previous publications^{61–63}. However, we can not rule out that more predictive PRS panels may be generated with the use of additional methodologies in the future.

To form this meta PRS we applied elastic net logistic regression within a repeated cross-validation framework. Across 10 repeats and 10 folds we split the validation dataset into training and testing partitions. In both partitions the model included independent variables of age, sex, the top four genetic principal components, family history of CAD, cohort and all 152 polygenic risk scores. Specifically, on the training partition we determined the optimal λ

value of the elastic net logistic regression and on the testing partition we fit an elastic net logistic regression model with the previously determined λ value. From this model fit to the testing partition we measured the coefficients of the polygenic risk score terms. At the end of the repeated, cross-validation process we generated 100 coefficients for each of the 152 PRSs, with each coefficient representing how salient or important the PRS was to the determination of the CAD case status. The average of the 100 coefficients became the final PRS weight used in the calculation of the meta PRS. The elastic net logistic regression was accomplished with *glmnet* and *cv.glmnet* functions from the *elasticnet* package within *R*. The repeated cross-validation process was organized with the *trainControl* function of the *caret* package within *R*.

While several other groups have either created metaPRSs by forming a weighted average of individual, predictive PRSs, or applied PRS-CSx upon a variety of multi-ancestry GWAS summary statistics, to the best of our knowledge no one else has previously combined both methodologies together. By doing so here, we present a novel way to generate a potentially maximally predictive PRS.

As described previously, to utilize the maximum amount of genetic information as possible (and thereby increase statistical power) we weighted the individuals in the *glmnet* models by their ancestry component. For example, when forming an African ancestry-specific meta PRS we would set the weights term in the model function call to the African ancestry component. We were then able to extend the number of individuals included in the model from those in just the ancestry group of the validation dataset to everyone in any ancestry-specific group of validation dataset.

Despite increasing the number of individuals within the meta PRS process, we found that the meta PRS could generate very uncertain risk predictions. To specify the degree of uncertainty, within each of the 100 total folds of the meta PRS generation process we formed a meta PRS from the coefficients estimated for the 152 PRS terms within that fold. We then regressed the scaled meta PRS against the phenotype and adjusted for the non PRS covariates of the *glmnet* model and lastly recorded the standard error of the meta PRS term. At the end of the meta PRS generation process we then had 100 standard error terms for each ancestry group. We decided that if the mean of the standard error terms was greater than 1 the meta PRS could not be relied upon as a precise risk predictor. To find an alternative PRS to apply within the testing dataset we utilized a bootstrapping procedure which subsampled the validation dataset, with replacement, 1000 times and fit a model for each of the 152 PRSs in each subsample. The PRS with the greatest average ORxSD was chosen as the single PRS to use for that ancestry group within the testing dataset.

All models computed upon the validation data included the covariates of the 152 PRSs, age, sex, top four genetic principal components, family history of CAD and cohort. Each of these covariates was normalized to a mean of zero and standard deviation of one. Any missing values were mean imputed. Each model was weighted by the *iAdmix* ancestry component corresponding to the group of individuals currently under analysis. Additionally, each PRS was ancestry adjusted by predictions made from a model fit to the top ten genetic principal components. This ancestry adjustment process is equivalent to that introduced by Hao et al.³¹, and provides an additional means of correcting for fine-scale ancestry confounding of the polygenic risk scores. The average coefficients resulting from this process, which were then used to calculate the meta PRS, are provided within Supplementary Table 6 for all scores.

PRS testing. Within each ancestry group of the testing dataset we aimed to fairly measure the discriminative ability of the polygenic risk score derived within the validation dataset. Within each of the finalized ancestry groups of the testing dataset we created the meta PRS score by combining the raw PRS values with each of the weights derived from

the validation process. However, before combining the weights with the testing dataset PRSs, we first performed ancestry adjustment and normalization, following the same series of steps applied within the validation dataset. Alternatively, for ancestry groups that were not amenable to the metaPRS process we directly identified the PRS which performed best in a validation dataset bootstrapping process to serve as the ancestry-specific PRS in the testing dataset. We repeated this process for each set of ancestry-specific group for all individuals in the testing dataset. Therefore, each individual in the testing dataset had five, ancestry-specific polygenic risk scores.

These testing dataset ancestry-specific polygenic-risk scores were derived in the entirely independent validation datasets, meaning that any score that worked well in the validation dataset just by chance would almost certainly not also work well in the testing dataset. Furthermore, for most of the ancestry groups we utilized not just different individuals, but individuals from different cohorts between the validation and testing datasets, reducing the chance that some cohort-specific confounding may limit the ability of our scores to perform well in other cohorts in the future.

Next, within each cohort of each ancestry-specific group we formed a logistic regression model of the corresponding ancestry-specific PRS against CAD disease status adjusted for age, sex, top four genetic principal components, and family history. All of the adjusting covariates were normalized to a mean of zero and standard deviation of one. Following other authors, we utilized the top four genetic principal components when using these covariates in different analyses^{64–67}. Furthermore, we conducted some sensitivity analyses and found that four genetic principal components were sufficient to predict the genetic ancestry assignments, within ten fold cross validation of the testing dataset, with an accuracy of 99%. The plots of all four genetic principal components graphically how certain combinations of components well split genetic ancestries (Supplementary Fig. 6).

The exponentiated coefficient of the PRS term in the model became the reported Odds Ratio per Standard Deviation (ORxSD). We also calculated the Brier score, Area Under the receiver operator Curve, Nagelkerke R^2 and threshold at which a portion of the sample was at a certain multiple of odds compared to the remainder of the sample. When multiple cohorts were available for a single ancestry group (Africans, East Asians and Europeans) we performed a fixed effects meta-analysis with the meta package within the R computing language. We chiefly analyze the results of this meta-analysis, and not the results of the component cohorts, for these ancestry groups where a meta-analysis was conducted (Supplementary Table 16).

Individuals who we were unable to assign a single ancestry group, because they did not meet the *iAdmix* cut-off value of any group, were placed within a group we considered to be of admixed individuals. For these individuals we calculated an ancestry weighted PRS. This ancestry weighted PRS is a sum of the product between an individual's ancestry-specific PRS and the $\log(\text{ORxSD})$ of that PRS measured within the corresponding single ancestry group and the corresponding ancestry component⁶⁸. So if an individual was determined to be admixed with 20% European and 80% African ancestry their weighted PRS would be 0.2 of their European specific PRS value multiplied by $\log(1.56)$ plus 0.8 of their African specific PRS value multiplied by $\log(1.53)$. We additionally performed ancestry adjustment and normalization upon this weighted PRS within each ancestry group. For each ancestry group we thereby generated an admixture-aware PRS with a mean of zero and standard deviation of one. We then applied each of the models fit upon the single ancestry groups to the group of individuals with admixed ancestry to form a set of ancestry-specific predictions. Although, in this prediction calculation, the standard PRS was replaced with the weighted PRS. We additionally summed the ancestry prediction weighted by each individual's ancestry component to form a final,

single set of predictions. The logistic transformed predictions were compared to the true disease status of the individuals to calculate brier scores.

Estimating the proportion of a population at twofold risk. An additional statistic that measures the discriminative ability of a polygenic risk score is the odds against the remainder. The odds against the remainder is calculated by splitting a group of individuals according to whether their polygenic risk score is above or below a given threshold value, then calculating the odds ratio of disease of the high risk group against the low risk group. This type of odds ratio is often calculated at a specific quantile of the polygenic risk score within the group of individuals. Alternatively, we can work backwards and attempt to measure the quantile that achieves a specific odds against the remainder. And just as the odds ratio has a corresponding error range, the quantile value also has an error range. The single, point estimate of the quantile value is not the absolute two fold threshold but rather the best guess within a range of likely values. This idea builds upon Aragam et al.¹¹, who estimated the odds ratio conferred by being in the top quintile against the remainder of the distribution.

Just as a measured odds ratio is only relevant to the polygenic risk score it was computed from, the threshold is only relevant to a specific polygenic risk score applied to a specific testing population. Therefore, if the metaPRS panel we developed for a specific ancestry group generates a threshold of 80%, then only a PRS value calculated from the same panel upon the same ancestry group can be used with this 80% threshold.

The relatively small samples available to us made it unfeasible for us to calculate the odds threshold directly. Rather, we utilized a bootstrapping process which repeatedly sampled with replacement from an ancestry group within the testing dataset. With this bootstrapped sample we measured the odds against the remainder at a range of quantile values. Where possible (the African, East Asian and European ancestry groups) the final odds against the remainder value was calculated with a fixed effects meta-analysis (specifically with the *metabin* function of the meta package within the R computing language). After the bootstrap process we had a sample of odds against the remainder values at a variety of polygenic risk score quantiles. We were then able to fit a trend to the data and determine the exact polygenic risk score quantile at which the trend surpassed the desired odds against the remainder. In this fashion we computed the value recorded as the 2X threshold percentile.

Estimating 10-year ASCVD risk with the Pooled Cohort Equations. To estimate 10-year clinical risk using the Pooled Cohort Equations (PCE) we extracted and recoded information about clinical covariates from data available in the UK Biobank and MESA cohorts. This included: age, sex, smoking, LDL, HDL and total cholesterol, treated and untreated systolic blood pressure, diabetes status, family history of heart disease, CAD outcome.

Individuals within the Testing cohort were combined and filtered for adherence with 2018 American College of Cardiology (ACC)/American Heart Association (AHA) guidelines using the following Inclusion criteria: LDL comprised between 70 and lower than 190 mg/dL; Age at baseline comprised between 40 and 75; No Diabetes status at Baseline; No LDL-lowering medications at baseline; No missing values at any PCE variables. Re-calibrated 10-year risk was calculated for the PCE testing dataset as per Elliot et al.⁶⁹. We estimated the baseline survival function in the PCE testing dataset through a Cox regression proportional hazard model (with PCE risk factors as control covariates) and combined it with published PCE Hazard Ratio values.

Comparison of risk category frequencies (Low, Borderline, Intermediate, and High 10-year PCE risk) between Not elevated and High genetic risk (genetic risk lower or equal/higher than twofold,

respectively) was assessed and no significant difference between risk frequencies in Low and High PRS groups was found (one-way χ^2 test; $P = 0.983$).

Finally, we performed separate confirmatory analyses using logistic regression of CHD and ASCVD as outcomes in the full PCE dataset recomputing the effect sizes (odds ratios) for all PCE covariates (age, sex, smoking, LDL, HDL and total cholesterol, treated and untreated systolic blood pressure, diabetes status, family history of heart disease) and a binary indicator variable denoting high CAD PRS (>2X threshold).

Integrating CAD PRS into the PCE. We used Allelica's *INTEGRATEpce* tool to assess the effect of integrating our ancestry-specific PRSs into the PCE model. The *INTEGRATEpce* software generally combines genetic risk predictions with non-genetic risk factors (such as cholesterol and medication status) in order to form a single, integrated disease risk prediction. In this investigation, we incorporated PRS-specific weights as the logarithm of the PRS hazard ratio (i.e. Log(HR)) with clinical risk weighted from the PCE model to output a single 10-year risk value. HRs for the PRSs were calculated in a Cox proportional hazards regression, with time of follow-up time variable and CAD as outcome. Cox regression was adjusted for age and first four Principal Components of ancestry, and used CAD as clinical outcome and time of follow-up as time variable. Cox regression was performed in the PCE testing dataset ($N = 9114$) and because separate HRs are available for sex and ethnicity, we stratified by sex and self-reported ethnicity (African Americans and non-African Americans). Of note, the African American group comprised all individuals with African major genetic ancestry fraction, while the non-African Americans comprised individuals with European, South American, East-, and South-Asian major genetic ancestry fractions. As before, we analyzed the Net Reclassification Improvement (NRI) of PRS, considering two NRI measures: categorical NRI (calculated for a risk threshold of 20%) and continuous NRI (Supplementary Table 13).

Computational considerations

We analyzed the PRS with the coding languages R and Python. With each of these languages multiple packages/libraries were utilized. These packages/libraries are listed within Supplementary Table 17. Finally, as we have utilized multiple abbreviations within this text we have compiled a glossary within Supplementary Table 18 for easier reading.

Statistical considerations

When not otherwise noted the ranges around a point statistic are 95% confidence intervals. P -value significance thresholds were obtained after Bonferonni Correction.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The UK Biobank is available to qualified researchers through <https://www.ukbiobank.ac.uk/enable-your-research/apply-for-access>. Both the MESA and ARIC datasets are freely available through dbGaP with accession codes phs000209.v2.p1 and phs000280.v3.p1, respectively. The 1000 Genomes dataset is freely available from multiple sources, including <https://www.internationalgenome.org/>. The polygenic risk scores described in this study are available under restricted access for non-commercial research use and can be obtained upon written request detailing their proposed use to the corresponding authors (giordano@allelica.com/george@allelica.com).

References

- Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *Lancet* **396**. [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(20\)30925-9/abstract](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)30925-9/abstract).
- Getz, G. S. & Reardon, C. A. Nutrition and cardiovascular disease. *Arterioscler. Thromb. Vasc. Biol.* **27**, 2499–2506 (2007).
- Zdravkovic, S. et al. Heritability of death from coronary heart disease: a 36-year follow-up of 20 966 Swedish twins. *J. Intern. Med.* **252**, 247–254 (2002).
- Schunkert, H. et al. Large-scale association analyses identifies 13 new susceptibility loci for coronary artery disease. *Nat. Genet.* **43**, 333–338 (2011).
- Abul-Husn, N. S. et al. Genetic identification of familial hypercholesterolemia within a single U.S. health care system. *Science* **354**, aaf7000 (2016).
- Deloukas, P. et al. Large-scale association analysis identifies new risk loci for coronary artery disease. *Nat. Genet.* **45**, 25–33 (2013).
- Nelson, C. P. et al. Association analyses based on false discovery rate implicate new loci for coronary artery disease. *Nat. Genet.* **49**, 1385–1391 (2017).
- Aragam, K. G. & Natarajan, P. Polygenic scores to assess atherosclerotic cardiovascular disease risk: Clinical perspectives and basic implications. *Circ. Res.* **126**, 1159–1177 (2020).
- Khera, A. V. et al. Genome-wide polygenic scores for common diseases identify individuals with risk equivalent to monogenic mutations. *Nat. Genet.* **50**, 1219–1224 (2018).
- Inouye, M. et al. Genomic risk prediction of coronary artery disease in 480,000 adults: implications for primary prevention. *J. Am. Coll. Cardiol.* **72**, 1883–1893 (2018).
- Aragam, K. G. et al. Limitations of contemporary guidelines for managing patients at high genetic risk of coronary artery disease. *J. Am. Coll. Cardiol.* **75**, 2769–2780 (2020).
- Bolli, A., Di Domenico, P., Pastorino, R., Busby, G. B. & Bottà, G. Risk of coronary artery disease conferred by low-density lipoprotein cholesterol depends on polygenic background. *Circulation* **143**, <https://www.ahajournals.org/doi/abs/10.1161/CIRCULATIONAHA.120.051843> (2021).
- Sud, A. et al. Realistic expectations are key to realising the benefits of polygenic scores. *BMJ* **380**, e073149 (2023).
- Martin, A. R., Daly, M. J., Robinson, E. B., Hyman, S. E. & Neale, B. M. Predicting polygenic risk of psychiatric disorders. *Biol. Psychiatry* **86**, 97–109 (2019).
- Fatumo, S. et al. A roadmap to increase diversity in genomic studies. *Nat. Med.* **28**, 243–250 (2022).
- Auton, A. et al. A global reference for human genetic variation. *Nature* **526**, 68–74 (2015).
- Ruan, Y. et al. Improving polygenic prediction in ancestrally diverse populations. *Nat. Genet.* **54**, 573–580 (2022).
- Hellenthal, G. et al. A genetic atlas of human admixture history. *Science* **343**, 747–751 (2014).
- Lewis, A. C. F. et al. Getting genetic ancestry right for science and society. *Science* **376**, 250–252 (2022).
- Weissbrod, O. et al. Leveraging fine-mapping and multipopulation training data to improve cross-population polygenic risk scores. *Nat. Genet.* **54**, 450–458 (2022).
- Bild, D. E. et al. Multi-ethnic study of atherosclerosis: objectives and design. *Am. J. Epidemiol.* **156**, 871–881 (2002).
- ARIC Investigators. The atherosclerosis risk in community (ARIC) study: design and objectives. *Am. J. Epidemiol.* **129**, 687–702 (1989).
- Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
- Goff, D. C. et al. 2013 ACC/AHA guideline on the assessment of cardiovascular risk: a report of the American College of Cardiology/

- American Heart Association Task Force on Practice Guidelines. *Circulation* **129**, S49–73 (2014).
25. Grundy, S. M. et al. 2018 AHA/ACC/AACVPR/AAPA/ABC/ACPM/ADA/AGS/APHA/ASPC/NLA/PCNA guideline on the management of blood cholesterol: executive summary: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *J. Am. Coll. Cardiol.* **73**, 3168–3209 (2019).
 26. Marston, N. A. et al. Predictive utility of a coronary artery disease polygenic risk score in primary prevention. *JAMA Cardiol.* **8**, 130–137 (2022).
 27. O’Sullivan, J. W., Ashley, E. A. & Elliott, P. M. Polygenic risk scores for the prediction of cardiometabolic disease. *Eur. Heart J.* **44**, 89–99 (2023).
 28. O’Sullivan, J. W. et al. Polygenic risk scores for cardiovascular disease: a scientific statement from the American Heart Association. *Circulation* **146**, e93–e118 (2022).
 29. Mujwara, D. et al. Integrating a polygenic risk score for coronary artery disease as a risk-enhancing factor in the Pooled Cohort Equation: A Cost-Effectiveness Analysis Study. *J. Am. Heart Assoc.* **11**, e025236 (2022).
 30. Senol-Cosar, O. et al. Considerations for clinical curation, classification, and reporting of low-penetrance and low effect size variants associated with disease risk. *Gen. Med.* **21**, 2765–2773 (2019).
 31. Hao, L. et al. Development of a clinical polygenic risk score assay and reporting workflow. *Nat. Med.* **28**, 1006–1013 (2022).
 32. Choi, S. W., Mak, T. S.-H. & O’Reilly, P. F. Tutorial: a guide to performing polygenic risk score analyses. *Nat. Protoc.* **15**, 2759–2772 (2020).
 33. Busby, G. B. et al. Admixture into and within sub-Saharan Africa. *eLife* **5**, e15266 (2016).
 34. Privé, F. et al. Portability of 245 polygenic scores when derived from the UK Biobank and applied to 9 ancestry groups from the same cohort. *Am. J. Hum. Genet.* **109**, 12–23 (2022).
 35. Kamiza, A. B. et al. Transferability of genetic risk scores in African populations. *Nat. Med.* **28**, 1163–1166 (2022).
 36. Mostafavi, H. et al. Variable prediction accuracy of polygenic scores within an ancestry group. *eLife* **9**, e48376 (2020).
 37. Riveros-Mckay, F. et al. Integrated polygenic tool substantially enhances coronary artery disease prediction. *Circ. Genom. Precis. Med.* **14**, e003304 (2021).
 38. Mars, N. et al. Genome-wide risk prediction of common diseases across ancestries in one million people. *Cell Genomics* **2**, 100118 (2022).
 39. Patel, A. P., Wang, M., Kartoun, U., Ng, K. & Khera, A. V. Quantifying and understanding the higher risk of atherosclerotic cardiovascular disease among South Asian individuals. *Circulation* **144**, 410–422 (2021).
 40. Kiflen, M. et al. Cost-effectiveness of polygenic risk scores to guide statin therapy for cardiovascular disease prevention. *Circ. Genom. Precis. Med.* **15**, e003423 (2022).
 41. Nikpay, M. et al. A comprehensive 1,000 Genomes-based genome-wide association meta-analysis of coronary artery disease. *Nat. Genet.* **47**, 1121–1130 (2015).
 42. Matsunaga, H. et al. Transethnic meta-analysis of genome-wide association studies identifies three new loci and characterizes population-specific differences for coronary artery disease. *Circ. Genom. Precis. Med.* **13**, e002670 (2020).
 43. Sakaue, S. et al. A cross-population atlas of genetic associations for 220 human phenotypes. *Nat. Genet.* **53**, 1415–1424 (2021).
 44. Jiang, L., Zheng, Z., Fang, H. & Yang, J. A generalized linear mixed model association tool for biobank-scale data. *Nat. Genet.* **53**, 1616–1621 (2021).
 45. Weissbrod, O. et al. Functionally informed fine-mapping and polygenic localization of complex trait heritability. *Nat. Genet.* **52**, 1355–1363 (2020).
 46. Gazal, S. et al. Functional architecture of low-frequency variants highlights strength of negative selection across coding and non-coding annotations. *Nat. Genet.* **50**, 1600–1607 (2018).
 47. Consortium, T. G. P. An integrated map of genetic variation from 1,092 human genomes. *Nature* **491**, 56–65 (2012).
 48. Zhao, B. & Zou, F. On polygenic risk scores for complex traits prediction. *Biometrics* **78**, 499–511 (2021).
 49. Dudbridge, F. Power and predictive accuracy of polygenic risk scores. *PLoS Genet.* **9**, e1003348 (2013).
 50. Clarke, T.-K. et al. Genome-wide association study of alcohol consumption and genetic overlap with other health-related traits in UK biobank (n=12117). *Mol. Psychiatry* **22**, 1376–1384 (2017).
 51. Ritchie, S. C. et al. Quality control and removal of technical variation of NMR metabolic biomarker data in ~120,000 UK biobank participants. *Scientific Data* **10**, <https://doi.org/10.1038/s41597-023-01949-y> (2023).
 52. Ong, J.-S. et al. Height and overall cancer risk and mortality: evidence from a mendelian randomisation study on 310,000 UK biobank participants. *Br. J. Cancer* **118**, 1262–1267 (2018).
 53. Tasa, T., Puustusmaa, M., Tonisson, N., Kolk, B. & Padrik, P. Recommendations for primary prevention of skin melanoma. <https://doi.org/10.1101/2020.08.25.20181610> (2020).
 54. Wilcox, N. et al. Exome sequencing identifies breast cancer susceptibility genes and defines the contribution of coding variants to breast cancer risk. *Nat. Genet.* <https://doi.org/10.1038/s41588-023-01466-z> (2023).
 55. Klein, R. J. et al. Prostate cancer polygenic risk score and prediction of lethal prostate cancer. *npj Precis. Oncol.* **6**, <https://doi.org/10.1038/s41698-022-00266-8> (2022).
 56. Bansal, V. & Libiger, O. Fast individual ancestry inference from DNA sequence data leveraging allele frequencies for multiple populations. *BMC Bioinform.* **16**, <https://doi.org/10.1186/s12859-014-0418-7> (2015).
 57. Rabinowitz, J. A. et al. Positive associations between cannabis and alcohol use polygenic risk scores and phenotypic opioid misuse among african-americans. *PLoS ONE* **17**, e0266384 (2022).
 58. Namjou, B. et al. Multiancestral polygenic risk score for pediatric asthma. *J. Allergy Clin. Immunol.* **150**, 1086–1096 (2022).
 59. Chande, A. T. et al. The impact of ethnicity and genetic ancestry on disease prevalence and risk in colombia. *Front. Genet.* **12**, <https://doi.org/10.3389/fgene.2021.690366> (2021).
 60. Maier, R. M. et al. Improving genetic prediction by leveraging genetic correlations among human diseases and traits. *Nat. Commun.* **9**, <https://doi.org/10.1038/s41467-017-02769-6> (2018).
 61. Privé, F., Vilhjálmsson, B. J., Aschard, H. & Blum, M. G. Making the most of clumping and thresholding for polygenic scores. *Am. J. Hum. Genet.* **105**, 1213–1221 (2019).
 62. Patel, A. P. et al. A multi-ancestry polygenic risk score improves risk prediction for coronary artery disease. *Nat. Med.* **29**, 1793–1803 (2023).
 63. Inouye, M. et al. Genomic risk prediction of coronary artery disease in 480,000 adults. *J. Am. Coll. Cardiol.* **72**, 1883–1893 (2018).
 64. Khera, A. V. et al. Polygenic prediction of weight and obesity trajectories from birth to adulthood. *Cell* **177**, 587–596.e9 (2019).
 65. Grimm, O. et al. Impulsivity and venturesomeness in an adult ADHD sample: relation to personality, comorbidity, and polygenic risk. *Front. Psychiatry* **11**, <https://doi.org/10.3389/fpsy.2020.557160> (2020).
 66. Fritsche, L. G. et al. On cross-ancestry cancer polygenic risk scores. *PLoS Genet.* **17**, e1009670 (2021).

67. Lu, X. et al. A polygenic risk score improves risk stratification of coronary artery disease: a large-scale prospective chinese cohort study. *Eur. Heart J.* **43**, 1702–1711 (2022).
68. Mrquez-Luna, C. & Loh, P.-R., South Asian Type 2 Diabetes (SAT2D) Consortium, SIGMA Type 2 Diabetes Consortium & Price, A. L. Multiethnic polygenic risk scores improve risk prediction in diverse populations. *Genet. Epidemiol.* **41**, 811–823 (2017).
69. Elliott, J. et al. Predictive accuracy of a polygenic risk score-enhanced prediction model vs a clinical risk score for coronary artery disease. *JAMA* **323**, 636–645 (2020).

Acknowledgements

We thank all individuals who contributed data to the publicly available datasets utilized in this study. UK Biobank data was accessed under application ID 40692. This study was funded by Allelica Inc.

Author contributions

Concept and design: G.B.B., S.K., P.D., and G.B. Acquisition, analysis or interpretation of data: G.B.B., S.K., A.B., J.K., P.D., and G.B. Drafting of the manuscript: G.B.B. and S.K. Critical revision of the manuscript for important intellectual content: G.B.

Competing interests

All authors are employees of Allelica, Inc.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-023-42897-w>.

Correspondence and requests for materials should be addressed to George B. Busby or Giordano Bottà.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023