

Defining the Race and Ethnic Standards for Federal Statistics and Administrative Reporting

[Rene Canady](#)^{1*}, [Jorge Jimenez](#)^{2*}, [Danesh Thirukumar](#)^{3*}

¹Washington University in St. Louis, Department of Sociology, St. Louis, MO

²University of Pittsburgh, Department of Bioengineering, Pittsburgh, PA

³Georgetown University, Department of Neuroscience, Washington, D.C.

*All authors contributed equally

<https://doi.org/10.38126/JSPG180406>

Corresponding author: daneshthirukumar@gmail.com

Keywords: race; ethnicity; health disparities; racial bias; social inequality; bioethics

Executive Summary: Race describes cultural, historical, and oppressive relationships in society. The use of race in biomedical and scientific studies has been a powerful tool that can reinforce and alter society's current assumptions about race. Some of the historical uses of race include evidence for race-based medicine, biological inferiority, and genocide. These uses have all used race as a crude proxy for genetic makeup, rather than a biological expression of the social environment that infiltrates the health and well-being of every American. By defining race and its social and cultural impacts on identity and the human experience within research, the field of biomedicine will improve clarity and integrity in addressing historical, scientific, and clinical inequalities. Currently, the Office of Management and Budget (OMB) does not contain a definition of race and uses homogeneous ethnic categories when reporting population statistics. We propose that the definition of race be added in the collection of race data as a requirement of the OMB for nationally conducted research.

I. Background

i. Statement of issue

Although humans have been deemed 99.9% similar at the DNA level (Collins and Mansoura 2001), race is still a key consideration in research due to its societal relevance. The US federal government retains a vested interest in race and research, as recently a Request for Information was put forth by the National Institutes of Health (NIH) requesting for suggestions to address racial inequalities present in biomedical research (NIH 2021).

Data and statistics at the federal level are coordinated by the OMB through the Statistical Programs and Standards (The White House 2021). Specifically, data collection, maintenance, and reporting standards on race and ethnicity are provided by directive 15 of the OMB statistical policy directives (CDC 2019). The OMB also oversees the NIH budget and allocation of

resources, making the OMB reliant upon NIH to fund studies that may provide useful statistics at the federal level. These statistics may be flawed, as discrepant conclusions may be drawn due to the lack of a clear definition on how the government defines race. Thus, the federal statistics which are relied upon to make decisions in both the public and private sectors may be skewed. This potentially misguided reliance has broad-ranging impacts both clinically and socioeconomically.

ii. Clinical considerations

Race as a variable for medical and clinical research has been a controversial issue since its introduction to US medical curricula as a type of leprosy by Benjamin Rush in 1790 (Ioannidis et al. 2021). Throughout history, several examples of the misuse of race as a biological variable are represented. For example, in 1843 Josiah Nott stated that "mulatto women are particularly delicate—are subject to

many chronic diseases, and especially derangement of catamenia... and other diseases particular to females" (Nott 2010), which specifically designates a diagnosis of disease specific to "mulatto" females who are of heterogeneous race. These examples were scientifically disproven by the Human Genome Project of 1999, which emphasized race as nonbiological with no basis in genetic code (Lin and Kelsey 2000). Unfortunately, the call to end the use of race as a biological category in genetic research and subsequently clinical research has been largely ignored. It is possible that the inconsistent definition and historical use of taxonomic categorization based on hereditary traits, e.g., skin color (Yudell et al. 2016), has led to the misuse of race in clinical research. Some specific examples include the development of race-based medicine practices like diagnostic algorithms, and misrepresentation of populations in clinical trials and their outcomes.

The insertion of race into medicine includes the use of diagnostic algorithms that adjust their outputs on the basis of a patient's race or ethnicity (Vyas et al. 2020). Physicians use these algorithms to individualize risk assessment and guide clinical decisions. By embedding race into the basic data and decisions of health care, these algorithms propagate race-based medicine. One example, in obstetrics health, includes the use of the Vaginal Birth after Cesarean (VBAC) Risk Calculator (Vyas et al. 2019; Grobman et al. 2007), which estimates the probability of successful vaginal birth after a patient's prior cesarean section. As a result of historically racist assumptions about pelvic inferiority in Africans (Caldwell and Moloy 2016) and indigenous Mexicans (O'Brien 2012), the use of race for African American and Hispanic birthing people includes a correction factor that is subtracted from the estimated success rate for any person who identifies as Black or Hispanic. The decreased value for Black (0.671) or Hispanic (0.680) is almost as large as the benefit from prior vaginal delivery (0.888) or prior VBAC (1.003) leading to lower estimate scores (Vyas et al. 2019). These estimates may guide clinicians to not offer vaginal labor delivery for people of color, causing an equity concern for vaginal delivery between populations. An additional use of race-based medicine includes BMI risk for diabetes in Asian patients. Asian patients are presumed to develop more visceral than peripheral adipose tissue than patients of other races at similar BMI levels (Cerdena

et al. 2020) influenced by the implicit hypothesis on the inherent deficiencies in South Asian descent DNA (Sniderman et al. 2007). This association perpetuates the practice that Asians are at high risk for insulin resistance during Type-2 Diabetes screening (Hsu et al. 2015). The race-based medicine and screening presented here are a few examples of many instances (Cerdena et al. 2020) that highlight the misuse of race as biological and is also falsely investigated in the recruitment and outcomes of clinical trials.

Clinical research provides the means to develop innovative solutions to address medical ailments plaguing society. The Food and Drug Administration (FDA) and the NIH have contributed extensive efforts to recruit racial and ethnic minorities into clinical trials that seek to approve medical products for human use. These organizations also aim to conduct studies that proportionally represent the US general population (FDA 2020), which includes historically marginalized communities. Often, the recruitment of patients from historically marginalized communities into randomized clinical trials is designed to conduct subgroup analysis between populations (FDA 2020). Unfortunately, the design of clinical trials in this manner contributes to the assumption that race and ethnicity represent valid biological constructs that may modify the effect of any drug studied in a randomized trial and thus necessitate race-specific treatments (FDA 2020). Utilizing poorly defined race-based categories in scientific research manifests detrimentally in clinical settings. Such is the case with the clinical trial testing of Bidil (Bibbins-Domingo and Fernandez 2007), a heart failure medication specific for African Americans that largely ignored the social constructs of race and, instead, used it as a proxy for genetic factors causing heart disease (Roberts 2008). Bidil's clinical trial did not clearly show that the drug worked better in African Americans, as it relied on "self-identified" African Americans and lacked a comparison group (Kahn 2007). Bidil is a concrete example of how the sociocultural factors in disease may be ignored with the current assumption of a race and presumed genetic difference (Brody and Hunt 2006). This is further detrimental since race-based categories in clinical trials ignore the heterogeneous composition of racial groups (Yudell et al. 2016).

As currently presented, the NIH and FDA follow federal reporting of race and ethnicity in their studies. Both agencies contribute largely to American

health and the training of our biomedical workforce, e.g., NIH training grants (T32) and fellowships (F30, F31–Diversity) (NIH Individual Fellowships 2021; NIH Institutional Training Grants 2021). Enforcing a new definition to race at the federal level may cause downstream confusion among states, particularly among educational and medical systems that benefit from NIH and FDA funding. In this memorandum, we explore the possible outcomes of two policy options: 1) alter the Information Quality Act Guidelines, and 2) include the definition of race in the Race and Ethnic Standards for Federal Statistics and Administrative Reporting. We also explore inaction as a model to support clarity and the importance of our recommendations.

II. Policy options:

i. Option 1: Alter the Information Quality Act guidelines
Section 515 of Public Law 106-554, known as the Information Quality Act Guidelines (“Guidelines for Ensuring...” 2002), prioritize quality, objectivity, utility, and integrity of information. This language must be edited to bolster the importance of subjectivity of information, acknowledge the scientific research consequences on marginalized identities, and define race. We propose that this section be amended to state that:

- Subjectivity be added to the list of priorities in the Information Quality Act Guidelines
- Scientific research poses unequal consequences on marginalized identities
- “Race” refers to physical differences that groups and cultures consider socially significant, while “ethnicity” refers to shared culture, such as language, ancestry, practices, and beliefs. Additionally, define race, gender, and sexuality.

Advantages

This amendment would clarify that subjectivity is present in all research and can hold as much value as objectivity. Accepting subjective practices in data collection, theory, and analysis gives biomedical researchers a wider arrangement of tools to mitigate racial bias and properly contextualize race within studies. This presents the opportunity to fix common misinterpretations presented by majority groups during data analysis and real-world applications of racial and ethnic differences in biology. This change

would enforce high level consideration of marginalized groups when investigating scientific differences in health and behavior. Adding specific definitions of demographics will improve the accuracy of, and provide context for, demographic-based differences.

Disadvantages

Successful uptake of subjectivity as a research priority requires sociological understandings of subjectivity that are currently not part of most research trainings. The knowledge required to comply with changes would need to come from additional training. Additionally, considering the little contact between the OMB and researchers, enforcement challenges are likely to ensue, as with all ethical guidelines. Given that objectivity is a universal concept taught to scientific researchers, it is unlikely that this language would be adopted with ease. Additionally, acknowledging that every researcher has biases may be mistaken as permission for researchers to use their own.

ii. Option 2: Include the definition of race in the Race and Ethnic Standards for Federal Statistics and Administrative Reporting.

The definition of race for Race and Ethnic Standards for Federal Statistics and Administrative Reporting (CDC 2019) should include a clear definition of race. We propose the standards should be amended to additionally state that:

- “Race” refers to physical differences that groups and cultures consider socially significant, while “ethnicity” refers to shared culture, such as language, ancestry, practices, and beliefs. Additionally, define race, gender, and sexuality.
- There are historical, social, and physical differences across races. Biological demarcations are frequent, yet inaccurate identifiers of race.

Advantages

Redefining race as a social construct would show racial differences in biological studies as functions of social environment, where the root of differences could be addressed. Scientists who use race data are encouraged to state social differences instead of biological differences. This would encourage

scientists to engage with racial disparities as social problems that require interdisciplinary solutions.

Disadvantages

This rule would push responsibility of racial differences into the social science field, potentially distancing natural scientists as well as technologists. However, the responsibility of social welfare is always of utmost importance in scientific research. Additional education may be required in order to introduce race in a productive manner.

iii. Option 3: Inaction

Do not alter the OMB's documentation providing guidance over scientific research.

Advantages

Inaction would be a convenient short-term solution for policymakers, as well as scientists who wish to be uninvolved with social issues. Researchers engage in race in the social traditions of their fields, allowing for more contextualized discussion.

Disadvantages

No one is held accountable for the inaccuracy of research using race. Race remains ill-defined and

continues to be used to make consequential decisions, harming racial minorities. Without an accurate definition of race, personalized medicine is unlikely to progress in a constructive manner. Lastly, racial ambiguity is unresolved, and interpretations are left to the convention of the field and researcher.

III. Policy recommendation

We recommend that the OMB alter the Race and Ethnic Standards for Federal Statistics and Administrative Reporting to include the definition of race (Option 2). We propose the use of the definition of race and ethnicity as defined by the American Sociological Association (American Sociological Association 2021) which further supports the scientific fact that race is non-biological in origin. While Option 1, altering the Information Quality Act Guidelines, gives researchers freedom to mitigate their own biases, it does not give clear guidance on how to do so. Option 2 provides guidance to researchers by defining and reporting a common misconception of race. Including the definition of race will improve the lives of Americans by increasing the accuracy of the scientific research that drives their health, socioeconomic status, and educational attainment.

References

- "67 FR 8451 - Guidelines for Ensuring and Maximizing the Quality, Objectivity, Utility, and Integrity of Information Disseminated by Federal Agencies; Republication - Content Details - R2-59." Accessed May 2, 2021. <https://www.govinfo.gov/app/details/FR-2002-02-22/R2-59>
- American Sociological Association. "Race and Ethnicity." Accessed May 2, 2021. <https://www.asanet.org/topics/race-and-ethnicity>.
- Bibbins-Domingo, Kirsten, and Alicia Fernandez. "BiDiI for Heart Failure in Black Patients: Implications of the U.S. Food and Drug Administration Approval." *Annals of Internal Medicine* 146, no. 1 (January 2, 2007): 52–56. <https://doi.org/10.7326/0003-4819-146-1-200701020-00009>.
- Brody, Howard, and Linda M. Hunt. "BiDiI: Assessing a Race-Based Pharmaceutical." *Annals of Family Medicine* 4, no. 6 (November 2006): 556–60. <https://doi.org/10.1370/afm.582>.
- Caldwell, W. E., and H. C. Moloy. "Anatomical Variations in the Female Pelvis and Their Effect in Labor with a Suggested Classification." *American Journal of Obstetrics and Gynecology* 26, no. 4 (October 1, 1933): 479–505. [https://doi.org/10.1016/S0002-9378\(33\)90194-5](https://doi.org/10.1016/S0002-9378(33)90194-5).
- Cerdeña, Jessica P., Marie V. Plaisime, and Jennifer Tsai. "From Race-Based to Race-Conscious Medicine: How Anti-Racist Uprisings Call Us to Act." *The Lancet* 396, no. 10257 (October 10, 2020): 1125–28. [https://doi.org/10.1016/S0140-6736\(20\)32076-6](https://doi.org/10.1016/S0140-6736(20)32076-6).
- Collins, F. S., and M. K. Mansoura. "The Human Genome Project. Revealing the Shared Inheritance of All Humankind." *Cancer* 91, no. 1 Suppl (January 1, 2001): 221–25. [https://doi.org/10.1002/1097-0142\(20010101\)91:1+<221::aid-cnrcr8>3.3.co;2-0](https://doi.org/10.1002/1097-0142(20010101)91:1+<221::aid-cnrcr8>3.3.co;2-0).

- Grobman, William A., Yinglei Lai, Mark B. Landon, Catherine Y. Spong, Kenneth J. Leveno, Dwight J. Rouse, Michael W. Varner, et al. "Development of a Nomogram for Prediction of Vaginal Birth After Cesarean Delivery." *Obstetrics & Gynecology* 109, no. 4 (April 2007): 806–12. <https://doi.org/10.1097/01.AOG.0000259312.36053.02>.
- Hsu, William C., Maria Rosario G. Araneta, Alka M. Kanaya, Jane L. Chiang, and Wilfred Fujimoto. "BMI Cut Points to Identify At-Risk Asian Americans for Type 2 Diabetes Screening." *Diabetes Care* 38, no. 1 (January 1, 2015): 150–58. <https://doi.org/10.2337/dc14-2391>.
- "Individual Fellowships | Research Training and Career Development." Accessed May 2, 2021. <https://researchtraining.nih.gov/programs/fellowships>.
- "Institutional Training Grants | Research Training and Career Development." Accessed May 2, 2021. <https://researchtraining.nih.gov/programs/training-grants>.
- Ioannidis, John P. A., Neil R. Powe, and Clyde Yancy. "Recalibrating the Use of Race in Medical Research." *JAMA* 325, no. 7 (February 16, 2021): 623. <https://doi.org/10.1001/jama.2021.0003>.
- Kahn, Jonathan. "Misreading Race and Genomics after BiDil." *Nature Genetics* 37, no. 7 (July 2005): 655–56. <https://doi.org/10.1038/ng0705-655>.
- Lin, Scarlett S., and Jennifer L. Kelsey. "Use of Race and Ethnicity in Epidemiologic Research: Concepts, Methodological Issues, and Suggestions for Research." *Epidemiologic Reviews* 22, no. 2 (January 1, 2000): 187–202. <https://doi.org/10.1093/oxfordjournals.epirev.a018032>.
- "NOT-OD-21-066: Request for Information (RFI): Inviting Comments and Suggestions to Advance and Strengthen Racial Equity, Diversity, and Inclusion in the Biomedical Research Workforce and Advance Health Disparities and Health Equity Research." Accessed May 2, 2021. <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-21-066.html>.
- Nott, J. C. "The Mulatto a Hybrid — Probable Extermination of the Two Races If the Whites and Blacks Are Allowed to Intermarry." Research-article. <http://dx.doi.org/10.1056/NEJM184308160290201>. Massachusetts Medical Society, June 2, 2010. World. <https://doi.org/10.1056/NEJM184308160290201>.
- O'Brien, Elizabeth. "Pelvimetry and the Persistence of Racial Science in Obstetrics." *Endeavour, Continuity and Change in the History of Mexican Public Health*, 37, no. 1 (March 1, 2013): 21–28. <https://doi.org/10.1016/j.endeavour.2012.11.002>.
- "OMB DIRECTIVE 15: RACE AND ETHNIC STANDARDS FOR FEDERAL STATISTICS AND ADMINISTRATIVE REPORTING." Accessed May 2, 2021. <https://wonder.cdc.gov/wonder/help/populations/bridged-race/directive15.html>.
- Rathore, Saif S, and Harlan M Krumholz. "Race, Ethnic Group, and Clinical Research." *BMJ: British Medical Journal* 327, no. 7418 (October 4, 2003): 763–64.
- Research, Center for Drug Evaluation and. "Conducting Clinical Trials." FDA. FDA, June 30, 2020. <https://www.fda.gov/drugs/development-approval-process-drugs/conducting-clinical-trials>.
- Roberts, Dorothy E. "Is Race-Based Medicine Good for Us?: African American Approaches to Race, Biomedicine, and Equality." *The Journal of Law, Medicine & Ethics* 36, no. 3 (August 1, 2008): 537–45. <https://doi.org/10.1111/j.1748-720X.2008.302.x>.
- Sniderman, Allan D, Raj Bhopal, Dorairaj Prabhakaran, Nizal Sarrafzadegan, and Andre Tchernof. "Why Might South Asians Be so Susceptible to Central Obesity and Its Atherogenic Consequences? The Adipose Tissue Overflow Hypothesis." *International Journal of Epidemiology* 36, no. 1 (February 1, 2007): 220–25. <https://doi.org/10.1093/ije/dyl245>.
- The White House. "Statistical Programs & Standards." Accessed April 15, 2021. <https://www.whitehouse.gov/omb/information-and-regulatory-affairs/statistical-programs-standards/>.
- Vyas, Darshali A., David S. Jones, Audra R. Meadows, Khady Diouf, Nawal M. Nour, and Julianna Schantz-Dunn. "Challenging the Use of Race in the Vaginal Birth after Cesarean Section Calculator." *Women's Health Issues: Official Publication of the Jacobs Institute of Women's Health* 29, no. 3 (June 2019): 201–4. <https://doi.org/10.1016/j.whi.2019.04.007>.
- Vyas, Darshali A., Leo G. Eisenstein, and David S. Jones. "Hidden in Plain Sight — Reconsidering the Use of Race Correction in Clinical Algorithms." *New England Journal of Medicine* 383, no. 9 (August 27, 2020): 874–82. <https://doi.org/10.1056/NEJMms2004740>.

Yudell, Michael, Dorothy Roberts, Rob DeSalle, and Sarah Tishkoff. "Taking Race out of Human Genetics." *Science* 351, no. 6273 (February 5, 2016): 564–65. <https://doi.org/10.1126/science.aac4951>.

Rene Canady's career goal is to understand Black liberation through the biomedical world. As an interdisciplinary scholar, she is interested in studying the racialization of the biomedical field in order to improve the holistic health of marginalized groups. She is currently a Sociology Ph.D. student at the University of Washington in Saint Louis and holds a B.S.E. in bioengineering from the University of Pittsburgh. Rene is both a fellow of the Spencer T. and Ann W. Olin Fellowship Program and the National Science Foundation Graduate Research Fellowship Program.

Jorge Jimenez is a queer Mayan interdisciplinary scholar. They are earning a Ph.D. in bioengineering at the University of Pittsburgh's Swanson School of Engineering. They earned an NIH T32 grant to study rare corneal disease and develop novel ocular drug delivery systems. Jorge is a National Science Foundation's Alliances for Graduate Education and the Professoriate (AGEP-KAT) scholar where they utilize their bilingualism (Spanish/English) in community engagement in Latinx healthcare and engineering education. Jorge utilizes their lived experience and scientific training to intersect health disparities topics in drug delivery education through the Center for the Integration of Research Teaching, and Learning (CIRTL).

Danesh Thirukumaran holds a B.S.E. in bioengineering from the University of Pittsburgh and an M.S. in Neuroscience from Georgetown University. He hopes to continue his career working at the center between biomedical innovation and clinical practice. He also hopes to develop underserved populations' understanding of the medical realm to assist in navigating the healthcare system.

Acknowledgements

The authors would like to share their deepest gratitude for the members of University of Pittsburgh Science Policy Group for organizing policy memo workshops. We would like to specifically thank Nathan Lampenfield (MPH candidate in Health Policy and Management) for providing feedback on initial draft revisions.

Disclaimer

The authors disclose that they have no conflict of interest in this form.