COMMENTARY

DEFINING "RACE" AS THE DEFINING PROBLEM

Timothy Caulfield[*]

TABLE OF CONTENTS

I. INTRODUCTION

It has been stated that race "is a thing of our world like no other."[1] It is a concept that has had a long, prominent, and (usually) less than constructive role in world history. And it remains the focus of intense academic debate. But despite this central role and continued scrutiny, it is a concept that remains awfully difficult to define. What is "race"? Is it how we view ourselves or how we are categorized by others? Does it relate to skin color or ethnic background? Is it the discrete and simplistic

* Timothy Caulfield is a Canada Research Chair in Health Law and Policy; Senior Health Scholar, Alberta Heritage Foundation for Medical Research; Professor, Faculty of Law and School of Public Health; and Research Director, Health Law Institute, University of Alberta. I would like to thank Ubaka Ogbogu for his insight and Genome Alberta for funding support.

1. Duana Fullwiley, *Race and Genetics: Attempts to Define the Relationship*, 2 BIOSOCIETIES 221, 225 (2007).

categories created by European biologists in the 1700s or a more nuanced social community generated by shared cultural experience? Of course, race can be all of these things and none of them. Its meaning and definition change depending on context, perspective, and time.

It is this complexity of meaning that is largely missing from the interesting and provocative paper by Professor Michael Malinowski.[2] As in all his work, Professor Malinowski has done a truly impressive job of drawing together a large amount of diverse commentary. And he takes a brave position. By advocating for race-based genetic research, he seems to be swimming against the prevailing currents in academia. In fact, Professor Malinowski's thesis is not as controversial as it initially sounds. He is not promoting research on the biological and genetic foundations of race (which is the research activity that causes the most uniform cautionary response). Rather, he suggests that we need to recognize the value of exploring genetic diversity and using race as a research variable. Research on genetic variability has already led to many interesting health related findings. Malinowski hints throughout his paper that we should not let politics and fear of misuse interfere with this potentially valuable research—a sentiment with which I have great sympathy.

But despite these and many other worthwhile points, Professor Malinowski's paper does miss some of the key themes in the current race and research debate. In this brief Commentary, I seek to highlight a few of these themes, with a particular emphasis on those that relate to the representation and definition of "race." Indeed, I believe it is how "race" is represented and defined (or *not* defined) that creates some of the primary challenges associated with the current wave of genetic variability research.

II. THE VALUE OF "RACE" IN RESEARCH

As a starting point, it is important to note that despite Professor Malinowski's main thesis (i.e., that many scholars are opposed to the use of race in research), it is my impression that almost everyone, regardless of disciplinary background, recognizes the value of race as a research variable (the study of health disparities, for example, demands an understanding of the

2. Michael J. Malinowski, *Dealing with the Realities of Race and Ethnicity: A Bioethics-Centered Argument in Favor of Race-Based Genetics Research*, 45 HOUS. L. REV. 1415 (2009).

role of race). As noted in a recent interdisciplinary consensus paper, which involved Professor Malinowski, "[W]hile there is much agreement that race (and other forms of social identification) matters to health, there is little agreement about why or how race matters, how best to study its effects, and how to translate and communicate research results from racially stratified studies."[3]

So, whether "race" matters is not really in dispute. What stirs the most controversy is the ways in which "race" is used and described (and, of course, the social implication of that use—but this latter point is largely beyond the scope of this short piece). For example, the social construct of race clearly has relevance to health outcomes.[4] If one is studying issues associated with access to health care, for instance, "race" can often be an important variable.[5] Likewise, the genetic variability of individuals and groups can be a factor in health disparities and health outcomes. But the use of "race" in the socioeconomic context of a study on access to health care is much different than using race as a variable in the exploration of genetic variation. The category of "race" is rarely (if ever) an appropriate approximation of underlying genetic variability—and this is particularly so if by "race" we mean those broad social categories such as black, white, and Asian.

Numerous commentators have noted the danger of using "race" as a proxy for describing genetic variability.[6] Francis Collins, former Director of the U.S. National Human Genome Research Institute, noted that "[a] true understanding of disease

---

3. Timothy Caulfield, Stephanie M. Fullerton, Sarah E. Ali-Khan et al., *Race and Ancestry in Biomedical Research: Exploring the Challenges*, 1 GENOME MED. (forthcoming 2009).

4. Of course, even when "race" is used in this context, there are challenges. *See, e.g.*, Margaret A. Winker, *Measuring Race and Ethnicity: Why and How?*, 292 J. AM. MED. ASS'N 1612, 1613 (2004) ("The use of race as a proxy for unmeasured confounders, such as cultural, social, and environmental influences, is commonplace, but race is a poor proxy for these measures. The life experience and cultural milieu of US immigrants may be completely different from those who grew up in the United States, despite being assigned to similar racial or ethnic categories. Socioeconomic status, not race, is likely the greater determinant of health and health-related qualities. Therefore, race is not a substitute for carefully assessed social and cultural characteristics.").

5. *See generally* David R. Williams & Pamela Braboy Jackson, *Social Sources of Racial Disparities in Health*, 24 HEALTH AFF. 325 (2005).

6. See Caulfield et al., *supra* note 3, for a review of the literature on either side of the "proxy" debate. To be fair, whether "race" is a good proxy for genotype also depends on the definition and context. If "race" is being used to refer to a discrete, geographic region, then its value as a proxy may increase. If, however, race is used to reflect broad—and, I believe, socially constructed—notions of black, white, and Asian, then its biological relevance decreases.

risk requires a thorough examination of root causes. 'Race' and 'ethnicity' are poorly defined terms that serve as flawed surrogates for multiple environmental and genetic factors in disease causation."[7] This idea was reiterated in another, recent, interdisciplinary consensus paper led by Stanford's Sandra Soo-Jin Lee: "We discourage the use of race as a proxy for biological similarity and support efforts to minimize the use of the categories of race and ethnicity in clinical medicine."[8]

III. TERMINOLOGY SLIPPAGE

Given the emerging agreement about the challenges of using race as a variable in biomedical research, what is the problem? In other words, because many in the research community seem relatively sensitive to the associated risks and problems (a reality emphasized by Malinowski in his telling of the James Watson debacle),[9] why all the fuss?

First, there is ample evidence that, despite this sensitivity, "race" is often used with less than ideal precision. In part, this is because there is no consistent definition. As noted above, "race" is a fluid term. This can lead to confusion about the relationship between "race" and genetic variability. For example, George Ellison and his colleagues have "found that there is a lack of consensus about what race and ethnicity mean and how these should be operationalised. As a result, researchers and practitioners may conflate the utility of racial and ethnic categories for sampling diverse study populations with their ability to identify and address aetiological variation therein."[10] A study by Hasan Shanawani and colleagues from Duke University supports this pessimistic view of the use of race in genetic research. The Shanawani team examined 268 published genomic studies involving "race" as a study variable and found that 72% did not explain their methods for assigning race or ethnicity as an independent variable. Despite this oversight, 67% of the studies reached conclusions about genetics, health outcomes, and race.[11]

---

7. Francis Collins, *What We Do and Don't Know About Race, Ethnicity, Genetics and Health at the Dawn of the Genome Era*, 36 NATURE GENETICS S13, S13 (2004).

8. Sandra Soo-Jin Lee, Joanna Mountain, Barbara Koenig et al., *The Ethics of Characterizing Difference: Guiding Principles on Using Racial Categories in Human Genetics*, 9 GENOME BIOLOGY 404, 404.2 (2008).

9. Malinowski, *supra* note 2, at 1457–59.

10. George Ellison, Andrew Smart, Richard Tutton, Simon M. Outram, Richard Ashcroft & Paul Martin, *Racial Categories in Medicine: A Failure of Evidence-Based Practice?*, 4 PLOS MED. e287, e287 (2007).

11. H. Shanawani, L. Dame, D.A. Schwartz, & R. Cook-Deegan, *Non-Reporting and*

Second, even when more precise variables are used and described—such as reference to a specific geographic ancestry—there can be slippage back to more ambiguous terms. Indeed, we see this in a wide variety of publications, including academic articles. In his paper in this volume, for example, Professor Malinowski describes research on the genetic causes of adverse drug reactions.[12] In describing this work (and in just a few paragraphs), he uses a wide variety of terms: "ethnic," "race," "Asian," "South Asian," and "Japanese." In the end, the relevant genetic variation appears to exist "in a small percentage of persons with Japanese origin."[13] In the context of this work, what is the right term? Race? Are the Japanese a "race"? Again, it depends on perspective.

There are numerous examples of this kind of ambiguity in the popular press.[14] There are times when a genetic variability study in a biomedical journal will make no mention of the term "race," but the corresponding media report will describe the work by referring to "race" or a term associated with a broad racial category (e.g., black, white, or Asian). To cite just one example, a recent study examined the ways in which genetic variation can impact how different populations respond to drugs and fight infections. The academic report of the study makes no mention of "race," "Caucasians," or racial difference.[15] On the contrary, the paper used very specific geographic regions, including "individuals of European ancestry from Utah" and "Yoruba individuals from Ibadan, Nigeria."[16] The media coverage, however, talked about the genetic differences between the two "races."[17]

A third reason that the representation and use of race in genetic research continues to be a problem relates to a lack of journal publication guidelines on point. While there are many recommendations in the literature about how researchers should

---

Inconsistent Reporting of Race and Ethnicity in Articles that Claim Association Among Genotype, Outcome and Race or Ethnicity, 32 J. MED. ETHICS 724, 724 (2006).

12. Malinowski, *supra* note 2, at 1441–42.

13. *Id.* at 1441.

14. *See, e.g.*, Timothy Caulfield & Simrat Harry, *Popular Representations of Race: The News Coverage of BiDil*, 36 J.L. MED. & ETHICS 485 (2008).

15. W. Zhang, S. Duan, E.O. Kistner, W.K. Bleibel et al., *Evaluation of Genetic Variation Contributing to Differences in Gene Expression Between Populations*, 82 AM. J. HUM. GENETICS, 631–40 (2008).

16. *Id.* at paras. 1–2.

17. *Immune System Differences Found*, BBC NEWS, Feb. 29, 2008, http://news.bbc.co.uk/go/pr/fr/-/2/hi/health/7270562.stm. Another recent example is a study on high blood pressure, done primarily on the Hutterite.

handle race in biomedical research,[18] few biomedical journals have clear instructions on point. A 2006 study by Simon Outram and George Ellison looked at 120 journals that published in the areas of genetics and heredity. "Of those, only two had instructions for authors that referred directly to race or ethnicity."[19] This lack of guidance likely contributes to the large amount of variability and ambiguity in the use of the term.

IV. THE SOCIAL CONCERNS

If the real challenge is not the use of "race" in research, but the *way* in which it is used, one can ask: "so what"? What is the social harm? Is this just a game of semantics?

When race is used without explanation or in an ambiguous manner, there is a danger that interpretations of data will lead to an inappropriate conflation between the social categories of "race" (such as black, white, and Asian) and genetic variability.[20] In other words, it can lead to a belief that these social categories are biologically significant and that they serve as an accurate proxy for underlying genetic variability.[21] In addition to being

---

18. See Caulfield et al., *supra* note 3, for a summary of recommendations. To be clear, no one is suggesting that the journals dictate what kind of research can be done. The recommendations are aimed at, for example, how variables are described and justified.

19. Simon M. Outram & George T. H. Ellison, *Improving the Use of Race and Ethnicity in Genetic Research: A Survey of Instructions to Authors in Genetics Journals*, 29 SCI. EDITOR 78, 78 (2006). The authors note the impact of the lack of guidelines thus: "That is extraordinary, given that over half (55.8%) the genetics journals surveyed had recently published empirical articles using 'race,' 'ethnicity,' and related census categories. However, our interviews with genetics journal editors suggest that many felt ill-equipped to tackle the issues involved and were reluctant to develop or enforce guidelines without the support of the genetics research community." *Id.* at 80.

20. For a discussion of the complex relationship between race and genetics, see Tim Ingold, *When Biology Goes Underground: Genes and the Spectre of Race*, 4 GENOMICS SOC. & POL'Y 23, 33 (2008) ("Genes and culture do not interact with the environment, nor are people and their behaviour the products of such interaction. If we ask what interacts with the environment, the answer is the people themselves. And through such interaction, people actively intervene in shaping the conditions of future development, both for themselves and for others. That is to say, people are the producers as much as the products of their own history, in a continuous process of social life.").

21. I think it is fair to say that there is a good deal of consensus that the broad social categories of race (such as black, white, and Asian) are a biological fiction. See, for example, Sandra Lee, Joanna Mountain & Barbara Koenig, *The Meanings of "Race" in the New Genomics: Implications for Health Disparities Research*, 1 YALE J. HEALTH POL'Y L. & ETHICS 33, 39 (2001), where they note that the "widely accepted consensus among evolutionary biologists and genetic anthropologists is that biologically identifiable human races do not exist." There is less consensus, however, on the latter point. While most agree that the use of skin color or self-indentified race is far from ideal, some have argued that it may be (at times) the best available proxy. See Caulfield et al., *supra* note 3, for a brief review of this literature.

scientifically misleading, Sandra Soo-Jin Lee notes that the "conflation of race with genetics opens the door to prejudice, racial stereotyping, and overly simplistic conceptualization of pharmacogenomic interactions, which could ultimately lead to poor health outcomes."[22]

There is at least some evidence that the concerns about prejudice and stereotyping are more than mere speculation.[23] A recent study by John Lynch et al. found that exposure to messages linking genetics and race increased some forms of racism (specifically, "genetically based racism").[24] These results correspond well to an earlier study that found "some messages linking race, genes, and health produce increases in racist attitudes in some audiences."[25] Such data, while preliminary, should remind us of the potential social consequences of genetic research that implicates racial categories.

V. CONCLUSION

Professor Malinowski's paper does a good job of underlining the value of exploring genetic variability. Such research requires the stratification of populations into groups. When doing so, there is mounting agreement that researchers should be cautious

---

22. S. Soo-Jin Lee, *Racializing Drug Design: Implications of Pharmacogenomics for Health Disparities*, 95 AM. J. PUB. HEALTH 2113, 2113 (2005). This is a point that has been made by many commentators. *See, e.g.*, Morris Foster & Richard Sharp, *Beyond Race: Towards a Whole-Genome Perspective on Human Populations and Genetic Variation*, 5 NATURE REVIEWS GENETICS 790, 790 (2004) ("[W]hen used to define populations for genetic research, race has the potential to confuse by mistakenly implying biological explanations for socially and historically constructed health disparities."); Lee et al., *supra* note 8, at 404.3 ("Overemphasizing the genetic contribution to complex human disease or behavioral traits can promote not only racism, but also a naive genetic essentialism—the notion that genes determine health status or behavior.").

See also L. Braun, A. Fausto-Sterling, D. Fullwiley, E.M. Hammonds, A. Nelson, et al., *Racial Categories in Medical Practice: How Useful Are They?*, 4 PLOS MED. e271, e271 (2007), in which it is noted that the current policy of including "race" in some forms of research should be reexamined, for the reasons noted in this footnote. "In the short run, the National Institutes of Health needs to re-examine its race-based research rules, weighing the balance between attempting to include minority populations in our health care system, on the one hand, without forcing us into a misconstrual of race as biology on the other." *Id.*

23. For an interesting analysis of the complex relationship between representations and public understanding of race, see Benjamin R. Bates et al., *Evaluating Direct-to-Consumer Marketing of Race-Based Pharmacogenomics: A Focus Group Study of Public Understandings of Applied Genomic Medication*, 9 J. HEALTH COMM. 541 (2004).

24. John Lynch, Jennifer Bevan, Paul Achter, Tina Harris & Celeste M. Condit, *A Preliminary Study of How Multiple Exposures to Messages About Genetics Impact on Lay Attitudes Towards Racial and Genetic Discrimination*, 27 NEW GENETICS & SOC'Y 43, 51 (2008).

25. Celeste Condit et al., *Exploration of the Impact of Messages About Genes and Race on Lay Attitudes*, 66 CLINICAL GENETICS 402 (2004).

in the use of the concept of "race." In his article, Malinowski underplays this definitional aspect of the current debate. Indeed, given the history of the use of race in biomedical research,[26] there are ample reasons to err on the side of caution. No one is suggesting that valuable research should not continue. But good science and good ethics demand that we be careful in how we describe, report on, and interact with the populations under study.

26. *See, e.g.*, Susan L. Smith, *Mustard Gas and American Race-Based Human Experimentation in World War II*, 36 J.L. MED. & ETHICS (2008).