Race as a Component of Cardiovascular Disease Risk Prediction Algorithms

Current Cardiology Reports (2023) 25:1131–1138
https://doi.org/10.1007/s11886-023-01938-y

ISCHEMIC HEART DISEASE (D MUKHERJEE, SECTION EDITOR)

Race as a Component of Cardiovascular Disease Risk Prediction Algorithms

Ramachandran S. Vasan[1,2] · Shreya Rao[1] · Edwin van den Heuvel[2,3]

Accepted: 7 August 2023 / Published online: 15 August 2023

Abstract

Purpose of Review Several prediction algorithms include race as a component to account for race-associated variations in disease frequencies. This practice has been questioned recently because of the risk of perpetuating race as a biological construct and diverting attention away from the social determinants of health (SDoH) for which race might be a proxy. We evaluated the appropriateness of including race in cardiovascular disease (CVD) prediction algorithms, notably the pooled cohort equations (PCE).

Recent Findings In a recent investigation, we reported substantial and biologically implausible differences in absolute CVD risk estimates upon using PCE for predicting CVD risk in Black and White persons with identical risk factor profiles, which might result in differential treatment decisions based solely on their race.

Summary We recommend the development of raceless CVD risk prediction algorithms that obviate race-associated risk misestimation and racializing treatment practices, and instead incorporate measures of SDoH that mediate race-associated risk differences.

Keywords Race · Risk · Prediction · Cohort studies · Cardiovascular disease

Introduction

Clinical prediction algorithms are useful for decision-making and include a set of variables related to the risk of the disease being diagnosed or forecasted [1, 2]. Management decisions are often guided by such clinical prediction tools, including guiding the choice of diagnostic tests or crossing a specific risk threshold that mandates intervention [1, 2]. Reference equations are also used for categorizing levels of function of organs such as the lungs and the kidneys and incorporate key covariates that influence pulmonary and renal function [3, 4••]. Several risk algorithms and reference equations have included race as a key variable in the estimation of the absolute risk of an incident event or for "normalizing" the level of function of organ systems [3, 4••, 5–7, 8••].

The incorporation of race in clinical algorithms has been debated vigorously and criticized extensively in the recent literature because of the risk of "biologizing" race with a real potential for exacerbating race-related health disparities and underlying racial inequities [8••, 9•, 10–12, 13••, 14, 15•, 16], resulting in calls for race-free prediction algorithms and reference equations. In this context, major risk prediction algorithms for estimating the absolute risk of cardiovascular disease (CVD) [17] and heart failure [5] continue to incorporate race as a variable. Our review focuses attention on the appropriateness or lack thereof of this practice in mainstream cardiovascular medicine.

✉ Ramachandran S. Vasan
vasan@uthscsa.edu

Shreya Rao
raos4@uthscsa.edu

Edwin van den Heuvel
e.r.v.d.heuvel@tue.nl

1 University of Texas School of Public Health and University of Texas Health Sciences Center, 8403 Floyd Curl Drive, Mail Code 7992, San Antonio TX 78229, USA

2 Section of Preventive Medicine and Epidemiology, Department of Medicine, Boston University School of Medicine, Boston, MA, USA

3 Department of Mathematics and Computer Science, Eindhoven University of Technology, Eindhoven, the Netherlands

Estimating the Absolute Risk of Cardiovascular Disease: Guiding Principles and Background

Risk prediction in cardiovascular medicine has been guided by two central tenets. First, several risk factors interact multiplicatively to elevate CVD risk and they can be combined mathematically to create a multivariable risk prediction equation [18]. Second, the intensity of lowering of CVD risk factors is related to the absolute risk of CVD experienced over a time horizon, with a greater intensity of pharmacological intervention in those at greatest risk of experiencing a CVD event [19]. The Framingham Heart Study was a forerunner in this domain with the formulation of risk scores that combined levels (or categories) of standard risk factors such as age, sex, systolic blood pressure, antihypertensive treatment, serum total cholesterol and high-density lipoprotein cholesterol levels, smoking, and diabetes [20–23]. Other risk equations such as the Reynolds risk score expanded the risk factor components to include newer risk markers and parental history of heart disease [24, 25]. These aforementioned risk scores were predominantly developed on cohorts of White men and women.

Given well-established race-associated health disparities [26] and the greater burden of CVD borne by Black people [27], a Risk Assessment Work Group convened by the National Heart, Lung and Blood Institute, the American Heart Association (AHA), and the American College of Cardiology (ACC) led efforts to develop more representative CVD risk prediction equations that were based on several large, racially, and geographically diverse cohort studies [17]. The result of these efforts was sex- and race-specific equations (referred to as pooled cohort equations, PCE) for estimating the 10-year risk of atherosclerotic cardiovascular disease for Black and White adults aged 40–79 years [17]. Subsequently in 2019, the AHA/ACC Task Force on Clinical Practice Guidelines recommended the use of PCE for the primary prevention of CVD [28].

Race as a Component of Pooled Cohort Equations for CVD Risk Prediction: Impact on Absolute Risk Estimates in Black Versus White People with Identical Risk Factors

An initial investigation [29•] explored the transportability of the PCE to a multi-cohort sample that added two additional recent cohort samples to the one used to develop the PCE and deleted data from the older Framingham Study Original cohort. The authors reported that for a third of Black adults, their absolute 10-year CVD risk estimates based on the PCE can vary widely from − 70 to + 250% of comparable estimates for White adults with identical risk factor levels. The investigators revised the PCE (using newer data and statistical methods) to mitigate the misestimation of absolute CVD risk [29•].

In a subsequent investigation [30••], our group assessed the extent to which the use of the PCE might shift absolute CVD risk estimates on opposite sides of the 7.5% threshold that typically guides pharmacological intervention with lipid-lowering agents for Black versus White individuals with identical risk factor profiles. In in silico analyses of over 50,000 risk factor profiles, we evaluated risk factor combinations that might result in divergent PCE-based 10-year CVD risk estimates (on opposite sides of the 7.5% threshold) for Black vs. White individuals (observed by switching the race term in the PCE). We investigated also how often individuals in samples from the Framingham Heart Study (FHS) and a recent cycle of the National Health and Nutrition Examination Survey (NHANES) would yield divergent CVD risk estimates for White individuals versus their same-sex counterparts and vice versa. Approximately, 20% of the in silico risk factor combinations in both men and women yielded divergent CVD risk estimates (on opposite sides of the 7.5% threshold for initiating pharmacological treatment) for Black vs. White individuals, being higher in the former; the converse (CVD risk estimates being higher in White versus Black people) was observed infrequently (~ 1% of the risk factor combinations). The differences in absolute risk (median of 6%) and relative risk (median 2.3) estimates for Black versus White individuals were striking and seemed biologically implausible and were influenced by select risk factors; the presence of higher systolic BP and diabetes magnified the absolute risk differences in Black people relative to their White counterparts [30••]. Furthermore, we observed that the risk misestimation in Black versus White individuals with identical risk factor profiles in our in silico analyses was prevalent in the two community-based samples (FHS and NHANES).

Inappropriateness of Including Race in CVD Risk Prediction Equations: Some Arguments

The foregoing investigations raise some fundamental questions beyond the misestimation of risk in Black versus White people when the PCE are used. First, PCE generate the sequel of differential treatment of individuals with identical risk factor profiles based on their race [29•, 30••], which fundamentally conflicts with *principles of individual fairness* [31].

Second, race is a sociocultural and political construct that reflects societal stratification—it is a marker of historical policies that potentially disadvantaged groups, and created adverse health outcomes due to inequities of resources (material, cultural, social, and political), agency, education and economic opportunities, and intergenerational social mobility [15•, 32, 33••]. From an epidemiological perspective, race serves as a crude group-level latent variable that captures the combination of a myriad of intergenerational and individual exposures and lived experiences, and sociocultural influences (including the effects of structural, institutional, interpersonal, and internalized racism) [34]. Therefore, the use of race as a valid epidemiological classifier in etiological research has been questioned because it provides a formal basis for a focus on *racism* (i.e., social relations) rather than *race* (i.e., innate biologic predisposition) in the interpretation of racial and ethnic "effects" [35•]. It is important to note, though, that because the effects of racism are "embodied," information regarding race offers important insights into the "epidemiology of social inequalities in health" [36], which is distinct from its use as a predictor.

Third, the terms race and ethnicity as defined in the literature often have unclear or unspecified definitions [37, 38]. Such categorization is typically based on self-reports or self-identification, or the implicit assumptions of researchers (based on the skin color of participants), ignores biracial or multiracial people, lumps some people as "others," and is agnostic to the dynamic status of self-perceptions regarding belonging to a specific group. The lack of explicit definitions of race/ethnicity leads one to question the validity of modeling these variables as reliable predictors in deterministic assessments of disease risk [35•,37, 38, 39•]. It is unclear what exactly the race/ethnicity taxonomy measures beyond the color of the skin, facial features, cultural practices, or social experiences [40].

Fourth, although race is somewhat correlated with social class and socioeconomic position, it is a poor global proxy for the concomitant burden of inequities related to the SDoH, such as employment status, social class, education level, food security, neighborhood environmental factors (safety, walkability, etc.), and access to healthcare [41].

A recent investigation [42••] evaluated how SDoH might mediate and explain premature mortality in a national probability-based sample (NHANES). Black participants had a four-fold greater prevalence of six or more unfavorable SDoH (compared to their White counterparts) and such individuals experienced a nearly eight-fold higher mortality risk relative to those with no adverse SDoH [42••]. Notably, adjustment for SDoH completely attenuated the association of the Black race with premature mortality [42••]. Thus, the use of the race term as a proxy for SDoH will divert focus on the truly causal and upstream fundamental factors [43, 44] that mediate disease risk. The unfavorable SDoH can be addressed via multisectoral multidisciplinary team efforts targeting access to education, nutritious food, shelter, health insurance, and transportation to healthcare along with modifications of the neighborhood environments [43, 45]. Additional research is warranted to elucidate if the term race in prediction models represents other factors beyond the SDoH (such as structural determinants and other causes and consequences of racism not easily captured at the individual level). Such research should evaluate how the association of race with outcomes is altered upon adjustment for these additional factors.

Fifth, the use of race in such prediction equations perpetuates the notion that Black people have a greater biological propensity to CVD due to the color of their skin per se, masking the causal role of unfavorable SDoH they face due to discrimination, intergenerational deprivation, and structural racism [45–47]. The use of race for prediction legitimizes its usage as a medically valid classifier, thereby promoting *race-based medicine* [48].

Sixth, the "biologization" of race (or the "racialization" of biology) [36, 40, 49] sustains and enhances the inherent biases in prediction algorithms such as aggregation biases and ecological fallacies [50]. It is noteworthy that the historical cohorts used to develop the PCE often reflect past prejudices and unfair practices/context; sensitive or protected attributes (such as race) may result in the "embodiment of health" [51] and affect the observed treatments and disease outcomes. Thus, the higher CVD rates observed for Black individuals in these datasets may reflect historical differences in embodied experiences and access to health care or may be unduly influenced by the outcomes of a small number of Black people represented in historical observational cohorts including the ones on which the PCE are based. Accordingly, classification algorithms that predict outcomes from these historical datasets tend to replicate and perpetuate these biases and stigmatize select racial groups. Linking predictions based on historical datasets to interventional decisions in contemporary cohorts could negatively impact future healthcare decisions for Black people and thereby extend the health inequities we want to eliminate.

Lastly, there is a distinction between risk equations used for prognostication purposes (distanced from management choices) versus those that are linked directly to decisions about therapeutic choices [52]. The latter context (prediction to guide treatment, as in the case of PCE) warrants the inclusion of true causal factors in the prediction equations rather than imperfect markers that serve as proxies. Such a causal framework for risk prediction [53] will enhance the *transportability* of the algorithms to other populations distinct from the ones used to develop the prediction Eqs. [54]. Moreover, such a modeling strategy facilitates *prediction invariance*, i.e., a prediction model generated from observational data (such as the PCE) will work equally well under

interventional circumstances if its components are causal risk factors [55].

Does Race Capture Genetic Variation that Might Be Linked to Differences in Disease Frequencies?

The completion of the Human Genome Project highlighted the genetic similarities between 99.9% of humans and highlighted that *within-race* genetic variation exceeded *across-race* variation. In our evolutionary history, our species has seldom if ever been totally fixed, isolated, or biologically discrete [56]. Our lineages are the result of complex transcontinental migratory patterns and resultant gene exchanges and flow. So, our genetic variation is mostly geographically continuous and clinal, rather than discrete units with reproductive barriers that can conform to epidemiological categories [56–58]. The genetic diversity of *Homo sapiens* does not map onto socio-politically constructed ethnoracial categories [40, 56, 59, 60].

Continental genetic ancestry has been suggested as a replacement for race in our quest for uncovering genetic determinants of disease risk. However, as noted above and discussed elsewhere [61, 62], this premise is questionable due to global migratory and admixture patterns that underlie human genetic diversity, the likely polygenic origins of complex diseases such as CVD, and the impact of environmental influences that can influence both phenotypic variation and transmissible epigenetic changes [63]. Thus, caution must be exercised in the use of "biogeographical ancestry" in making deterministic assertions about sources of phenotypic variability [64], and risk prediction.

However, it is important to use race, ethnicity, and ancestry to ensure adequate representation of different segments of the population in genetic research to ensure inclusivity and preclude inequities resulting from inadequate representation of historically marginalized groups. A necessary component of inclusive strategies in genetic research is to ensure the optimal standardization and harmonization of definitions of race, ethnicity, and ancestry and methods of collection in genetic studies [39•].

Use of Race to Describe Health Disparities Versus Risk Prediction

Historically marginalized people of color experience a substantial burden of CVD in the USA [65]. Race as a self-identified macro-characteristic often serves as an indicator for describing the health disparities in CVD burden in the community [65]. The use of race may be critical when it is employed as a noncausal descriptor of societal patterns of disease with a view to understanding and elucidating the bases of these patterns, studying health and healthcare outcomes associated with race and ethnic categories, planning healthcare policy and resource allocation, assessing healthcare utilization, and for public health campaigns [13••]. Furthermore, the collection and presentation of data disaggregated by race and ethnicity may be essential to discerning societal patterns of disease [66], as noted in the context of the COVID-19 pandemic [67]. In all such non-etiological research contexts, it is critical to understand and describe how information on race and ethnicity is collected, analyzed, and interpreted to preclude the potential racialization of disease correlates and determinants [35•, 38, 68].

In contrast to its clarified and contextualized use in descriptive and analytical epidemiology to target health disparities, the use of race in etiological research, prediction algorithms, and reference equations is best avoided unless its purpose and effect on achieving more equitable care can be clearly articulated due to the various reasons discussed in the foregoing sections.

Suggested Improvements to the Pooled Cohort Equations: Towards Equity-Centered Race-Conscious CVD Risk Prediction

We and others have presented above several arguments for not using the race term in PCE and other clinical prediction algorithms, clinical decision tools, and reference equations. Scientists, National Foundations, and Academies have endorsed a movement in this direction recently [3, 4••, 7, 8••, 9•, 13••, 16, 69•, 70–74]. How do we address incorporating SDoH (as a replacement for race) in newer raceless enhancements of the PCE, given their overarching role in mediating CVD risk? Among the proposed solutions are incorporating one or more composite measures of social needs and adversity at the individual level, indicators of SDoH (group level), and even an aggregate polysocial risk score [9•, 30••, 75•, 76], exercising caution to avoid aggregation bias and ecological fallacies [13••, 77]. Implementation of such strategies is achievable only with improved capture of social and lived circumstances of patients in research and clinical settings, but successful execution may help to redirect clinical attention from race to causal SDoH factors while potentially improving risk estimation [78•]. Another area for improvement is to develop causal frameworks for counterfactual risk prediction based on longitudinal observations of individuals treated based on their predicted risk (preferably in recent large-sized diverse cohorts) to complement the current strategy based on observational data and historical cohorts [30••].

Conclusion

Racialized health inequities in CVD burden result from the conjoint influence of standard risk factors, social needs at the individual level, and SDoH at the neighborhood and geographic levels. The inclusion of race as a component of CVD risk prediction algorithms (such as the PCE) reifies race as a biologically valid construct and raises concerns about risk misestimation based on race, while concomitantly institutionalizing racialized treatment practices and distracting from the SDoH that drive a substantial proportion of the race-associated differences in CVD burden. Accordingly, current approaches to CVD risk prediction warrant revaluation for their inclusion of race as a predictor variable and exclusive focus on individual-level "biological" risk factors distanced from the social needs of individuals and the SDoH they encounter. We advocate for the development of raceless PCE that incorporate composite measures of adversity at an individual level and SDoH at the neighborhood level, are based on multi-ethnic diverse contemporary cohorts (including data from historically under-represented groups), and use a causal framework supported by counterfactual risk prediction.

Compliance with Ethical Standard

Conflict of Interest The authors declare that they have no conflict of interest.

Human and Animal Rights and Informed Consent This article does not contain any studies with human or animal subjects performed by any of the authors.

References

Papers of particular interest, published recently, have been highlighted as:
- • Of importance
- •• Of major importance

1. Chen L. Overview of clinical prediction models. Ann Transl Med. 2020;8(4):71.
2. Bonnett LJ, Snell KIE, Collins GS, Riley RD. Guide to presenting clinical prediction models for use in clinical settings. BMJ. 2019;365: l737.
3. Bowerman C, Bhakta NR, Brazzale D, Cooper BR, Cooper J, Gochicoa-Rangel L, et al. A race-neutral approach to the interpretation of lung function measurements. Am J Respir Crit Care Med. 2023;207(6):768–74.
4. •• Delgado C, Baweja M, Crews DC, Eneanya ND, Gadegbeku CA, Inker LA, et al. A unifying approach for GFR estimation: recommendations of the NKF-ASN task force on reassessing the inclusion of race in diagnosing kidney disease. Am J Kidney Dis. 2022;79(2):268–88. e1. **This report from the National Kidney Foundation and American Society of Nephrology summarizes their recommendation to replace existing race-based GFR calculations with the non-race-based CKD-Epi eGFR-cystatin-C formula. The authors, writing on behalf of the joint task force, discuss their approach to evaluating various alternative formulas based on considerations of performance as well as how each reflected population diversity and impacted clinical care, surveillance efforts, and research.**
5. Khan SS, Ning H, Shah SJ, Yancy CW, Carnethon M, Berry JD, et al. 10-year risk equations for incident heart failure in the general population. J Am Coll Cardiol. 2019;73(19):2388–97.
6. Medrano D, Matsumoto M, Flores E, Dako F. Understanding race-based medicine and its impact on radiology. J Am Coll Radiol. 2023.
7. Siddique SM, May FP. Race-based clinical recommendations in gastroenterology. Gastroenterology. 2022;162(2):408–14. e2.
8. •• Vyas DA, Eisenstein LG, Jones DS. Hidden in plain sight—reconsidering the use of race correction in clinical algorithms. New England Journal of Medicine: Mass Medical Soc. 2020;874–82. **This perspective presents a description of the application of race corrections in various clinical risk estimations across specialties, and outlines ways in which incorporation of race in each of these scenarios may contribute to differential provision of resources and care by race. The authors caution against indiscriminate incorporation of race into clinical risk formulas and argue that where race is associated with clinical outcomes, investigators should consider what this relationship may represent and how adjustment for race may impact allocation of care and resources.**
9. • Baugh AD, Shiboski S, Hansel NN, Ortega V, Barjaktarevic I, Barr RG, et al. Reconsidering the utility of race-specific lung function prediction equations. Am J Respir Crit Care Med 2022;205(7):819–829. **This study examines the use of race-specific equations for estimating lung function in the care and management of patients with COPD. Authors report on results using both race-specific and universally applied formulas and find that use of a race-specific formula resulted in underestimates of disease severity in Black individuals as assessed by symptoms, 6-min walk distance, and radiographic findings. They additionally report that adjustment for SDOH information attenuated the previously reported relationships between race and lung function, suggesting that these differences may be at least partly attributed to social factors.**
10. Elmaleh-Sachs A, Balte P, Oelsner EC, Allen NB, Baugh A, Bertoni AG, et al. Race/ethnicity, spirometry reference equations, and prediction of incident clinical events: the Multi-Ethnic Study of Atherosclerosis (MESA) Lung Study. Am J Respir Crit Care Med. 2022;205(6):700–10.
11. Eneanya ND, Yang W, Reese PP. Reconsidering the consequences of using race to estimate kidney function. JAMA. 2019;322(2):113–4.
12. Kelso JM. How to deconstruct "race" and spirometry. J Allergy Clin Immunol Pract. 2022;10(9):2488–9.
13. •• Mohottige D, Boulware LE, Ford CL, Jones C, Norris KC. Use of race in kidney research and medicine: concepts, principles, and practice. Clin J Am Soc Nephrol. 2022;17(2):314–22. **In this review, authors define the terms race essentialism, racism, race, and genetic ancestry as they impact disparities in health outcomes and the provision of health services. The authors emphasize that genetic bases of race have no scientific validity and implications that racial groups are defined by unique biological traits reflect antiquated notions of race essentialism. They advocate for a race-conscious approach that acknowledges the social importance of race as a reflection of the effects of racism and promotes a health**

equity–minded approach to studying and using race in clinical care.

14. Norris KC, Eneanya ND, Boulware LE. Removal of race from estimates of kidney function: first, do no harm. JAMA. 2021;325(2):135–7.
15.• Powe NR. Black kidney function matters: use or misuse of race? JAMA. 2020;324(8):737–8. **This editorial summarizes the challenges surrounding both inclusion and simple elimination of race in the estimation of GFR. The author acknowledges the role for "raceless" markers, such as muscle mass in the context of GFR estimation, in providing more precise estimations of population heterogeneity in kidney filtration than race alone.**
16. Vyas DA, James A, Kormos W, Essien UR. Revising the atherosclerotic cardiovascular disease calculator without race. The Lancet Digital health. 2022;4(1):e4–5.
17. Goff DC Jr, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, 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(25_suppl_2):S49-S73.
18. Gordon T, Kannel WB. Multiple risk functions for predicting coronary heart disease: the concept, accuracy, and application. Am Heart J. 1982;103(6):1031–9.
19. Executive summary of the third report of the National Cholesterol Education Program (NCEP) expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel III). JAMA. 2001;285(19):2486–97.
20. Anderson KM, Wilson P, Odell PM, Kannel WB. An updated coronary risk profile. A statement for health professionals. Circulation. 1991;83(1):356–62.
21. Committee NCEPC. National Cholesterol Education Program. Second report of the expert panel on detection, evaluation, and treatment of high blood cholesterol in adults (adult treatment panel II). Circulation. 1994;89(3):1333–445.
22. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation. 1998;97(18):1837–47.
23. D'Agostino RB Sr, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation. 2008;117(6):743–53.
24. Ridker PM, Buring JE, Rifai N, Cook NR. Development and validation of improved algorithms for the assessment of global cardiovascular risk in women: the Reynolds Risk Score. JAMA. 2007;297(6):611–9.
25. Ridker PM, Paynter NP, Rifai N, Gaziano JM, Cook NR. C-reactive protein and parental history improve global cardiovascular risk prediction: the Reynolds Risk Score for men. Circulation. 2008;118(22):2243–51.
26. Velarde G, Bravo-Jaimes K, Brandt EJ, Wang D, Douglass P, Castellanos LR, et al. Locking the revolving door: racial disparities in cardiovascular disease. J Am Heart Assoc. 2023;12(8):e025271.
27. Carnethon MR, Pu J, Howard G, Albert MA, Anderson CA, Bertoni AG, et al. Cardiovascular health in African Americans: a scientific statement from the American Heart Association. Circulation. 2017;136(21):e393–423.
28. Arnett DK, Blumenthal RS, Albert MA, Buroker AB, Goldberger ZD, Hahn EJ, 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(11):e596–646.
29.• 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(1):20–29. **This study was critical in highlighting the shortcomings of the traditional PCE in risk estimation, particularly among under-represented minority groups. The authors present a revised PCE using more contemporary cohort data and statistical methods intended to avoid overfitting, resulting in a reduction, but not elimination, in risk overestimation across populations and in extreme and implausible values in subpopulations.**
30.•• Vasan RS, van den Heuvel E. 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. Lancet Digit Health. 2022;4(1):e55–e63. **In this in silico analysis of the Pooled Cohort Equations, authors demonstrate substantial and implausible differences in cardiovascular disease risk estimation between White and Black individuals using similar hypothetical risk factor profiles. Overall, they find that the PCE more commonly overestimated risk among Black men and women, frequently resulting in divergent clinical recommendations by categorizing patients on either side of the statin treatment threshold.**
31. Fleisher W, editor. What's fair about individual fairness? Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society; 2021.
32. Braun L, Grisson R. Race, Lung function, and the historical context of prediction equations. JAMA Network Open. 2023;6(6):e2316128-e.
33.•• Wright JL, Davis WS, Joseph MM, Ellison AM, Heard-Garris NJ, Johnson TL et al. Eliminating race-based medicine. Pediatrics. 2022;150(1):e2022057998. **In this review, authors describe examples of race-based practices across medical specialties on behalf of the American Association of Pediatrics. In discussing the role of race in risk estimation using the PCE, authors call for revision of the race-based PCE to address previously identified misestimation of cardiovascular disease risk on the basis of race alone.**
34. Jones CP. Levels of racism: a theoretic framework and a gardener's tale. Am J Public Health. 2000;90(8):1212–5.
35.• Kaufman JS, Cooper RS. Commentary: considerations for use of racial/ethnic classification in etiologic research. Am J Epidemiol. 2021;154(4):291–298. **In this commentary, authors review methodological considerations for the use of race and ethnicity in biological research. They emphasize the importance of distinguishing between studies of race, which imply biological causation in race-related differences, and those of racism, that draw attention to causal social factors and relationships that drive race-related differences.**
36. Krieger N. Refiguring, "race": epidemiology, racialized biology, and biological expressions of race relations. Int J Health Serv. 2000;30(1):211–6.
37. Bhopal R. Glossary of terms relating to ethnicity and race: for reflection and debate. J Epidemiol Community Health. 2004;58(6):441–5.
38. Lin SS, Kelsey JL. Use of race and ethnicity in epidemiologic research: concepts, methodological issues, and suggestions for research. Epidemiol Rev. 2000;22(2):187–202.
39.• Popejoy AB, Crooks KR, Fullerton SM, Hindorff LA, Hooker GW, Koenig BA, et al. Clinical genetics lacks standard definitions and protocols for the collection and use of diversity measures. Am J Hum Genet. 2020;107(1):72–82. **This survey of clinical genetics professionals demonstrates heterogeneity in beliefs around the information captured in measures of race, ethnicity, and genetic ancestry, and in the appropriate use of such data in clinical genetics practice. The authors highlight the need for standardization and evidence-based**

approaches to interpreting and utilizing this data given the potential implications for precision-based medical care.

40. Malinowska JK, Żuradzki T. Towards the multileveled and processual conceptualisation of racialised individuals in biomedical research. Synthese. 2022;201(1):11.
41. Solar O, Irwin A. A conceptual framework for action on the social determinants of health. Geneva, Switzerland: WHO; 2010. [September 22, 2016]. (Social determinants of health discussion paper 2 (policy and practice)).
42.•• Bundy JD, Mills KT, He H, LaVeist TA, Ferdinand KC, Chen J, et al. Social determinants of health and premature death among adults in the USA from 1999 to 2018: a national cohort study. Lancet Public Health. 2023;8(6):e422–e431. **This study examined the mediating role of SDOH in contributing to racial disparities in premature all-cause mortality in a community-based cohort sampled from the National Health and Nutrition Examination Survey. The authors demonstrate that adjustment for eight self-reported SDOH variables accounted for all of the difference in mortality observed by race, and that adverse SDOH traits had a multiplicative effect on risk, such that each additionally SDOH risk factor reported by a participant resulted in a significantly higher risk of premature death. The findings are novel in demonstrating how incorporation of causative socioeconomic factors in observational data may explain differences observed by race and how adverse SDOH factors may interact in driving higher risk of disease.**
43. Hacker K, Auerbach J, Ikeda R, Philip C, Houry D. Social determinants of health-an approach taken at CDC. Journal of public health management and practice : JPHMP. 2022;28(6):589–94.
44. Braveman P, Gottlieb L. The social determinants of health: it's time to consider the causes of the causes. Public Health Rep. 2014;129(1_suppl2):19–31.
45. Brandt EJ. Social determinants of racial health inequities. The Lancet Public Health. 2023;8(6):e396–7.
46. Braveman P, Parker Dominguez T. Abandon "race." Focus on racism. Front Public Health. 2021;9:689462.
47. Duster T. Race and reification in science. Am Assoc Adv Sci. 2005;1050–1.
48. Cerdeña JP, Plaisime MV, Tsai J. From race-based to race-conscious medicine: how anti-racist uprisings call us to act. The Lancet. 2020;396(10257):1125–8.
49. Peressini AF. Against the philosophical project of "biologizing" race. Metaphilosophy. 2021;52(5):593–615.
50. Robinson WS. Ecological correlations and the behavior of individuals. Int J Epidemiol. 2009;38(2):337–41.
51. Krieger N. Embodiment: a conceptual glossary for epidemiology. J Epidemiol Community Health. 2005;59(5):350–5.
52. Prosperi M, Guo Y, Sperrin M, Koopman JS, Min JS, He X, et al. Causal inference and counterfactual prediction in machine learning for actionable healthcare. Nature Machine Intelligence. 2020;2(7):369–75.
53. Schooling CM, Jones HE. Clarifying questions about "risk factors": predictors versus explanation. Emerg Themes Epidemiol. 2018;15(1):10.
54. 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(1):179.
55. Peters J, Bühlmann P, Meinshausen N. Causal inference by using invariant prediction: identification and confidence intervals. J R Stat Soc Ser B (Statistical Methodology). 2016:947–1012.
56. Fuentes A, Ackermann RR, Athreya S, Bolnick D, Lasisi T, Lee SH, et al. AAPA statement on race and racism. Am J Phys Anthropol. 2019;169(3):400–2.
57. Sankar P, Cho MK. Toward a new vocabulary of human genetic variation. Am Assoc Adv Sci. 2002;1337–8.
58. Barbujani G. Human races: classifying people vs understanding diversity. Curr Genomics. 2005;6(4):215–26.
59. Collins FS. What we do and don't know about 'race', 'ethnicity', genetics and health at the dawn of the genome era. Nat Genet. 2004;36(11):S13–5.
60. Yudell M, Roberts D, DeSalle R, Tishkoff S. Taking race out of human genetics. Science. 2016;351(6273):564–5.
61. Cerdeña JP, Grubbs V, Non AL. Genomic supremacy: the harm of conflating genetic ancestry and race. Hum Genomics. 2022;16(1):18.
62. Duster T. A post-genomic surprise. The molecular reinscription of race in science, law and medicine. Br J Sociol. 2015;66(1):1–27.
63. Leatherman T, Goodman A. Building on the biocultural syntheses: 20 years and still expanding. American journal of human biology : the official journal of the Human Biology Council. 2020;32(4):e23360.
64. Lee SS-J, Mountain J, Koenig B, Altman R, Brown M, Camarillo A, et al. The ethics of characterizing difference: guiding principles on using racial categories in human genetics. Genome Biol. 2008;9(7):404.
65. Tsao CW, Aday AW, Almarzooq ZI, Anderson CA, Arora P, Avery CL, et al. Heart disease and stroke statistics—2023 update: a report from the American Heart Association. Circulation. 2023;147(8):e93–621.
66. Kauh TJ, Read JNG, Scheitler AJ. The critical role of racial/ethnic data disaggregation for health equity. Popul Res Policy Rev. 2021;40(1):1–7.
67. Association AM. Top physician orgs urge COVID-19 mortality data by race, ethnicity. 2020.
68. Kaufman JS. Epidemiologic analysis of racial/ethnic disparities: some fundamental issues and a cautionary example. Soc Sci Med (1982). 2008;66(8):1659–69.
69.• Wright JL, Freed GL, Hendricks-Muñoz KD, Jarvis JN, MaldonadoYA, Raphael JL, et al. Achieving equity through science and integrity: dismantling race-based medicine. Pediatr Res. 2022;91(7):1641–1644. **This editorial examines the historical roots of race-based medicine and presents a case for moving away from biologized notions of race in favor of approaches that incorporate macro- and individual-level SDOH.**
70. Boulware LE, Mohottige D, Maciejewski ML. Race-free estimation of kidney function: clearing the path toward kidney health equity. JAMA. 2022;327(23):2289–91.
71. Essien UR, Corbie G. Getting under the skin: race-based guidelines and the pursuit of pharmacoequity. J Gen Intern Med. 2022;37(16):4035–6.
72. Yearby R. Race based medicine, colorblind disease: how racism in medicine harms us all. Am J Bioeth. 2021;21(2):19–27.
73. Rondini AC, Kowalsky RH. "First do no harm": clinical practice guidelines, mesolevel structural racism, and medicine's epistemological reckoning. Soc Sci Med. 2021;279:113968.
74. Gill JS, Kelly B, Tonelli M. Time to abolish metrics that sustain systemic racism in kidney allocation. JAMA. 2023;329(11):879–80.
75.• Yancy CW, Khan SS. Replacing race with social determinants of health in risk prediction—getting it right. JAMA Cardiology. 2022;7(8):856. **This editorial highlights the fallacies inherent in the biologization of race and endorses an approach to risk estimation incorporating SDOH information for which race is a likely proxy in the medical literature. The authors highlight one such tool, developed by Segar et al., in which incorporation of individual-, neighborhood-, and hospital-level SDOH data using a machine learning approach improved risk prediction following HF hospitalization.**

76. Figueroa JF, Frakt AB, Jha AK. Addressing social determinants of health: time for a polysocial risk score. JAMA. 2020;323(16):1553–4.
77. Subramanian SV, Jones K, Kaddour A, Krieger N. Revisiting Robinson: the perils of individualistic and ecologic fallacy. Int J Epidemiol. 2009;38(2):342–60.
78.• Segar MW, Hall JL, Jhund PS, Powell-Wiley TM, Morris AA, Kao D, et al. Machine learning-based models incorporating social determinants of health vs traditional models for predicting in-hospital mortality in patients with heart failure. JAMA Cardiol. 2022;7(8):844–854. **This study presents a novel approach to estimating in-hospital heart failure mortality using a race-agnostic tool in concert with the American Heart Association Get with the Guidelines-Heart Failure registry. The authors demonstrate improved risk prediction using a machine learning algorithm with incorporation of patient- and neighborhood-level SDOH data.**