

## Race and Ethnicity

### A Part of the Equation for Personalized Clinical Decision Making?

Jessica K. Paulus, ScD; David M. Kent, MD, MS

Two 48-year-old men, Jake and Earl, walk into a physician's office. They both have a total cholesterol of 200 mg/dL, a high-density lipoprotein level of 40 mg/dL, systolic blood pressure of 130 mm (despite antihypertension treatment), and no history of diabetes mellitus. Their physician applies the same primary prevention guidelines, and—despite their identical cardiac risk factors—recommends an aspirin and a statin for Earl but not Jake. These patients might be surprised to find their discrepant treatment derives from the fact that Jake is white and Earl is African American. Thus, according to the pooled cohort equations, the calculated 10-year risk of atherosclerotic cardiovascular disease is 4.3% for Jake (below the 7.5% American College of Cardiology/American Heart Association guideline threshold<sup>1</sup>) and 8.7% for Earl. The inclusion of race in the pooled cohort equations as a predictor of 10-year atherosclerotic cardiovascular disease risk has generated this race-based treatment difference.

In this commentary, we consider the use of race/ethnicity in clinical prediction models and discuss the similarities (and differences) between racial profiling in healthcare and in other settings, such as law enforcement and the insurance market.

#### Race Effects are Sometimes Included in Clinical Predictive Models But Usually Not: A Lack of Consensus Regarding Best Practice

Since the 1993 National Institutes of Health Revitalization Act mandated the inclusion of minorities in federally-funded clinical research, and specified that all research proposals describe study populations' racial and ethnic composition, race/ethnicity is a widely available variable for study. Disparities in treatments and differences in outcome risk and treatment effects are routinely examined. Risk differences by race/ethnicity have been documented in most major cardiovascular conditions, often of a magnitude comparable with that of clinical cardiovascular risk factors. These studies have generated controversy about potential underlying causal mechanisms, and how to approach race as an etiologic variable.<sup>2-4</sup>

Generally, variables encoding race and ethnicity are assumed to summarize a complex combination of factors related to social class, culture, biology, and the health impact of racism. Regardless of the mechanisms, racial/ethnic designation is frequently a strong predictor of outcomes, and

so—given its wide availability and the differences in associated outcomes—it seems a natural candidate variable in risk prediction. Indeed, calibrating models to different racial/ethnic groups may be thought to represent good practice in the broad dissemination of prediction models. Although its predecessor, the Framingham risk score, was known to perform poorly in some racial/ethnic groups, the pooled cohort equations underwent rigorous validation in multiethnic populations, and are in fact race-stratified, with unique equations for African Americans and whites specifically developed given known differences in event rates and possible differences in coefficients.<sup>1</sup> Race is also included in popular risk models that help determine the need for osteoporotic screening, prostate cancer biopsy, breast cancer screening (<http://www.cancer.gov/bcrisktool>), and other common decisions.

Nevertheless, despite these notable exceptions, we have found evidence of a deep reluctance to include race in CPMs. In the Tufts Predictive Analytics and Comparative Effectiveness CPM database (<http://pace.tuftsmedicalcenter.org/cpm>), which provides a comprehensive summary of reported predictive models in the cardiovascular literature, only 3% (23/854) of models include a coefficient for race or ethnicity, or present race/ethnicity-stratified equations. Even in models for conditions where prognostic effects of race have been well-documented, race/ethnicity is rarely included—including examples in which authors explicitly acknowledge excluding race from the model, despite a significant association with outcomes because of concern about the implications of using patient race for clinical decision making.

#### Rationale for Excluding Race Effects From Clinical Prediction Models: Similarities to Racial Profiling in Other Contexts

The reluctance to include race in prediction models is easy to understand. Including race in prediction models to support race-sensitive decision making recapitulates some major elements of racial profiling in other clearly objectionable contexts, such as law enforcement. In both settings, by using race predictively, individuals are judged differentially based on differences in averages in the racial/ethnic groups to which they belong, supporting decision making that (literally) discriminates on the basis of race, leading to apparently unequal

From the Predictive Analytics and Comparative Effectiveness Center, Institute for Clinical Research and Health Policy Studies, Tufts Medical Center/Tufts University School of Medicine, Boston, MA.

Correspondence to Jessica K. Paulus, ScD, Predictive Analytics and Comparative Effectiveness Center, Institute of Clinical Research and Health Policy Studies, Tufts Medical Center, 800 Washington St #63, Boston, MA 02111. E-mail [j paulus@tuftsmedicalcenter.org](mailto:j paulus@tuftsmedicalcenter.org) (*Circ Cardiovasc Qual Outcomes*. 2017;10:e003823. DOI: 10.1161/CIRCOUTCOMES.117.003823.)

© 2017 American Heart Association, Inc.

*Circ Cardiovasc Qual Outcomes* is available at <http://circoutcomes.ahajournals.org>

DOI: 10.1161/CIRCOUTCOMES.117.003823

treatment. Further, using race predictively particularly in the medical setting may seem to reify race from a social construct to a meaningful biological variable and legitimize discredited notions of racial essentialism.<sup>5</sup>

Although omitting factors that can contribute to prediction can lead to model misspecification, miscalibration, and harmful decision making, improving predictive accuracy is clearly insufficient on its own to justify using race as a predictive tool. In actuarial practice, for example, it is generally proscribed to use race in estimating life expectancy (or other risks) when setting insurance premiums, despite known racial differences in survival curves. It is simply unfair (and illegal) to be disadvantaged on the basis of a suspect characteristic—such as race or national origin—that cannot itself be a direct cause of some undesired predicted outcome, even if and when this characteristic might be a convenient and salient (although unreliable) proxy for other important, and perhaps unknown, causes. Thus, even if discriminatory policies lead to better prognostication, improve efficiency, and optimize an important social good, and even if these policies do not arise from bigotry or racist theories or beliefs, the basic ethical issue of fairness persists. This is true broadly across all spheres of human commerce, including law enforcement, employment, finance, and housing.

At least some legal scholars have opined that the same laws that make racial profiling illegal in other contexts apply also to healthcare decision making; the 42 US code § 1981 bars racial discrimination in all contracts (including the patient–provider relationship) and can be violated by race-based decision making when the clinician denies the patient the opportunity for meaningful choice.<sup>6</sup> Meaningful choice in this context presumably involves informing the patient when a decision or recommendation is made that might change if the patient was of a different racial/ethnic group. However, in our experience, tools are rarely applied in this way. Indeed, in the era of predictive analytics, prediction models may operate behind the scenes—automatically extracting information from the patient's electronic health record—not only to inform therapeutic decision making, but also to target care management interventions to high-risk populations.

### Important Distinctions Between Racial Profiling in Clinical Prediction and Other Domains

Nevertheless, it is important to be mindful of key differences between racial profiling in healthcare practice and in law enforcement and other areas, despite the compelling similarities. For example, in the context of insurance and law enforcement, profiled individuals would be unambiguously punished (or rewarded) based on the average risk in the racial group to which they belong. Individuals estimated to be at higher risk because of their race/ethnicity are only victims (and never beneficiaries) of profiling. This is true regardless of the theoretical effectiveness of racial profiling because any hypothetical benefits of predictive policing or actuarial risk disaggregation accrue to the collective society and not to the high-risk individual being profiled. In contrast, in the medical context, there is generally no conflict between efficiency concerns and fairness concerns because

for a risk-sensitive decision, there are—by definition—harms for both over- and undertreatment (and, therefore, harms to both over- and underestimation of risk). Therefore, when accuracy can be improved by the inclusion of race, decision making is improved (on average) for patients across racial groups. In the medical context, more accurate prognostication should always permit better alignment with the patient's own goals to receive the best, most individualized care (with the potential exception of when predictions are used to prioritize rationed healthcare services). When prognostic differences exist across different racial groups, using race-blind predictions harms all groups, but especially disadvantages minority groups, because predictions will more closely reflect those in the majority.

The potential import of including race in risk prediction is underscored by its effect on disparities research where—for example—equal treatment rates in white and nonwhite patients adjusting for all risk factors might obscure an actual disparity if, for example, nonwhite race itself is associated with a higher risk of the clinical outcome justifying more intensive therapy (such as with Earl compared with Jake).<sup>7</sup> Similarly, a 2014 National Quality Forum Expert Panel recommended that outcome-based, and some process-based, performance measures be adjusted for socioeconomic and sociodemographic factors (including race) under certain conditions. Not adjusting for markers of social disadvantage can exacerbate existing disparities by penalizing practices caring for the underserved or even by creating incentives for adverse selection of patients because it is easier to achieve high-quality scores among more socially privileged patients.<sup>8</sup>

### Caveats

Of course, race is (logically) just a surrogate for other causes and not a cause itself (except as mediated through racism or other indirect effects).<sup>2</sup> For this reason and others, decomposing the effects of race more precisely into its components (eg, genetic, socioeconomic, cultural factors, etc) has been recommended.<sup>9</sup> Although this may be useful for exploring causal hypotheses, socioeconomic and genetic factors are often unavailable in the relevant databases and difficult to ascertain in routine care when applying prediction models (genetic factors may also be unknown). Further, given the multidimensionality of race, we anticipate that including these other factors will likely leave a substantial residual predictive effect of race because of unmeasured variables and complex and poorly understood interactions.

To be sure, given the extremely compelling surface similarities between using race predictively in medical care and in the other contexts, it is legitimate to worry about collateral harms for society by giving cover for other forms of racial profiling, even if race-sensitive decision making helps optimize decisions for each patient. Additionally, notwithstanding recent changes in federal standards for how race/ethnicity data are collected and presented, race is a poorly defined concept, making it a nonideal predictor. Although it is generally defined by self-report in research databases (and predictive models), we suspect that clinicians using models tend to substitute their own impression, which might result in misclassification. Further, because

world populations are increasingly multiethnic, it is progressively more challenging to assign a single ethnicity to an individual. Although biogeographical ancestry (determined from a panel of ancestral informative genetic markers) may have some definitional advantages compared with self-identified race in terms of being more objectively (and continuously) defined, it is not widely available and may misleadingly emphasize genetic or biological underpinnings of race effects. More modern classification concepts (with less shameful histories) may ultimately replace the concept of race in medical research,<sup>9</sup> and do similar predictive work.

### Summary and Conclusion

There is clearly a lack of consensus in how to cope with race as a predictor in CPMs. More discussion, involving a wide variety of stakeholders, is needed to develop consensus about best practice. In this commentary, we focused on the narrow but important issue of fairness (ie, the harms caused to one group versus other groups), which would militate against the use of race predictively in other contexts but not necessarily in medical decision making. We point out that some of the reluctance about using race predictively in CPMs comes from overgeneralizing from these other contexts, which may substantially misrepresent the concerns in the medical decision making context.

Although racial equity might be taken to imply identical (race-neutral) care, the careful modeling on multiethnic populations of the pooled cohort equations points to a different paradigm of equity that accommodates the manifold and important influence of race/ethnicity in health and disease. This influence has been broadly recognized and is the impetus behind both new and longstanding regulations to include diverse populations in clinical research, rather than relying on the assumption that results from racially and ethnically homogeneous studies should generalize to excluded groups. Such regulations seem to make little sense if we as researchers do not explore and leverage the diversity, when it is clinically relevant, to improve decision making. A foundational definition of personalized medicine urges us to maximize patient-centered health outcomes using as much knowledge about the individual's state as is available.<sup>10</sup> If patients are to receive the best, most individualized care, then it seems race and ethnicity should sometimes be part of the equation.

### Sources of Funding

Supported by the National Institutes of Health Administrative Supplement Grant U01NS086294-03S1 and a Patient-Centered Outcomes Research Institute Methods Research Award (grant number: ME-1606-35555).

### Acknowledgments

We thank Leonard Katz, PhD, for his thoughtful review and comments on a draft of the article.

### Disclosures

None.

### References

1. Goff DC Jr, Lloyd-Jones DM, Bennett G, Coady S, D'Agostino RB, Gibbons R, Greenland P, Lackland DT, Levy D, O'Donnell CJ, Robinson JG, Schwartz JS, Shero ST, Smith SC Jr, Sorlie P, Stone NJ, Wilson PW, Jordan HS, Nevo L, Wnek J, Anderson JL, Halperin JL, Albert NM, Bozkurt B, Brindis RG, Curtis LH, DeMets D, Hochman JS, Kovacs RJ, Ohman EM, Pressler SJ, Sellke FW, Shen WK, Smith SC Jr, Tomaselli GF; American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 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. doi: 10.1161/01.cir.0000437741.48606.98.
2. VanderWeele TJ, Robinson WR. On the causal interpretation of race in regressions adjusting for confounding and mediating variables. *Epidemiology*. 2014;25:473-484. doi: 10.1097/EDE.0000000000000105.
3. Kaufman JS, Cooper RS. Commentary: considerations for use of racial/ethnic classification in etiologic research. *Am J Epidemiol*. 2001;154:291-298.
4. Krieger N. Discrimination and health inequities. In: Berkman LF, Kawachi I, Glymour M, eds. *Social Epidemiology*. New York, NY: Oxford University Press; 2014:63-125.
5. Gould SJ. *The Mismeasure of Man*. New York, NY: WW Norton & Company; 1996.
6. Lillquist E, Sullivan CA. The law and genetics of racial profiling in medicine. *Harv Civ Rights-Civil Lib Law Rev*. 2004;39:391-483.
7. Paulus JK, Shah ND, Kent DM. All else being equal, men and women are still not the same: using risk models to understand gender disparities in care. *Circ Cardiovasc Qual Outcomes*. 2015;8:317-320. doi: 10.1161/CIRCOUTCOMES.115.001842.
8. Ryan AM. Will value-based purchasing increase disparities in care? *N Engl J Med*. 2013;369:2472-2474. doi: 10.1056/NEJMp1312654.
9. Yudell M, Roberts D, DeSalle R, Tishkoff S. Science and society. Taking race out of human genetics. *Science*. 2016;351:564-565. doi: 10.1126/science.aac4951.
10. Pauker SG, Kassirer JP. Decision analysis. *N Engl J Med*. 1987;316:250-258. doi: 10.1056/NEJM198701293160505.

KEY WORDS: cardiovascular diseases ■ clinical decision making ■ ethnic groups ■ evidence-based medicine ■ race ■ risk factors