Race

Ritual, Regression, and Reality

Jay S. Kaufman

*T*he Bell Curve, published exactly 20 years ago by Herrnstein and Murray,[1] used regression analyses to argue for the genetically determined intellectual inferiority of black Americans. It is difficult to think of any other book with over 100 pages of statistical appendices that spent so much time on the New York Times best-seller list (14 weeks) and entered so deeply into the popular discourse on the meaning of race in American society. The book's argument for an inherent and immutable component of black inferiority is a counterfactual one in which blacks, if provided the same social environment as whites, would still have poorer educational outcomes and therefore lower social status achievement. The analyses reported in the book, however, did not involve direct measure of any genetic data or any experimental manipulations of education. How then does one move from passively observed data on groups of individual persons with different measured and unmeasured characteristics to a statement about what would happen under an intervention that has never actually taken place? The nascent field of "causal inference" aims to define the assumptions that would be necessary to make such a leap and provide precise definitions for quantitative estimators of the effects caused by such (hypothetical) interventions.

In the field of epidemiology (and in biomedicine more generally), race is a ubiquitous element of data description and analysis, usually following the peculiar American system of 5 racial categories plus Hispanic ethnicity.[2] As an indicator of socially defined subpopulations, its use as a stratification or adjustment variable is relatively uncontroversial.[3] However, when the inferential focus of the analysis is race itself, in particular a partial regression coefficient for a racial-group indicator variable, then it becomes much more difficult to explain what we are trying to accomplish and on what logical basis we hope to achieve this.

The new article by VanderWeele and Robinson[4] goes a long way to address these questions, using the formalisms that are by now the convention in the epidemiologic literature for defining and interpreting causal effects. They achieve an admirable clarity of exposition, and their unique contribution is to propose a causal interpretation of the adjusted coefficient for race without implying any hypothetical intervention on race itself. Rather, hypothetical intervention is made only on correlates of race, to unconfound a sort of total effect, or on consequences of race, to estimate a sort of direct effect (ie, that part of the total effect that is not mediated through a specified pathway). The "sort of" qualifiers derive from the fact that standard definitions of total and direct effects do involve hypothetical manipulation of the exposure, and so the quantities expressed by VanderWeele and Robinson[4] are not standard. Nonetheless, in accordance with the specified directed acyclic graphs (DAGs), as long as one does not probe too deeply into what is meant specifically by the arrow coming out of the race node, the quantities seem eminently reasonable. But are they helpful in practice?

Despite the admittedly oversimplistic DAGs, even just recognizing the distinction between a total effect versus direct effect interpretation, delineated so painstakingly and effectively by these authors, would dramatically improve the bulk of the published literature. There are innumerable deficiencies in the use of race in biomedical studies, but my unsystematic assessment is that a failure to appreciate this rudimentary distinction is the most common failing.[5] Another real contribution of the article by VanderWeele and Robinson[4] is the careful specification of the assumptions necessary to make this model work. These include, among other considerations, the correct specification of the DAG, the absence of residual confounding, and the absence of measurement error.[6] A sane reader might immediately recognize these conditions as hopelessly unattainable in practice and even impossible to approximate. VanderWeele and Robinson[4] do not dwell on this thorny problem, but they are dutifully honest about these impossible criteria, and this is a commendable strength of their article, even though such honesty serves to discourage the wider application of the quantities they define.

Soon after the publication of *The Bell Curve*, various critical reviews appeared in scholarly journals and books. A good number focused on technical questions over the interpretation of the regression coefficients, including a merciless thrashing from Heckman[7] and a similarly pointed reanalysis of the same data set by Korenman and Winship.[8] In these and other critiques, measurement error and residual confounding were among the most commonly cited concerns. For example, Heckman wrote:

> It would be miraculous if 15–23 years of environmental influences could be summarized by a composite of education, occupation, and family income measured in one year. If environment is poorly measured, Herrnstein and Murray's evidence that IQ has a stronger impact on socioeconomic outcomes than their measure of environment could rise for spurious reasons. We have already seen that education affects test scores. With a standard errors-in-variables argument, their measure of IQ may proxy the mismeasured environmental variable, and if so, the importance of IQ will be overstated …. Similar remarks apply to the authors' study of racial and ethnic differentials in socioeconomic outcomes. If racial differentials in environments affect ability and influence measured test scores … evidence that racial differentials weaken when ability is accounted for using regression methods does not rule out an important role for the environment in explaining performance in society. In the presence of measurement error in the environmental variables, the authors' analysis will overstate the "true" effect of ability on those outcomes. [7,pp1113–4]

Thus, a careful reader of the article by VanderWeele and Robinson[4] will readily dismiss estimates from models such as those used by Herrnstein and Murray,[1] not because they require intervention on race (which in fact they do not) but rather because they fail to approximate any reasonable identification of meaningful inference in a real-world context. Even if we could characterize the complete environment at birth in hopes of unconfounding a total effect estimate, this would probably not be anywhere close to sufficient. If the 20th century belonged to Charles Darwin, it is looking increasingly as if the 21st century will be handed back to Jean-Baptiste Lamarck, given the explosion of recent developments in epigenetics.[9] To the extent that disadvantages can be acquired during the lifetimes of parents and grandparents and passed on to their offspring, then no accounting for status at birth will ever suffice to remove environmental influences and reveal the innate aspects of race through the adjusted regression coefficients.

Another particularly unforgiving critique of *The Bell Curve* was published by Glymour[10] in 1997. Deploying the term "pseudo-science" 20 times in as many pages, Glymour took aim not only at the work of Herrnstein and Murray[1] but also at the analytic rituals in social sciences that permitted such a work to appear normative in the technical sense, even if uncomfortable in its substantive conclusions. Like VanderWeele and Robinson,[4] Glymour reviewed the assumptions necessary to grant a causal interpretation to an adjusted regression coefficient, and then declared that social science was generally so far from approximating these assumptions that any application in this arena could not qualify as science. Glymour was by no means alone in this pessimism.[11,12]

Echoing similar sentiments now 17 years later, Glymour and Glymour[13] express several profound doubts about the new article by VanderWeele and Robinson. They argue that socioeconomic status (SES) is not so easy to equalize by any practical intervention because status is complex and multidimensional. In a racially stratified society, one could argue that race itself is a key component of social status. But the commentators[13] reject the notion that hypothetical intervention is a foundation of causal inference in any case and argue instead for a kind of qualitative inference, which they refer to as "etiological." An adjusted regression coefficient in this framework would not predict the results of a hypothetical intervention. Instead, it would merely motivate us to make changes in the direction of greater social justice, without pretending to inform us exactly how the world might look after such changes.

The argument in favor of a qualitative or etiological notion of causation remains vague in this brief comment, but it does raise the important question of what biomedical researchers are intending when they ritualistically adjust race comparisons for measured covariates. Are they trying to predict the contrast that would remain after an intervention on race or on the covariates? Probably not. They are probably aiming for something much closer to the suggestion by Glymour and Glymour.[13] The problem is that many authors are following rote analytic traditions without articulating a specific

intention of any kind. If pinned down, they might say that the greater degree of justice sought is not thorough intervening on the race variable nor on its downstream consequences but rather on the arrow that connects them. That is, the intervention in mind is to redraw the DAG, such that we would live in a society in which there is no longer an arrow going from race to SES at all.[14]

There is another distinction that VanderWeele and Robinson[4] mention only briefly, which is between race conceived of as a social interaction and race conceived of as a set of immutable or essential characteristics hidden within the biology of an individual person. In the first case, race is a kind of caste system mediated through social presentation and recognition. Under this model, the causal mechanism is one of discrimination, and race is in principle manipulable.[15] Race could therefore function like any other observed exposure that epidemiologists study. A number of randomized trials have been based on this idea, a few of which are referenced by Vander-Weele and Robinson.[4] It is only the second model, of immutable or essential biologic characteristics, that is impervious to standard causal inference assumptions and which cannot be assigned in a trial. There remains some disagreement about whether race and sex are causes, but this is only in the case of the second model. In contrast, there is broad agreement that racism can be studied, even if some continue to argue that race cannot. And so, to the extent that the race variable in a regression model represents social treatment and the judgments made by others, there is no impasse whatsoever.

The theologian Reinhold Niebuhr famously prayed for the serenity to accept things that could not be changed, the courage to change things that must be changed, and the wisdom to know the difference. Our field owes a tremendous debt of gratitude to VanderWeele and Robinson[4] and to Glymour and Glymour[13] for extending and refining that wisdom.

ABOUT THE AUTHOR

Jay S. Kaufman is Professor and Canada Research Chair in Health Disparities in the Department of Epidemiology, Biostatistics, and Occupational Health at McGill University. His work focuses on social epidemiology, analytic methodology, causal inference, and evaluation of interventions. He is an editor of Epidemiology, *an associate editor of the* American Journal of Epidemiology, *and co-editor of the textbook* Methods in Social Epidemiology.

REFERENCES

1. Herrnstein RJ, Murray CA. *The Bell Curve: Intelligence and Class Structure in American Life*. New York, NY: Free Press; 1994:845 pp.
2. Executive Office of the President, Office of Management and Budget. Race and ethnic standards for federal statistics and administrative reporting. Statistical policy directive no. 15. Federal Register 1994 (June 9);59:FR 29831–5. Available at: http://wonder.cdc.gov/wonder/help/populations/bridged-race/directive15.html. Accessed 20 February 2014.
3. Kaufman JS, Cooper RS. Commentary: considerations for use of racial/ethnic classification in etiologic research. *Am J Epidemiol*. 2001; 154:291–298.
4. VanderWeele TJ, Robinson WR. On causal interpretation of race in regressions adjusting for confounding and mediating variables. *Epidemiology*. 2014;25:473–484.
5. Kaufman JS. Dissecting disparities. *Med Decis Making*. 2008;28:9–11.
6. Kaufman JS, Cooper RS, McGee DL. Socioeconomic status and health in blacks and whites: the problem of residual confounding and the resiliency of race. *Epidemiology*. 1997;8:621–628.
7. Heckman JJ. Lessons from the bell curve. *J Polit Econ*. 1995:103: 1091–1120.
8. Korenman S, Winship C. A reanalysis of the bell curve: intelligence, family background, and schooling. In: Arrow KJ, Bowles S, Durlauf SN, eds. *Meritocracy and Economic Inequality*. Princeton, NJ: Princeton University Press; 2000:137–178.
9. Mill J, Heijmans BT. From promises to practical strategies in epigenetic epidemiology. *Nat Rev Genet*. 2013;14:585–594.
10. Glymour C. Social statistics and genuine inquiry: reflections on the bell curve. In: Devlin B, Fienberg SE, Resnick DP, Roeder K, eds. *Intelligence, Genes, and Success*. Chapter 12. New York: Springer; 1997:257–280.
11. Clogg CC, Haritou A. The regression method of causal inference and a dilemma confronting this method. In: McKim VR, Turner PS, eds. *Causality in Crisis*. Notre Dame, IN: University of Notre Dame Press; 1997:83–112.
12. Freedman DA. Statistical models and shoe leather. *Sociol Methodol*. 1991;21:291–313.
13. Glymour C, Glymour MR. Race and sex are causes. *Epidemiology*. 2014;25:488–490.
14. Krieger N. Epidemiology and the web of causation: has anyone seen the spider? *Soc Sci Med*. 1994;39:887–903.
15. Kaufman JS. Epidemiologic analysis of racial/ethnic disparities: some fundamental issues and a cautionary example. *Soc Sci Med*. 2008;66:1659–1669.