Differences in estimates for 10-year risk of cardiovascular disease in Black versus White individuals with identical risk factor profiles using pooled cohort equations: an in silico cohort study

Ramachandran S Vasan, Edwin van den Heuvel

Summary

Background Sex-specific and race-specific pooled cohort equations (PCEs) are recommended for estimating the 10-year risk of cardiovascular disease, with an absolute risk of more than 7·5% indicating a clinical decision threshold. We compared differences between Black and White individuals in PCE-estimated absolute cardiovascular disease risk across various plausible risk factor combinations with the aim of evaluating if using the PCE might result in different clinical decisions in Black versus White individuals with identical risk profiles.

Methods We generated in silico patient risk profiles by combining numerical risk factors (age [5-year intervals], total cholesterol [20-mg/dl intervals], HDL cholesterol [5-mg/dl intervals], systolic blood pressure [10-mm Hg intervals]) and binary risk factors (smoking, diabetes, and antihypertensive treatment). We compared PCE-estimated 10-year cardiovascular disease risk in Black versus White individuals with identical risk profiles. We did similar comparisons using eligible participants in the Framingham Heart Study (FHS) third generation cohort and the National Health and Nutrition Examination Survey (NHANES) 2017–18.

Findings For our in silico analysis, we evaluated 29 515 risk factor combinations for women and 30 565 for men, after excluding profiles that generated 10-year cardiovascular disease risk estimates below 1% or above 30%. There were 6357 risk profiles associated with 10-year cardiovascular disease risk above 7·5% for Black men but not for White men (median risk difference [RD] 6·25%, range 0·15–22·8; median relative risk [RR] 2·40, range 1·02–12·6). There were 391 profiles with 10-year cardiovascular disease risk above 7·5% for White men but not Black men (median RD 2·68%, range 0·07–16·9%; median RR 1·42, range 1·01–3·57). There were 6543 risk profiles associated with 10-year estimated cardiovascular disease risk above 7·5% for Black women but not for White women (median RD 6·14%, range 0·35–26·8%; median RR 2·29, range 1·05–12·6). There were 318 profiles with 10-year cardiovascular disease risk above 7·5% for White women but not Black women (median RD 3·71%, range 0·22–20·1%; median RR 1·66, range 1·03–5·46). For the population-based samples, we calculated the PCE-based 10-year cardiovascular disease risk for 1272 eligible participants (378 women; median age 48 years [IQR 44–52]; 100% White) in the FHS third generation cohort and 550 participants (223 women [36·8% Black] and 327 men [40·4% Black]; median age 61 years [IQR 52–67]) in the NHANES cohort. The population-based samples showed similar risk differences to that of the in silico analyses.

Interpretation The PCE might generate substantially divergent cardiovascular disease risk estimates for Black versus White individuals with identical risk profiles, which could introduce race-related variations in clinical recommendations for cardiovascular disease prevention.

Funding US National Institutes of Health.

Lancet Digit Health 2022; 4: e55–63

Boston University's and National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, MA, USA (Prof R S Vasan MD); **Department of Medicine, School of Medicine** (Prof R S Vasan, Prof E van den Heuvel PhD) **and Department of Epidemiology, School of Public Health** (Prof R S Vasan), **Boston University, Boston, MA, USA; Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, Netherlands** (Prof E van den Heuvel)

Correspondence to: Prof Ramachandran S Vasan, Boston University School of Medicine, Boston, MA 02118, USA
vasan@bu.edu

Introduction

In 2013, the American College of Cardiology (ACC) and the American Heart Association (AHA) formulated sex-specific and race-specific equations—known as pooled cohort equations (PCEs)—for estimating the 10-year risk of atherosclerotic cardiovascular disease using representative community-based cohorts with White and Black individuals.[1] The PCEs represent an important methodological improvement over other risk scores (such as the Framingham and Reynold's risk scores) that are based predominantly on observations in White individuals. The PCE incorporates race and standard vascular risk factors—ie, sex, age, systolic blood pressure, antihypertensive treatment, blood total cholesterol and high-density lipoprotein cholesterol concentrations, diabetes, and smoking. In 2019, the ACC–AHA endorsed the use of PCE in primary care settings to guide clinical decisions, such as the

Research in context

Evidence before this study

No formal systematic literature review was done. When used in primary care settings, the pooled cohort equations (PCEs) for risk prediction can result in both under-estimation and over-estimation of cardiovascular disease risk. These prediction equations include race, a social construct that might not be a causal biological risk factor for cardiovascular disease. It is unknown if using the PCEs in Black and White individuals with identical risk factor profiles can result in substantial differences in their predicted 10-year cardiovascular disease risk. It is also unclear whether these race-related differences in predicted cardiovascular disease risk among individuals with an identical risk factor profile could translate into the differential treatment of Black versus White individuals.

Added value of this study

Our in silico analyses evaluating 51 840 potential risk factor combinations and our counterfactual prediction in two population-based samples showed substantial differences in estimated 10-year cardiovascular disease risk (using PCEs) in Black versus White individuals with identical risk factor profiles. The magnitude of these predicted risk differences seems biologically implausible based on race alone. Furthermore, these race-related differences in predicted cardiovascular disease risk could result in the differential prescription of statins in Black versus White individuals with identical risk profiles. Such a result conflicts with causal definitions of fairness according to which comparable individuals should not be treated differentially based on attributes that might predispose them to discrimination, such as a specific race or ethnicity.

Implications of all the available evidence

The PCEs should be improved to preclude differential treatment of Black and White individuals with identical risk factor profiles. There is a need to replace the race term in cardiovascular disease prediction equations with causal factors that are associated with race (such as structural racism, health-care access, education, economic challenges, and other social determinants of health), which might mediate the risk differences among Black versus White individuals with identical risk factor profiles.

prescription of statins for cardiovascular disease prevention.[2]

In parallel, an emerging body of literature[3–10] has strongly criticised the use of race in medical risk assessment for several acute and chronic diseases because of concerns regarding algorithmic fairness,[11] including the risk of exacerbating health inequities. Others have emphasised that prediction equations explicitly intended to guide treatment decisions should incorporate causal risk factors.[12–14] Predictor variables in the PCE are biological in nature, except race, which is a social construct.[3] Incorporating race into PCE might equate racial differences in predicted cardiovascular disease risk with true biological differences in disease susceptibility between the races, translating into over-treatment versus under-treatment with pharmacological agents (statins or aspirin) of one racial group versus the others.

To our knowledge, this is the first and largest in silico analysis of the potential for race-related differential clinical decisions resulting from the use of PCE for cardiovascular disease risk assessments. We aimed to evaluate this premise by comparing differences between Black and White individuals in PCE-estimated absolute cardiovascular disease risk across various plausible risk factor combinations using an in silico approach, complemented by data from two distinct community-based studies.

See Online for appendix

Methods

We used hypothetical data for in silico analyses. We evaluated two community-based cohorts that vary in their racial composition, are geographically distinct, recruited in different calendar decades, and encompass a broad age range. For the first cohort, we used de-identified, publicly available data from the National Health and Nutrition Examination Survey (NHANES; recruited 2017–18). The NHANES cohort is a more contemporary, multi-racial cohort (34% Black individuals), represents a US national probability sample, and has a mean age of 49·9 years (SD 18·8). For the second cohort, we used data from the Framingham Heart Study (FHS) third generation cohort participants at their first examination cycle (recruited in 2002–05). The FHS third generation cohort is a cohort of White individuals predominantly living in Massachusetts and the greater New England region in the USA, with a mean age of 40·2 years (SD 8·8).

All FHS participants provided written informed consent, and the Institutional Review Board at the Boston University Medical Center approved the study protocol.

Creation of a SAS macro using risk functions from the PCE

We extracted the risk prediction functions for the 10-year risk of atherosclerotic cardiovascular disease for the combined strata of sex and race from table A of the 2013 ACC–AHA guidelines.[1] We created a SAS macro that programs the risk functions specified by the PCE using the prediction equations provided by Goff and colleagues.[1] Full details of the SAS macro are given in the appendix (pp 19–21).

Creation of risk factor categories and their combinations

As specified for the ACC web-based risk estimator, we considered a wide range of permissible values for cardiovascular disease risk factors.[15] We created profiles by

combining risk factors as follows: age, 40 to ≤80 years in 5-year increments (eight categories); systolic blood pressure, 100–200 mm Hg in 10-mm Hg increments (ten categories); total cholesterol concentrations, 130–290 mg/dL in 20-mg/dL increments (eight categories); HDL cholesterol concentrations, 20–90 mg/dL in 10-mg/dL increments (seven categories); diabetes, yes versus no (two categories); smoking status, yes versus no (two categories); and treatment for elevated systolic blood pressure, yes versus no (two categories). We restricted values of treated systolic blood pressure empirically to a maximum of 180 mm Hg. For each of four strata (men *vs* women; Black versus White individuals), we created 51840 possible risk factor combinations (also referred to as risk profiles).

Estimation of 10-year cardiovascular disease risk for risk factor combinations with PCE

We calculated the 10-year absolute cardiovascular disease risk for each of the 51840 risk factor combinations by inputting risk factor values into the published PCE risk functions, as detailed in the appendix (pp 19–21).[1] We excluded risk profiles that yielded 10-year cardiovascular disease risk estimates that were below 1% or above 30%, as recommended by the 2013 ACC–AHA guidelines.[1]

Differences in risk for Black versus White individuals with the identical risk factor combinations

All analyses were sex specific. First, we calculated differences in the PCE-based estimates of 10-year absolute risk and the relative risk of a cardiovascular disease event for Black versus White individuals with identical risk factor combinations. Next, we evaluated four possible scenarios (two for each sex) in which the 10-year absolute cardiovascular disease risk estimates for the two races were on opposite sides of the critical 7·5% threshold for 10-year cardiovascular disease risk that triggers clinical decisions when exceeded (ie, discussions with patients regarding initiating treatment with statins).[1,2]

For each scenario, we plotted histograms to describe the sex-specific distributions of the differences in absolute and relative risks (of cardiovascular disease) for Black versus White individuals. We identified specific risk factor combinations that yielded divergent estimates of 10-year cardiovascular disease risk for Black versus White individuals, including those that maximised these risk differences. We used sex-specific box plots to visualise the distributions of numerical risk factors (age, systolic blood pressure, total cholesterol, and HDL cholesterol) for these combinations. We also formulated sex-specific box plots showing differences in 10-year absolute cardiovascular disease risk between Black versus White individuals according to the categories of each of the numerical risk factors.

Additionally, we repeated all our analyses using risk factor combinations within the normal range (ie, for individuals without diabetes, high blood pressure, or history of smoking), and for values of systolic blood pressure 100–130 mm Hg, total cholesterol concentrations 130–170 mg/dl, and HDL concentrations 40–70 mg/dl in men or 50–80 mg/dl in women.

Analysis of FHS data

The design and selection criteria of the third generation of the FHS cohort have been detailed elsewhere.[16] All participants in the cohort self-reported as White. Eligible participants attended their first examination cycle and had a complete risk factor profile. We excluded individuals who were younger than 40 years, those with risk factor values outside the range of our in silico analyses, and those with PCE-based risk estimates outside the recommended range per ACC–AHA guidelines.[1] We calculated the PCE-based 10-year cardiovascular disease risk for eligible participants in the FHS third generation cohort[16] using their risk factor data and their theoretical Black counterfactuals of the same sex, assuming they had an identical risk factor profile. We displayed the differences in absolute cardiovascular disease risk for the four scenarios where White participants versus their Black counterfactuals had absolute cardiovascular disease risk on opposite sides of the 7·5% risk threshold, paralleling our in silico approach.

Analysis of NHANES 2017–18 data

We accessed NHANES 2017–18 public-use data.[17] Eligible participants had complete risk factor profiles. We excluded individuals with risk factor values outside the range for our in silico analysis and individuals with PCE-based risk estimates outside the recommended range per guidelines.[1] Participants in NHANES self-report their race (from a set of fixed choices). We calculated the PCE-based 10-year cardiovascular disease risk for Black and White participants and their same-sex counterfactuals with identical risk profiles. We displayed the differences in the PCE-estimated 10-year absolute cardiovascular disease risk for these participants and their counterfactuals where the factual and the counterfactual 10-year cardiovascular disease risks were on opposite sides of the 7·5% risk threshold, consistent with our in silico approach.

Statistical analysis

All analyses were sex specific. For the in silico analysis, we first created a dataset of risk factor profiles for which we calculated the predicted cardiovascular disease risk for Black and White participants. Next, we created two new datasets, one where Black individuals have a PCE-estimated 10-year cardiovascular disease risk exceeding 7·5% but White individuals do not, and another dataset where the converse was true. For each dataset, we calculated the PCE-estimated cardiovascular disease risk difference and the relative risk (where the race with the lower risk was the referent). All analyses on

these datasets were descriptive, using means, medians, SDs, and related visualisations for risk differences and relative risks. The complete SAS code for the analysis is provided in the appendix (pp 19–40).

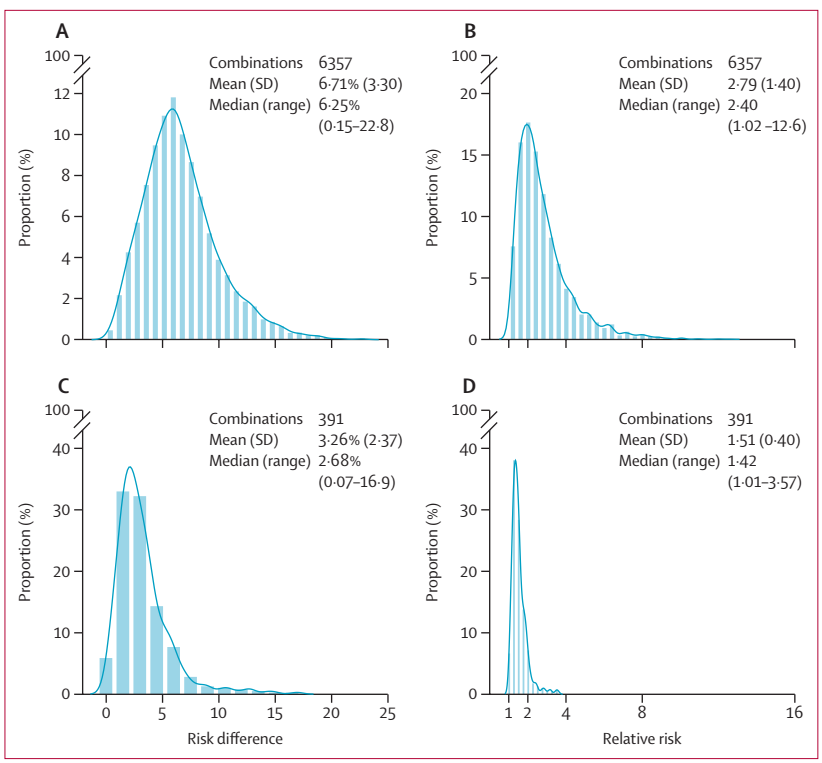

Figure 1: **Distributions of risk differences and risk ratios for Black versus White men with divergent risks**
(A) Differences in PCE-based 10-year cardiovascular disease risk in men when absolute cardiovascular disease risk exceeds 7·5% in Black but not in White men. (B) Relative risks (Black *vs* White men) for 10-year incidence of cardiovascular disease when PCE-based absolute risk exceeds 7·5% in Black but not in White men. (C) Differences in PCE-based 10-year cardiovascular disease risk in men when absolute cardiovascular disease risk exceeds 7·5% in White but not in Black men. (D) Relative risks (White *vs* Black men) for 10-year incidence of cardiovascular disease when PCE-based absolute risk exceeds 7·5% in White but not in Black men. PCE=pooled cohort equation.

Role of the funding source
The funder had no role in study design, data collection, data analysis, data interpretation, or writing of the report.

Results

For our in silico analysis, we evaluated 29 515 risk factor combinations for women and 30 565 for men (appendix p 4), after excluding profiles that generated 10-year cardiovascular disease risk estimates below 1% or above 30%. In both sexes, 45% of the putative risk factor combinations included smoking, whereas 40% included diabetes.

We evaluated the extent of divergence in PCE-estimated 10-year CVD risks in Black versus White individuals with identical risk factor profiles. There were 6357 risk factor combinations where a Black man had an estimated 10-year absolute cardiovascular disease risk exceeding 7·5%, but a White man with an identical risk factor profile had an estimated risk below that threshold (appendix p 5); the proportions of smoking and diabetes were 39·7% and 49·4%, respectively. Differences in absolute cardiovascular disease risk between Black and White individuals can be as large as 22·8% (median 6·25%; figure 1A), and the Black versus White relative risk for cardiovascular disease can be as large as 12·6 (median 2·40; figure 1B).

There were 391 risk factor combinations where a White man had an estimated 10-year absolute cardiovascular disease risk exceeding 7·5%, but a Black man with an identical risk factor profile had an estimated risk below that threshold (appendix p 5); the proportions of smoking and diabetes were 32·5% and 7·2%, respectively. White versus Black differences in absolute cardiovascular disease risk can be as large as 16·9% (median 2·68%; figure 1C), and the White versus Black relative risk of cardiovascular disease can be as large as 3·57 (median 1·42; figure 1C).

Table 1 shows specific risk factor combinations that yield maximal differences in absolute and relative risks

	Age, years	Hypertension	Smoking	Diabetes	Systolic blood pressure, mm Hg	Total cholesterol, mg/dl	HDL cholesterol, mg/dl	10-year risk, Black	10-year risk, White	Maximal value
Black men are at risk* but White men are not										
Black *vs* White, risk difference	45	1	1	1	180	130	80	28·84%	6·00%	22·84%
Black *vs* White, relative risk	40	1	0	1	180	150	90	13·76%	1·10%	12·55
White men are at risk* but Black men are not										
White *vs* Black, risk difference	40	0	1	0	120	290	20	6·92%	23·87%	16·95%
White *vs* Black, relative risk	40	0	1	0	100	290	20	5·03%	17·94%	3·57
Black women are at risk* but White women are not										
Black *vs* White, risk difference	40	1	0	1	160	130	30	29·93%	3·17%	26·75%
Black *vs* White, relative risk	40	1	0	1	180	130	40	28·98%	2·30%	12·60
White women are at risk* but Black women are not										
White *vs* Black, risk difference	80	1	0	0	120	130	20	6·38%	26·43%	20·06%
White *vs* Black, relative risk	40	0	1	0	100	290	30	1·83%	9·99%	5·46

For binary risk factors, 1=present and 0=absent. PCE=pooled cohort equation. *>7·5% 10-year risk of atherosclerotic cardiovascular disease as defined by the PCEs.

Table 1: **Risk factor combinations for which the difference in PCE-estimated 10-year risk of cardiovascular disease is maximal for Black versus White individuals (or vice versa)**

of cardiovascular disease between Black versus White men, within the constraints of our inclusion criteria. The appendix (pp 9–10) shows the box plot distributions of numerical risk factors for combinations that yield divergent (and convergent) cardiovascular disease risk estimates for Black versus White men. Box plots are also shown for the Black versus White differences in absolute cardiovascular disease risk estimates for various risk factor categories (appendix pp 11–13). Risk factor combinations that yield an estimated 10-year cardiovascular disease risk of 7·5% or more for Black, but not White, men typically occur at younger ages, with a higher frequency of non-smokers and higher systolic blood pressure values. Risk factor combinations that yield an absolute 10-year cardiovascular disease risk exceeding 7·5% for White, but not Black, men are associated with a high prevalence of hypertension treatment, lower HDL levels, and the absence of diabetes and smoking. Even with risk factor values within the normal range, some profiles can yield moderate to large differences in absolute cardiovascular disease risk between Black and White individuals (appendix p 6).

There were 6543 risk factor combinations associated with a 10-year absolute risk estimate greater than 7·5% in Black women, but not in White women with identical risk factor profiles (appendix p 7); the proportions of smoking and diabetes were 31·1% and 48·3%, respectively. Differences between Black and White individuals in 10-year absolute and relative risk of cardiovascular disease for these risk factor combinations are shown in figure 2. The differences in absolute cardiovascular disease risk between Black and White individuals can be as large as 26·8% (median 6·14%; figure 2A), and the difference in relative risk of cardiovascular disease can be as large as 12·6 (median 2·29; figure 2B).

There were 318 risk factor combinations associated with estimated 10-year absolute cardiovascular disease risk that exceeds 7·5% in White, but not in Black, women (appendix p 7); the proportions of smoking and diabetes were 68·2% and 38·7%, respectively. White versus Black differences in estimated absolute cardiovascular disease risk can be as large as 20·1% (median 3·17%; figure 2C), and the relative risks of cardiovascular disease between White and Black individuals can be as large as 5·46 (median 1·66; figure 2D).

Table 1 shows select risk factor combinations that yield maximal differences between Black and White individuals in estimated absolute and relative risks of cardiovascular disease in women, within the boundaries of our inclusion criteria. Box plots for numerical risk factors for the combinations that yield divergent (and convergent) estimated absolute cardiovascular disease risk for Black versus White women are shown in the appendix (pp 14–15). Box plots are also shown for the differences in PCE-estimated 10-year cardiovascular disease risk for Black versus White women for the individual risk factor categories (appendix pp 16–18). Estimated 10-year

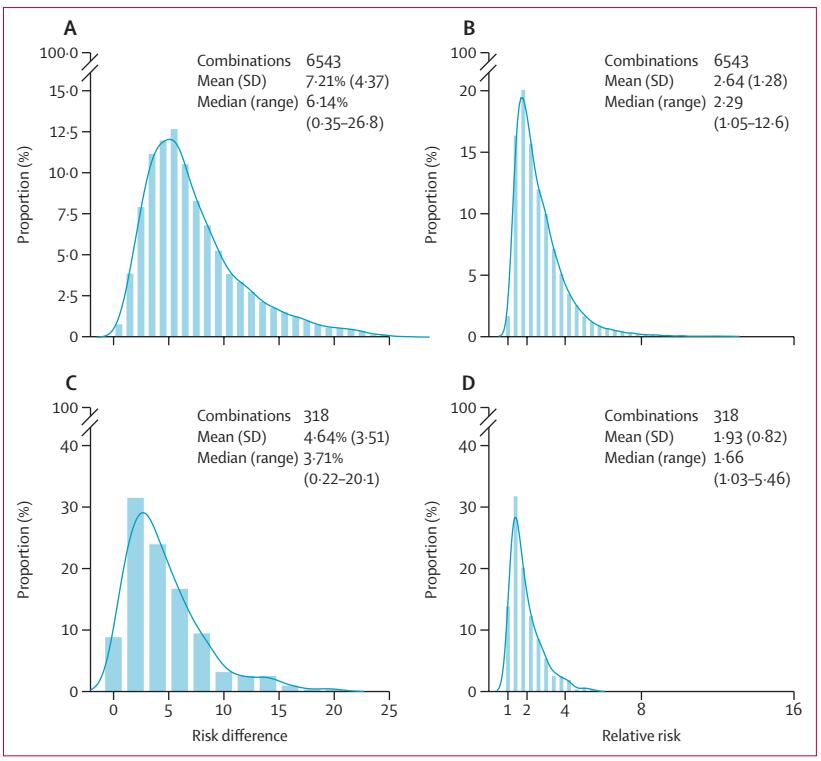

Figure 2: **Distributions of risk differences and risk ratios for Black versus White women with divergent risks**
(A) Differences in PCE-based 10-year cardiovascular disease risk in women when absolute cardiovascular disease risk exceeds 7·5% in Black but not in White women. (B) Relative risks (Black *vs* White women) for 10-year incidence of cardiovascular disease when PCE-based absolute risk exceeds 7·5% in Black but not in White women. (C) Differences in PCE-based 10-year cardiovascular disease risk in women when absolute cardiovascular disease risk exceeds 7·5% in White but not in Black women. (D) Relative risks (White *vs* Black women) for 10-year incidence of cardiovascular disease when PCE-based absolute risk exceeds 7·5% in White but not in Black women. PCE=pooled cohort equation.

cardiovascular disease risk exceeded 7·5% in Black, but not White, women more frequently in non-smoking adults and when diabetes is present. Conversely, a higher calculated absolute cardiovascular disease risk for White, but not Black, women occurred more frequently among smokers. Even across the normal range of risk factors, we observed moderate differences in absolute cardiovascular disease risk estimates for Black versus White women (appendix p 8).

Sensitivity analyses using a narrow set of increments for risk factor categories generated a much larger number of risk factor combinations, but our overall results remained robust (data not shown).

Of 4073 participants attending their first examination cycle with complete risk factor profiles (appendix p 2), we excluded 1897 individuals younger than 40 years, 258 participants with risk factor values outside the range for our in silico analysis, and 646 individuals (611 for participants and 35 for Black counterfactuals) with PCE-based risk estimates outside the recommended range per ACC–AHA guidelines.[1]

Of 1272 eligible FHS participants (378 women; median age 48 years [IQR 44–52; appendix p 2), 121 White

participants (9·5%; 34 women) had a PCE-estimated 10-year cardiovascular disease risk below 7·5%, whereas their Black counterfactuals had an estimated absolute cardiovascular disease risk above that threshold. Conversely, 17 White individuals (1·3%; two women) had a PCE-estimated 10-year cardiovascular disease risk above 7·5%, whereas their Black counterfactuals had an estimated absolute risk below that threshold. Figure 3 shows the distributions of estimated absolute cardiovascular disease risk differences between White FHS participants and their Black counterfactuals.

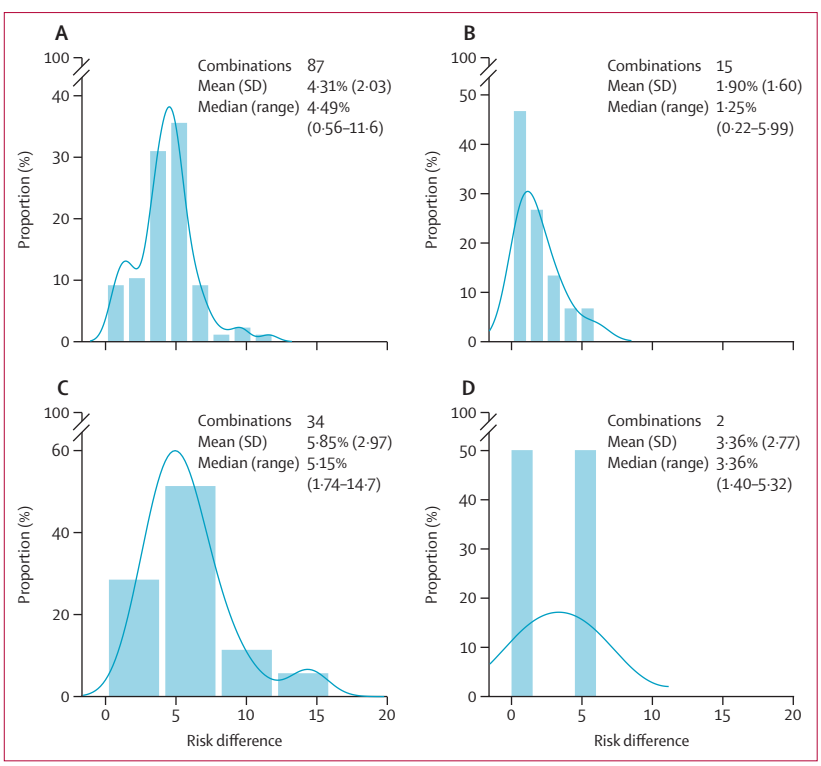

Figure 3: **Distribution of risk differences in PCE-based 10-year cardiovascular disease risk for White versus Black individuals, Framingham Heart Study**
Counterfactuals refers to theoretical participants who have the same risk factor combinations as the actual Framingham Heart Study participants (factual) but belong to the comparison race of interest. (A) Risk differences when absolute cardiovascular disease risk does not exceed 7·5% in White men but exceeds this threshold in their Black counterfactuals. (B) Risk differences when absolute cardiovascular disease risk exceeds 7·5% in White men but does not exceed this threshold in their Black counterfactuals. (C) Risk differences when absolute cardiovascular disease risk does not exceed 7·5% in White women but exceeds this threshold in their Black counterfactuals. (D) Risk differences when absolute cardiovascular disease risk exceeds 7·5% in White women but does not exceed this threshold in their Black counterfactuals. PCE=pooled cohort equation.

Of 1065 NHANES participants with complete risk factor profiles (appendix p 3), we excluded 365 individuals with risk factor values outside the range for our in silico analysis and 150 individuals with PCE-based risk estimates outside the recommended range per guidelines.[1] Of 550 eligible NHANES participants (223 women [36·8% Black] and 327 men [40·4% Black]; median age 61 years [IQR 52–67]; appendix p 3), 34 White participants (10·1%; 19 women) had a PCE-estimated 10-year cardiovascular disease risk below 7·5%, whereas their Black counterfactuals had an estimated absolute cardiovascular disease risk above that threshold (table 2). Conversely, seven White individuals (2·1%; one woman) had a PCE-estimated 10-year cardiovascular disease risk above 7·5%, whereas their Black counterfactuals had an estimated absolute risk below that threshold (table 2). Furthermore, 31 Black participants (14·5%, 19 women) had an estimated cardiovascular disease risk above 7·5%, whereas their White counterfactuals had an estimated cardiovascular disease risk below 7·5% (table 2). Only one Black participant, a woman, had an estimated risk below 7·5%, with a counterfactual White risk of cardiovascular disease above 7·5% (table 2).

Discussion

Our analysis yielded three main findings. First, PCE can result in major differences in predicted cardiovascular disease risk for Black versus White individuals who have identical risk factor profiles. This situation might frequently occur, as seen both in silico and in two independent community-based samples recruited in different calendar decades, encompassing a wide age range and geographical diversity. In males and females, PCE-estimated cardiovascular disease risk estimates for Black individuals exceeded that for White individuals with the same risk profile more frequently than the converse. These differences between Black and White individuals in risk (median 6%) and relative risks

	Counterfactual*	Sex	Black >7·5% and White <7·5%†		Black <7·5% and White >7·5%†	
			n (%)‡	Median risk difference (range)	n (%)‡	Median risk difference (range)
Black	White	Female	19 (23·2%)	3·82 (0·94–7·03)	1 (1·2%)	2·63 (2·63–2·63)
Black	White	Male	12 (9·1%)	2·13 (0·71–7·25)	0	NA
White	Black	Female	19 (13·5%)	3·72 (1·26–7·32)	1 (0·7%)	0·45 (0·45–0·45)
White	Black	Male	15 (7·7%)	2·38 (0·57–9·16)	6 (3·1%)	1·77 (0·69–3·02)

NA=not applicable. NHANES=National Health and Nutrition Examination Survey. *Theoretical participants who have the same risk factor combinations as the actual NHANES participants (factual) but belong to the comparison race of interest. †Comparisons are between members of a race group (Black or White) against their counterfactuals belonging to the same sex stratum. ‡% refers to the percentage of participants in that race–sex stratum who have divergent risk prediction with respect to the total number of participants in that stratum. We evaluated data for 550 participants. Of 327 men, 132 were Black. Of 223 women, 82 were Black.

Table 2: **Summary statistics for divergent, pooled cohort equations-based, estimated 10-year cardiovascular risk in White and Black participants of the NHANES 2017–18 cycle**

(median 2·3) are substantial and seem biologically implausible.

Second, specific risk factor combinations in each sex might exacerbate differences between Black and White individuals in PCE-estimated absolute cardiovascular disease risks. In both sexes, higher PCE-based cardiovascular disease risk estimates for Black individuals—relative to White individuals with identical risk factor profiles—were seen in younger age groups, with higher systolic blood pressure, and when diabetes was present. In the older age groups and in individuals with low HDL concentrations, PCE-estimated cardiovascular disease risk was higher in White women than in Black women with identical risk factor profiles. Risk factor combinations with lower total cholesterol or higher HDL concentrations were associated with higher PCE-estimated cardiovascular disease risks in Black versus White men.

Third, the race-related differences in PCE-estimated cardiovascular disease risk might be clinically meaningful—ie, they could result in different treatment decisions (such as statin prescription) in individuals with identical risk factor profiles based solely on their race.

It is widely acknowledged that PCE can both underestimate and over-estimate cardiovascular disease risk in different contexts.[18–25] Yet, data are scant regarding how frequently PCE generates considerably divergent risk estimates for Black versus White individuals with identical risk factor profiles that could impact clinical decisions. Yadlowsky and colleagues[25] reported divergent PCE-derived cardiovascular disease risk estimates for Black versus White individuals in the NHANES 2013–14 sample. The authors noted that updating the PCE with data from more recent cohort samples and reducing model overfitting attenuated, but did not eliminate, these race-related differences.[25] To our knowledge, no previous report has evaluated whether specific risk factor combinations might exaggerate differences in PCE-estimated cardiovascular disease risk between Black and White individuals.

There are several clinical implications of our findings. First, the use of PCE could result in Black individuals with select risk factor combinations becoming more eligible for receiving statin treatment than their White counterparts with identical risk profiles. Although the direction of this potential bias might seem somewhat reassuring (relative to the opposite scenario of Black individuals not receiving statins relative to their White counterparts), the risks associated with over-treatment—ie, financial, psychological, side-effects, and quality of life—are not trivial.

Second, race is a social construct created by humans "to group individuals with certain observable physical characteristics, such as skin color or facial features, who evolved from different geographies in the world."[3] It is widely accepted that race is associated with health outcomes and health inequities, in part mediated by "exposure or vulnerability to behavioral, psychosocial, material, or environmental risk factors and resources."[26] Therefore, differences in PCE-based cardiovascular disease risk estimates between Black and White individuals could be a surrogate for structural racism, differences in health-care access, educational achievement and economic challenges, and other sources of health inequities.[27] As such, by using race in the PCE, we might be normalising and legitimising a social construct as a medically valid classifier (Black *vs* White), leading the uninitiated to equate race-related differences in estimated cardiovascular disease risk with actual biological differences in disease susceptibility. Therefore, race should be replaced in any risk prediction equation by the various potentially causal factors that race represents, and that can be targeted with interventions. If replaced by appropriate causal variables, race should no longer improve prediction in the risk algorithms. For example, at least three UK-based risk scores[28–30] incorporate a social deprivation index (instead of race and ethnicity) that more directly addresses the social determinants of health. Additional research should consider evaluating the performance of the PCE with a social deprivation index substituted for race.

Third, it is important to distinguish between risk prediction unrelated to interventional decisions (eg, for prognostication) and prediction equations tied to decisions for intervention (such as the PCE).[13] The latter requires a causal framework for multiple reasons,[14] including the importance of addressing true root causes, enhancing transportability of the prediction algorithm across settings,[12] and ensuring prediction invariance.[31] Prediction invariance refers to the concept that a prediction from a causal model will work as well under interventions as it does for observational data.[31]

Fourth, some scientists[32] have argued that, from an outcomes perspective, incorporating race in risk estimation algorithms might permit better concordance with the patient's own goals, thereby facilitating individualised and optimal care. On the other hand, causal definitions of individual fairness would argue that comparable individuals (ie, those with identical risk factor profiles) should not be treated differently based on attributes that might predispose them to discrimination, such as a particular race or ethnicity. Expert consensus is needed regarding how a clinician should optimally balance an algorithmic fairness framework against an outcomes-based approach.

Scientists have stipulated several criteria for the inclusion of race in clinical prediction algorithms, as follows: the race-based measures are reproducible at an individual level; there is a sound scientific rationale for a causal role of race in disease etiology that is supported by robust scientific and statistical evidence; data indicate that incorporation of race would mitigate rather than exacerbate harm for a group that is at greater risk of poor health outcomes; there is evidence that such benefit

cannot be achieved by other feasible means; it is transparent, and accommodates in fair manner individuals who reject race categorisation.[4,5,10]

Fifth, additional research is also needed to evaluate if refinements to the PCE are necessary for some scenarios (eg, for specific risk factor combinations) that are associated with race-related differences in risk estimates that are extreme and deemed to be biologically implausible. For example, risk calculators could offer either race-less predictions (using variables other than race, leaving the clinician to make their best judgment; the approach of fairness through unawareness)[3] or provide estimates of predicted risk for both Black and White individuals, regardless of the race of the person in front of the clinician. In the latter situation, the risk estimates for individuals of both races might be divergent but below a threshold where no clinical intervention is suggested (eg, <5% 10-year cardiovascular disease risk in both White and Black individuals) or above a level that requires action on the part of the clinician (eg, >10% 10-year cardiovascular disease risk in both White and Black individuals). In these circumstances, the race-related differences might result in non-polar decisions, rendering the variances somewhat less germane. However, there could be zones of 10-year cardiovascular disease risk estimates (eg, individuals at intermediate risk with a 5–10% PCE-based absolute risk) that might be conducive to differential clinical decisions in individuals from different race groups. The estimates from the cardiovascular disease risk calculator could be flagged to alert the clinician about entering a potential race-based medicine zone (ie, fairness through awareness). Examples of such flags could be a zone for risk estimates where between Black versus White individuals (or the converse) in absolute risk exceed 2·5%, or the corresponding relative risks exceed 1·5, a magnitude generally regarded as significant in epidemiology for most exposures.

Last, developing a causal framework for risk prediction involves estimating the risk for an individual whose risk factor profile we would like to alter with an intervention of known effect size in a group of individuals with similar characteristics; this requires counterfactual risk prediction, which could delineate the risk experienced by a similar individual with an altered (post-intervention) risk profile. Such counterfactual prediction will clarify whether, and to what extent, we can mitigate the higher cardiovascular disease risk experienced by Black individuals by intervening on their risk profile. The PCEs were developed using only observational, factual data and did not include the estimation of reduced risk after treatment (ie, counterfactual data). Thus, the PCE could be improved by including longitudinal data that consider risk modification with interventions and by replacing race with the underlying potentially causal factors that can be targeted.

We studied numerous risk factor combinations determined by our choice of pragmatic cut points for binning the continuous range of numerical risk factors. Our in silico approach facilitated the evaluation of PCE-based risk estimation across various risk factor profiles, which is not readily feasible using data from individual studies or pooled cohorts. Yet, our in silico results might be questioned as theoretical. Accordingly, we repeated our analyses using a range of risk factors within the normal range. We also evaluated two independent community-based samples spanning a wide age range to support our findings. Overall, our FHS and NHANES data analyses suggest that differences in the 10-year absolute risk of cardiovascular disease between Black and White individuals with identical profiles might frequently occur within the typical range of risk factors in community-dwelling ambulatory individuals.

In summary, the race term in the PCE can result in substantially divergent absolute and relative risks for cardiovascular disease for Black versus White individuals with identical risk factor profiles. Such divergence in estimated cardiovascular disease risk might introduce race-related variations in physician recommendations for the prevention of cardiovascular disease.

Contributors

RSV contributed to funding acquisition, project administration, methodology, and supervision. EvdH contributed to the formal data analysis, data curation, methodology, and data visualisation. Both authors conceptualised the study, contributed to the methodology, wrote and drafted the manuscript, had access to and verified all study data, and jointly decided to submit the manuscript for publication.

Declaration of interests

We declare no competing interests.

Data sharing

Data collected from the FHS and NHANES 2017–18, including de-identified individual participant data and corresponding data dictionaries defining each field in the datasets, are available to others at https://biolincc.nhlbi.nih.gov/studies/gen3/ and https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2017, respectively. The authors have made available in the appendix (pp 19–40) their program codes and macros for the calculation of cardiovascular disease risk using PCEs.

Acknowledgments

This work is supported by the National Heart, Lung, and Blood Institute (contracts NO1-HC-25195, HHSN268201500001I, 75N92019D00031, and U01HL146382). RSV is supported by the Evans Medical Foundation and the Jay and Louis Coffman Endowment from the Department of Medicine, Boston University School of Medicine.

References

1 Goff DC Jr, Lloyd-Jones DM, Bennett G, 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* 2014; **129** (suppl 2): S49–73.

2 Arnett DK, Blumenthal RS, Albert MA, et al. 2019 ACC/AHA Guideline on the Primary Prevention of Cardiovascular Disease: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Circulation* 2019; **140:** e596–646.

3 Powe NR. Black kidney function matters: use or misuse of race? *JAMA* 2020; **324:** 737–38.

4 Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight—reconsidering the use of race correction in clinical algorithms. *N Engl J Med* 2020; **383:** 874–82.

5 Norris KC, Eneanya ND, Boulware LE. Removal of race from estimates of kidney function: first, do no harm. *JAMA* 2021; **325:** 135–37.

6 Paulus JK, Kent DM. Predictably unequal: understanding and addressing concerns that algorithmic clinical prediction may increase health disparities. *NPJ Digit Med* 2020; **3:** 99.

7 Yudell M, Roberts D, DeSalle R, Tishkoff S. Science and society. Taking race out of human genetics. *Science* 2016; **351:** 564–65.

8 Obermeyer Z, Powers B, Vogeli C, Mullainathan S. Dissecting racial bias in an algorithm used to manage the health of populations. *Science* 2019; **366:** 447–53.

9 Diao JA, Wu GJ, Taylor HA, et al. Clinical implications of removing race from estimates of kidney function. *JAMA* 2021; **325:** 184–86.

10 Eneanya ND, Yang W, Reese PP. Reconsidering the consequences of using race to estimate kidney function. *JAMA* 2019; **322:** 113–14.

11 Mitchell S, Potash E, Barocas S, D'Amour A, Lum K. Algorithmic fairness: choices, assumptions, and definitions. *Annu Rev Stat Appl* 2021; **8:** 141–63.

12 Piccininni M, Konigorski S, Rohmann JL, Kurth T. Directed acyclic graphs and causal thinking in clinical risk prediction modeling. *BMC Med Res Methodol* 2020; **20:** 179.

13 Prosperi M, Guo Y, Sperrin M, et al. Causal inference and counterfactual prediction in machine learning for actionable healthcare. *Nat Mach Intell* 2020; **2:** 369–75.

14 Schooling CM, Jones HE. Clarifying questions about "risk factors": predictors versus explanation. *Emerg Themes Epidemiol* 2018; **15:** 10.

15 American College of Cardiology. ASCVD risk calculator. 2013. http://tools.acc.org/ASCVD-Risk-Estimator-Plus/#!/calculate/estimate/ (accessed April 16, 2021).

16 Splansky GL, Corey D, Yang Q, et al. The third generation cohort of the National Heart, Lung, and Blood Institute's Framingham Heart Study: design, recruitment, and initial examination. *Am J Epidemiol* 2007; **165:** 1328–35.

17 Centers for Disease Control. National Health and Nutrition Examination Survey 2017–2018. 2020. https://wwwn.cdc.gov/nchs/nhanes/continuousnhanes/default.aspx?BeginYear=2017 (accessed April 16, 2021).

18 Crowson CS, Gabriel SE, Semb AG, et al. Rheumatoid arthritis-specific cardiovascular risk scores are not superior to general risk scores: a validation analysis of patients from seven countries. *Rheumatology (Oxford)* 2017; **56:** 1102–10.

19 Dalton JE, Perzynski AT, Zidar DA, et al. Accuracy of cardiovascular risk prediction varies by neighborhood socioeconomic position: a retrospective cohort study. *Ann Intern Med* 2017; **167:** 456–64.

20 Muntner P, Colantonio LD, Cushman M, et al. Validation of the atherosclerotic cardiovascular disease pooled cohort risk equations. *JAMA* 2014; **311:** 1406–15.

21 DeFilippis AP, Young R, Carrubba CJ, et al. An analysis of calibration and discrimination among multiple cardiovascular risk scores in a modern multiethnic cohort. *Ann Intern Med* 2015; **162:** 266–75.

22 Kavousi M, Leening MJ, Nanchen D, et al. Comparison of application of the ACC/AHA guidelines, Adult Treatment Panel III guidelines, and European Society of Cardiology guidelines for cardiovascular disease prevention in a European cohort. *JAMA* 2014; **311:** 1416–23.

23 Rana JS, Tabada GH, Solomon MD, et al. Accuracy of the atherosclerotic cardiovascular risk equation in a large contemporary, multiethnic population. *J Am Coll Cardiol* 2016; **67:** 2118–30.

24 Cook NR, Ridker PM. Further insight into the cardiovascular risk calculator: the roles of statins, revascularizations, and underascertainment in the Women's Health Study. *JAMA Intern Med* 2014; **174:** 1964–71.

25 Yadlowsky S, Hayward RA, Sussman JB, McClelland RL, Min Y-I, Basu S. Clinical implications of revised pooled cohort equations for estimating atherosclerotic cardiovascular disease risk. *Ann Intern Med* 2018; **169:** 20–29.

26 Williams DR, Lavizzo-Mourey R, Warren RC. The concept of race and health status in America. *Public Health Rep* 1994; **109:** 26–41.

27 Bailey ZD, Krieger N, Agénor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: evidence and interventions. *Lancet* 2017; **389:** 1453–63.

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

29 Joint British Societies Board. Joint British Societies' consensus recommendations for the prevention of cardiovascular disease (JBS3). *Heart* 2014; **100** (suppl 2)**:** 1–67.

30 Woodward M, Brindle P, Tunstall-Pedoe H. Adding social deprivation and family history to cardiovascular risk assessment: the ASSIGN score from the Scottish Heart Health Extended Cohort (SHHEC). *Heart* 2007; **93:** 172–76.

31 Peters J, Bühlmann P, Meinshausen N. Causal inference by using invariant prediction: identification and confidence intervals. *J R Stat Soc Series B Stat Methodol* 2016; **78:** 947–1012.

32 Paulus JK, Kent DM. Race and ethnicity. *Circ Cardiovasc Qual Outcomes* 2017; **10:** e003823.