Using Population Descriptors in Genetics and Genomics Research

A New Framework for an Evolving Field

RECOMMENDATIONS FOR TRANSFORMING THE USE OF POPULATION DESCRIPTORS IN HUMAN GENETICS AND GENOMICS RESEARCH

Requisites for Sustained Change

RECOMMENDATION 1

Researchers should not use race as a proxy for human genetic variation. In particular, researchers should not assign genetic ancestry group labels to individuals or sets of individuals based on their race, whether self–identified or not.

RECOMMENDATION 2

When grouping people in studies of human genetic variation, researchers should avoid typological thinking, including the assumption and implication of hierarchy, homogeneity, distinct categories, or stability over time of the groups.

RECOMMENDATION 3

Researchers, as well as those who draw on their findings, should be attentive to the connotations and impacts of the terminology they use to label groups.

- As an example, the term *Caucasian* should not be used because it was originally coined to convey white supremacy,[1] and is often mistakenly interpreted today as a "scientific" term, thus erroneously conferring empirical legitimacy to the notion of a biological white[2] race.

- Another example of a term that should not be used is *black race* because it wrongly implies the existence of a discrete group of human beings, or race, who could be objectively identified as "black."

RECOMMENDATION 4

Researchers conducting human genetics studies should directly evaluate the environmental factors or exposures that are of potential relevance to their studies, rather than rely on population descriptors as proxies. If it is not possible to make these direct measurements and it is necessary to use population descriptors as proxies, researchers should explicitly identify how the descriptors are employed and explain why they are used and are relevant. Genetics and genomics researchers should collaborate with experts in the social sciences, epidemiology, environmental sciences, or other relevant disciplines to aid in these studies, whenever possible.

RECOMMENDATION 5

Researchers, especially those who collect new data or propose new courses of study for a data set, should work

[1] Johann Friedrich Blumenbach (1752–1840) named Europeans Caucasian because he felt the most beautiful skull in his collection came from the Caucasus region and was thus a fitting symbol for a superior race (Marks, 1995; Painter, 2010).

[2] The committee chose not to capitalize "black" and "white" throughout the report to recognize and emphasize that they do not signify biological or ethnic groups.

in ongoing partnerships with study participants and community experts to integrate the perspectives of the relevant communities and to inform the selection and use of population descriptors.

Guidance for the Selection and Use of Population Descriptors in Genetics and Genomics Research

RECOMMENDATION 6

Researchers should tailor their use of population descriptors to the type and purpose of the study, in alignment with the guiding principles, and explain how and why they used those descriptors. Where appropriate for the study objectives, researchers should consider using multiple descriptors for each study participant to improve clarity.

RECOMMENDATION 7

For each descriptor selected, labels should be applied consistently to all participants. For example, if ethnicity is the descriptor, all participants should be assigned an ethnicity label, rather than labeling some by race, others by geography, and yet others by ethnicity or nationality. If researchers choose to use multiple descriptors, each descriptor should be applied consistently across all individuals in that study.

RECOMMENDATION 8

Researchers should disclose the process by which they selected and assigned group labels and the rationale for any grouping of samples. Where new labels are developed for legacy samples, researchers should provide descriptions of new labels relative to old labels.

Implementation and Accountability

RECOMMENDATION 9

Funding agencies, research institutions, research journals, and professional societies should offer tools widely to their communities to facilitate the implementation of these recommendations; these tools should be publicly available, especially when they are supported by public funds. Such tools could include:

- educational modules for inclusion in human research protection training;[3]

- manuscript submission and review guidelines;

- grant submission and review criteria;

- training and education of trainees at all levels;

- opportunities for continuing education for researchers; and

- informatics tools, such as data structure standards for sharing labels and labeling procedures used within a study.

RECOMMENDATION 10

Research institutions and funding agencies should embed incentives for fostering interdisciplinary collaboration among researchers with different areas of expertise, including genetics and genomics, social sciences, epidemiology, and community-based research, to facilitate the inclusion of environmental measures and the engagement of diverse communities in genomics research. Funding agencies and research institutions should develop strategies to encourage and reward such collaborations.

RECOMMENDATION 11

Given the persistent need to address this dynamic, high-stakes component of genomics research, funders and research institutions should create new initiatives to advance the study and methods development of best practices for population descriptor usage in genetics and genomics research, including the public availability of resources.

RECOMMENDATION 12

Key partners, including funding agencies, research institutions, and scientific journals, should ensure that policies and procedures are aligned with these recommendations and invest in developing new strategies to support implementation when needed.

---

[3] Often called "human subjects" research training. See also https://www.hhs.gov/ohrp/education-and-outreach/human-research-protection-training/index.html.

RECOMMENDATION 13

Because the understanding of population descriptors in genomics research is continuously evolving, responsibility for periodic reevaluation of these recommendations should be overseen by effective, multidisciplinary advisory groups. The advisory groups could:

- periodically reevaluate established best practices on the use of descent-associated population descriptors to ensure they reflect the current state of the science and an ongoing commitment to ethical and empirical principles;

- advise funders and other interested parties on the use of population descriptors and their implementation;

- facilitate the coordination of international best practice sharing;

- provide a venue for input from the broader community, including research participants; and

- monitor and measure changes adopted by funders, researchers, journals, societies, and other relevant parties based on the uptake of best practices identified.

REFERENCES

Marks, J. 1995. *Human biodiversity: Genes, race, and history.* New York: Aldine de Gruyter.

Painter, N.I. 2010. *The history of white people.* New York: W.W. Norton & Company.

COMMITTEE ON THE USE OF RACE, ETHNICITY, AND ANCESTRY AS POPULATION DESCRIPTORS IN GENOMICS RESEARCH

ARAVINDA CHAKRAVARTI (*Co-Chair*), New York University Grossman School of Medicine; **CHARMAINE ROYAL** (*Co-Chair*), Duke University; **KATRINA ARMSTRONG,** Columbia University; **MICHAEL BAMSHAD,** University of Washington and Seattle Children's Hospital; **LUISA N. BORRELL,** City University of New York; **KATRINA CLAW,** University of Colorado Anschutz Medical Campus; **CLARENCE C. GRAVLEE,** University of Florida; **MARK D. HAYWARD,** The University of Texas at Austin; **RICK KITTLES,** Morehouse School of Medicine; **SANDRA SOO-JIN LEE,** Columbia University; **ANDRÉS MORENO-ESTRADA,** Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Mexico; **ANN MORNING,** New York University; **JOHN P. NOVEMBRE,** University of Chicago; **MOLLY PRZEWORSKI,** Columbia University; **DOROTHY E. ROBERTS,** University of Pennsylvania; **SARAH A. TISHKOFF,** University of Pennsylvania; **GENEVIEVE L. WOJCIK,** Johns Hopkins Bloomberg School of Public Health

STAFF **SARAH H. BEACHY,** Study Director; **SAMANTHA SCHUMM,** Associate Program Officer; **LEAH CAIRNS,** Study Co-Director (*until October 2022*); **KATHRYN ASALONE,** Associate Program Officer; **MEREDITH HACKMANN,** Associate Program Officer; **LYDIA TEFERRA,** Research Assistant; **APARNA CHERAN,** Senior Program Assistant (*from June 2022*); **MICHAEL K. ZIERLER,** Science Writer; **ANDREW M. POPE,** Senior Director, Board on Health and Sciences Policy (*until July 2022*); **CLARE STROUD,** Senior Director, Board on Health and Sciences Policy (*from July 2022*); **MALAY K. MAJMUNDAR,** Director, Committee on Population

Health and Medicine Division
Division of Behavioral and Social Sciences and Education

NATIONAL ACADEMIES *Sciences Engineering Medicine*