VIEWPOINT

Recalibrating the Use of Race in Medical Research

John P. A. Ioannidis, MD, DSc
Department of Medicine, Stanford University School of Medicine, Stanford, California; and Department of Epidemiology and Population Health, Stanford University School of Medicine, Stanford, California.

Neil R. Powe, MD, MPH, MBA
Zuckerberg San Francisco General Hospital and University of California–San Francisco.

Clyde Yancy, MD, MSc
Division of Cardiology, Department of Internal Medicine, Northwestern University, Feinberg School of Medicine, Chicago, Illinois.

Viewpoint pages 625 and 627

Multimedia

Corresponding Author: John P. A. Ioannidis, MD, DSc, Stanford University School of Medicine, 1265 Welch Rd, Medical School Office Bldg, Room X306, Stanford, CA 94305 (jioannid@stanford.edu).

Race was originally introduced in US medical curricula in 1790 by Benjamin Rush, who asserted that blackness was a particular kind of leprosy. In 1857 Josh Nott characterized slaves as a biologically appropriate phenotype for hard labor under trying conditions. In the 1870s, the Jim Crow era of race exclusion from most societal venues reinforced medical segregation. This sordid history, although painful to recite, is the underpinnings of race in medicine, including its use in medical research.

Race as a variable in medical research has long been a contentious issue.[1] It is widely accepted that race is an indistinct construct that is not always measured accurately and standardized. In 1999, the Human Genome Project emphasized race as nonbiological with no basis in the genetic code. What, then, does race define?

Race is a poor surrogate of social constructs and even more so, if not abjectly, of biology. Differences observed in research studies between "races" may result from the multifarious consequences of long-entrenched and continuously transformed racism. As the crisis of coronavirus disease 2019 has revealed once again, long-standing effects of racism have tremendous effects on the propagation of inequalities and injustice at all levels, including health and health care. Racism, tragically, remains a chronic and acute problem of modern societies, and the use of race in medical research and practice is now being brandished as a surrogate for racism. Eradicating racism should be a moral imperative in medicine.

> If for no other reason than the further exposition of health inequities and systemic racism, the use of race should for now persist in medical research. But the imperfectness of race as a tool is problematic.

However, is any progress addressing inequities possible if race as a measure is banned? Is there still some room for using race variables? How much would be lost if these variables were eliminated? Is there a better tool in research and policy efforts? Are there some situations in which race variables remain valuable? What strategy would generate research that diminishes rather than increases inequalities and injustice? The time has come to recalibrate the use of race in medical research.

The call to entirely abandon race from medical research endeavors began several decades ago but is a simplistic solution to a complex set of concerns.[2,3] Dislodgement of race from research may hide still-evident and often egregious episodes of health disparities. If for no other reason than the further exposition of health inequities and systemic racism, the use of race should for now persist in medical research. But the imperfectness of race as a tool is problematic.

One school of thought asserts that because race (and ethnicity) is so weakly measured and even more poorly analyzed and reported, efforts should focus on trying to strengthen measurement, analysis, and reporting. A series of initiatives, including self-identification, especially in clinical trials and registries and in specifications of requirements for publicly funded research, ensured that more attention would be given toward obtaining more data on racial minority populations. However, empirical evaluations show that race information can be fragmented, inconsistent, and eventually not very usable.

The medical literature that uses or discusses race is vast, but is it really informative? On December 21, 2020, a search of PubMed with "race OR ethnicity" yielded 518 842 items, whereas one with focused terms such as "African American" and "Hispanic OR Latino" yielded 44 674 and 61 933 items, respectively. However, a recent evaluation[4] of a random sample of 1000 Cochrane systematic reviews on various medical interventions showed that only 14 (1.4%) had proposed to perform race- or ethnicity-based subgroup analyses for treatment effects. Only 1 of those 14 analyses was completed but yielded noninformative results.[4] Despite the poor performance of race as a measure, numerous passionate, burgeoning health professionals, many of whom are underrepresented in medicine, have been attracted to biomedical research, lured by life experiences to study with enthusiasm the interrelation of race and ethnicity with social and biological factors. Their work should go forward.

A second school of thought argues that race is a painful historical relic and lost cause. With this approach, race as a measure should be abandoned, and efforts should be diverted toward finding variables that are more robust and informative, both for the biological constructs (eg, genetic ancestry) and the sociologic ones (eg, discrimination, deprivation, socioeconomic status) for which race has failed to provide useful, reproducible insights. Does scientific theory support this approach?

On the frontiers of biology, the rapid advent of genetics has transformed the concept of ancestry. A spectrum of genetic granularity through whole-genome sequencing makes the surrogate of traditional races potentially obsolete. However, genetics, despite its tremendous accuracy of measurement and massive information, has been sluggish in making much progress in yielding useful medical tools for everyday practice and for improving patient and population outcomes that

matter to many. If anything, genetics may be contributing to worsening inequalities, especially when most genetic architecture databases overrepresent people of European ancestry (88% of genome-wide data had European ancestry as of 2018),[5] when genomic tools are too expensive to use for race-based research, and when both biological scientists and social scientists default to White as a reference standard to which others are normalized.[6]

Race may well be a surrogate, albeit imperfect, for sociologic constructs. However, the most important sociologic variables (eg, social determinants of health) and, in particular, differential opportunities (eg, good access to and quality of care) fail to associate with sufficient precision when race is used as the placeholder. A long list of variables has emerged that try to capture socioeconomic aspects, access to care, health insurance, discrimination, deprivation, geography and place, perceived identity, opportunities, social interactions, financial mobility, health behaviors, and more. Although many of these variables probably come closer to causal relationships than race, they too are still largely nonstandardized, are often crudely measured, and unfortunately do not fully explain differences by race. Limited translational potential and transferability ensue.

Perhaps it is possible to find a middle ground between these 2 schools of thought, improvement vs elimination, in navigating this conundrum. The research corpus can be separated into 2 components: past research investigations in which race has been incorporated in medical textbooks, clinical algorithms, guidelines, recommendations, and other evidence that may or may not be applied in practice; and future research investigations.

For past investigations, a large amount of research involving race variables has been, in hindsight, pedestrian and arguably lies among the greater waste of spurious, nonusable biomedical evidence. However, there are examples for which race variables have become part of the norm of accepted medical knowledge and practice. This applies to both therapeutics (incorporation of race to identify clinically meaningful treatment effect modification for various interventions, as in hypertension or heart failure)[7-9] and other clinical tools (incorporation of race to improve diagnosis or prognosis in, for example, calculation of kidney function or pulmonary function).[10] Expert specialty medical societies and methodologists should jointly systematically reexamine evidence involving race that is already accepted as core knowledge. For some applications, race may continue to be the best variable to capture the influence on health; quick dismissal or normalization of values to the majority group may worsen outcomes, especially for the most disadvantaged populations. For other situations, it may be realized that these race variables have become obsolete: what they were supposed to presage when they were first proposed may no longer be relevant in the current social and biological science landscape. Alternatively, perhaps some race variables continue to offer incremental, useful information, including the further elucidation of health disparities. However, other, better variables should be developed to replace race per se. Such replacements need to proceed with rigorous validation practices, ensuring the generalizability of the results and solidifying that whatever changes are made will help reduce, rather than exacerbate, existing inequalities.

For future investigations, it is important to think carefully about the fundamental question. Why should race variables be used, if at all? Consider 4 steps: (1) execute a systematic review of prior research because race may have been exhausted as a tool and is futile to study again, or may offer insight for how a new study may best leverage past work, or create novel hypotheses; (2) if race measurements are deemed appropriate, carefully consider collateral, explanatory biological and sociologic variables appropriate to include in the same investigation, and how standardization, accuracy, and relevance may be enhanced in explaining race-based signals; (3) in any comparative analyses, investigators should consider whether White race should be the reference standard because normative values are reasonable, but normal designations that characterize some humans as aberrant are problematic; and (4) carefully consider the potency of any race-related research and gauge a holistic portfolio of clinical and social consequences, including the amelioration or aggravation of existing inequalities.

In a volatile social landscape, it may not be possible to determine exactly how race-specific research efforts may lead to a better, more fair world. At a minimum, however, medical research should not aggravate already embedded gaps between the privileged and the disadvantaged. Just as the lens of science was used to establish a flawed premise of biological race-based differences, so should science now focus on illuminating that which is represented by race and become a trailblazer toward better health equity.

ARTICLE INFORMATION

Published Online: January 25, 2021. doi:10.1001/jama.2021.0003

Conflict of Interest Disclosures: Dr Yancy reports spousal employment at Abbott Labs Inc. No other disclosures were reported.

REFERENCES

1. 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. doi:10.1093/oxfordjournals.epirev.a018032

2. Fullilove MT. Comment: abandoning "race" as a variable in public health research—an idea whose time has come. *Am J Public Health*. 1998;88(9):1297-1298. doi:10.2105/AJPH.88.9.1297

3. Bhopal R, Donaldson L. White, European, Western, Caucasian, or what? inappropriate labeling in research on race, ethnicity, and health. *Am J Public Health*. 1998;88(9):1303-1307. doi:10.2105/AJPH.88.9.1303

4. Liu P, Ross JS, Ioannidis JP, Dhruva SS, Vasiliou V, Wallach JD. Prevalence and significance of race and ethnicity subgroup analyses in Cochrane intervention reviews. *Clin Trials*. 2020;17(2):231-234. doi:10.1177/1740774519887148

5. Mills MC, Rahal C. A scientometric review of genome-wide association studies. *Commun Biol*. 2019;2:9. doi:10.1038/s42003-018-0261-x

6. Martin AR, Kanai M, Kamatani Y, Okada Y, Neale BM, Daly MJ. Clinical use of current polygenic risk scores may exacerbate health disparities. *Nat Genet*. 2019;51(4):584-591. doi:10.1038/s41588-019-0379-x

7. Bloche MG. Race-based therapeutics. *N Engl J Med*. 2004;351(20):2035-2037. doi:10.1056/NEJMp048271

8. Taylor AL, Wright JT Jr. Should ethnicity serve as the basis for clinical trial design? importance of race/ethnicity in clinical trials: lessons from the African-American Heart Failure Trial (A-HeFT), the African-American Study of Kidney Disease and Hypertension (AASK), and the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). *Circulation*. 2005;112(23):3654-3660. doi:10.1161/CIRCULATIONAHA.105.540443

9. Ramamoorthy A, Pacanowski MA, Bull J, Zhang L. Racial/ethnic differences in drug disposition and response: review of recently approved drugs. *Clin Pharmacol Ther*. 2015;97(3):263-273. doi:10.1002/cpt.61

10. Diao JA, Wu GJ, Taylor HA, et al. Clinical implications of removing race from estimates of kidney function. *JAMA*. 2020;325(2):184-186. doi:10.1001/jama.2020.22124