



HHS Public Access

Author manuscript

Biol Psychiatry. Author manuscript; available in PMC 2023 June 15.

Published in final edited form as:

Biol Psychiatry. 2022 June 15; 91(12): e51–e52. doi:10.1016/j.biopsych.2022.02.953.

From Evolutionary History to the Concepts of Race and Ancestry: Shifting our Perspective in Clinical Research

Angela M. Haeny, PhD¹, Renato Polimanti, PhD, MSc^{1,2}

^[1]Yale School of Medicine, Department of Psychiatry, New Haven, CT

^[2]Veteran Affairs CT Healthcare System, West Haven, CT

Biomedical studies continue to confound race with ancestry and treat the former as a biological indicator of risk as opposed to a social construct (1-4). Thus, research guidelines are needed to ensure researchers understand the difference between racial categories defined by historical events and ancestry reflecting the continuum of human diversity (5).

Racism in the U.S. is a social caste system used to group people based on shared physical or social characteristics that privileges White people and is ever-evolving (6). According to the U.S. Census Bureau, race can be classified into five major categories: Black/African American, American Indian/Alaska Native, Asian, Pacific Islander/Native Hawaiian, and White/European American (8). Further, the U.S. Census Bureau recognizes only one ethnicity: Hispanic/Latine,ⁱ though Latine was considered a racial group in the past (9). The Bureau offers an approach for re-categorizing multiracial people into a single category in the following order: Black, Indigenous, Native Hawaiian/Pacific Islander, or Asian.ⁱⁱ That is, if a person identifies as Black and Indigenous, they would be categorized as Black. If someone identifies as White and Asian, they would be categorized as Asian. A person would only be considered White if they do not endorse any other racial category.

There are multiple issues with these categorizations. First, the approach to recategorize multiracial people is consistent with the “one-drop rule” that anyone with a single relative who is Black would be considered Black (10), which stems from racism. Further, ethnicity represents shared food, music, language, history, and cultural traditions that distinguish groups of people (11). Thus, everyone holds an ethnic identity (e.g., Nigerian, Puerto Rican, Italian, Filipino, German, Chinese). For Indigenous peoples, ethnic identity may be better described as tribal identity (e.g., Karuk Tribe, Penobscot Nation, Shawnee Tribe). These racial/ethnic categorizations do not apply across countries and they are fluid. How a person is racially categorized may not represent how they self-identify. Notably, although some White people may self-identify as Caucasian, this is an archaic term referring only to people

Correspondence for this article can be addressed to Angela M. Haeny, Ph.D. at 34 Park St., SAC-210, New Haven, CT 06519, phone: 475-434-0629, angela.haeny@yale.edu.

Disclosures: None

ⁱWe used the term Latine as a gender-neutral descriptor that is more widely accepted and used in the Spanish-speaking community than the term Latinx.

ⁱⁱSee NESARC-III Data Notes: [https://www.niaaa.nih.gov/sites/default/files/NESARC-III Data Notesfinal_12_1_14.pdf](https://www.niaaa.nih.gov/sites/default/files/NESARC-III%20Data%20Notesfinal_12_1_14.pdf)

from the area between the Black and the Caspian seas and should not continue to be used (11,12).

The diversity of human populations represents a continuum reflecting evolutionary history. *Homo sapiens* originated in Africa around 200,000-300,000 years ago. Early humans migrated across landmasses and created bottlenecks where human genetic variation moved to different continents. This results in populations with the same continental origin (ancestry) with higher molecular similarity to each other than other continental groups. The difference between ancestry groups accounts for ~9% of the human genetic variation while the remaining represents differences among individuals within the same population (~89%) and among populations within the same continental group (~2%) (13). Similar variation distributions are also observed with other molecular domains, including epigenome, transcriptome, proteome, metabolome, and microbiome (14-16).

Ancestry categories reflect a small portion of the molecular diversity of populations that is mostly accounted by inter-individual variability. The assignment of ancestry categories becomes more complicated considering population groups affected by recent admixture events such as those that occurred on the North American continent in the last 500 years. Individuals self-identified as Black or Latine present genomic segments inherited from different ancestral groups with a variable admixture degree. Black people inherited genomic segments from African and European ancestries with variation ranging from 10% to 90% (17). Latine individuals present genomic segments from three ancestries, African (5% to 15%), European (45% to 80%), and Native American (10% to 50%) (18).

Thus, molecular studies should not use racial categories because they do not adequately reflect biological diversity. For example, terms related to skin color categories reflect very poorly the spectrum of skin pigmentation, which is only one of the many products of human adaptation. For these reasons, we strongly recommend investigators focused on human biology to use the geographical origin (ancestry) to refer to the diversity of their participants. To adequately model the patterns observed within ancestry groups and among recently admixed populations, statistical methods should be used that account for human molecular variation to avoid results biased by the unaccounted population structure of the cohorts investigated.

Racially minoritized people are more likely to experience inequitable access to opportunities (e.g., quality education), goods (e.g., nutritious foods), and services (e.g., quality healthcare) (6,19), which contribute to racial disparities in health (6,20). Instead of interpreting race as a risk factor for health outcomes, race should be understood as a proxy for structural inequities. The structural factors that contribute to differences across race categories should be incorporated in studies aiming to address inequities.

We recommend the use of race is limited to ascertaining racial disparities and ensuring representation in studies. We suggest researchers specify how race was operationalized (e.g., an open-ended self-report question or a list of options--specify the list if used). For child studies, specify whether the child or caregiver reported the child's race given these reports may differ. Longitudinal studies might re-assess race at each time point to capture the

fluidity of race given that racial identity development theories (21-23) suggest racial identity can change over time particularly among multiracial people.

Researchers should explicate how race and ethnicity were categorized in their analyses and note the limitations. Consider the pros and cons of combining everyone who identifies as Latine and asking a follow-up question to determine which group multiracial people identify with most vs. using the algorithm vs. maintaining a multiracial subgroup. Specifying how race and ethnicity were assessed and analyzed will inform comparisons of findings across studies.

In summary, race should be used to characterize the sample and identify disparities whereas biological studies should use genetically-informed ancestry classifications (Figure 1). A greater understanding of the implications of race- and ancestry-related terms will better inform biomedical research questions utilizing these constructs. Appropriately using race and ancestry in research will further the science and contribute to ending scientific racism.

Acknowledgements:

Supported by R25DA035163, K23AA028515, L30DA049246, R33DA047527, R21DC018098 from the National Institutes of Health.

References

1. Maglo KN, Mersha TB, Martin LJ. Population genomics and the statistical values of race: An interdisciplinary perspective on the biological classification of human populations and implications for clinical genetic epidemiological research. *Front Genet* [Internet]. 2016 Feb 17 [cited 2020 Oct 29];7:1–13. Available from: www.frontiersin.org
2. Bamshad M, Wooding S, Salisbury BA, Stephens JC. Deconstructing the relationship between genetics and race. *Nat Rev Genet* [Internet]. 2004 Aug [cited 2020 Oct 7];5(8):598–609. Available from: www.nature.com/reviews/genetics
3. Mersha TB, Abebe T. Self-reported race/ethnicity in the age of genomic research: its potential impact on understanding health disparities. *Hum Genomics* [Internet]. 2015 Dec 7 [cited 2019 Feb 8];9(1):1. Available from: <https://humgenomics.biomedcentral.com/articles/10.1186/s40246-014-0023-x>
4. Frank R. What to make of it? The (Re)emergence of a biological conceptualization of race in health disparities research. *Soc Sci Med*. 2007 May 1;64(10):1977–83. [PubMed: 17368894]
5. Boyd RW, Lindo EG, Weeks LD, McLemore MR. On Racism: A New Standard For Publishing On Racial Health Inequities [Internet]. *Health Affairs Blog*. 2020 [cited 2020 Dec 10]. Available from: <https://www.healthaffairs.org/doi/10.1377/hblog20200630.939347/full/>
6. Haeny AM, Holmes SC, Williams MT. The Need for Shared Nomenclature on Racism and Related Terminology in Psychology. *Perspect Psychol Sci*. 2021;
7. McIntosh P. White privilege: Unpacking the invisible knapsack. 1988;191–6.
8. Jones NA, Bullock J. The Two or More Races Population: 2010 2010 Census Briefs [Internet]. 2012 [cited 2021 Mar 20]. Available from: <http://2010.census.gov/2010census>
9. McKenney NR, Bennett CE. Issues regarding data on race and ethnicity: the Census Bureau experience. *Public Health Rep* [Internet]. 1994 [cited 2021 Dec 21];109(1):16. Available from: [/pmc/articles/PMC1402238/?report=abstract](https://pubmed.ncbi.nlm.nih.gov/1402238/)
10. Hickman CB. The Devil and the One Drop Rule: Racial Categories, African Americans, and the U.S. Census. *Mich Law Rev*. 1997 Mar;95(5):1161.
11. Wadsworth LP, Morgan LP, Hayes-skelton SA, Roemer L, Suyemoto KL. Ways to Boost Your Research Rigor Through Increasing Your Cultural Competence (Part 1 of 2). *Behav Ther*. 2016;39(3):76–82.

12. Flanagin A, Frey T, Christiansen SL, Committee AM of S. Updated Guidance on the Reporting of Race and Ethnicity in Medical and Science Journals. *JAMA* [Internet]. 2021 Aug 17 [cited 2021 Aug 20];326(7):621–7. Available from: <https://jamanetwork.com/journals/jama/fullarticle/2783090>
13. Li JZ, Absher DM, Tang H, Southwick AM, Casto AM, Ramachandran S, et al. Worldwide human relationships inferred from genome-wide patterns of variation. *Science* (80-) [Internet]. 2008 Feb 22 [cited 2021 Feb 3];319(5866):1100–4. Available from: <https://pubmed.ncbi.nlm.nih.gov/18292342/>
14. Carja O, MacIsaac JL, Mah SM, Henn BM, Kobor MS, Feldman MW, et al. Worldwide patterns of human epigenetic variation. *Nat Ecol Evol* [Internet]. 2017 Oct 1 [cited 2021 Mar 20];1(10):1577–83. Available from: <https://pubmed.ncbi.nlm.nih.gov/29185505/>
15. Hughes DA, Kircher M, He Z, Guo S, Fairbrother GL, Moreno CS, et al. Evaluating intra- and inter-individual variation in the human placental transcriptome. *Genome Biol* [Internet]. 2015 Mar 19 [cited 2021 Mar 20];16(1). Available from: <https://pubmed.ncbi.nlm.nih.gov/25887593/>
16. Sjaarda J, Gerstein HC, Kutalik Z, Mohammadi-Shemirani P, Pigeyre M, Hess S, et al. Influence of Genetic Ancestry on Human Serum Proteome. *Am J Hum Genet* [Internet]. 2020 Mar 5 [cited 2021 Mar 20];106(3):303–14. Available from: <https://pubmed.ncbi.nlm.nih.gov/32059761/>
17. Bryc K, Durand EY, Macpherson JM, Reich D, Mountain JL. The genetic ancestry of african americans, latinos, and european Americans across the United States. *Am J Hum Genet* [Internet]. 2015 Jan 8 [cited 2021 Mar 20];96(1):37–53. Available from: <https://pubmed.ncbi.nlm.nih.gov/25529636/>
18. Gravel S, Zakharia F, Moreno-Estrada A, Byrnes JK, Muzzio M, Rodriguez-Flores JL, et al. Reconstructing Native American Migrations from Whole-Genome and Whole-Exome Data. *PLoS Genet* [Internet]. 2013 [cited 2021 Mar 20];9(12). Available from: <https://pubmed.ncbi.nlm.nih.gov/24385924/>
19. Jones CP. Levels of racism: A theoretic framework and a gardener’s tale. *Am J Public Health* [Internet]. 2000 [cited 2019 Jul 12];90(8):1212–5. Available from: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1446334/pdf/10936998.pdf>
20. Yearby R. Racial Disparities in Health Status and Access to Healthcare: The Continuation of Inequality in the United States Due to Structural Racism. *Am J Econ Sociol* [Internet]. 2018 May 1 [cited 2020 Sep 2];77(3–4):1113–52. Available from: <http://doi.wiley.com/10.1111/ajes.12230>
21. Miville ML, Constantine MG, Baysden MF, So-lloyd G. Chameleon Changes : An Exploration of Racial Identity Themes of Multiracial People. 2005;52(4):507–16.
22. Ponterotto JG, Gretchen D, Utsey SO, Stracuzzi T, Saya R. The Multigroup Ethnic Identity Measure (MEIM): Psychometric Review and Further Validity Testing. *Educ Psychol Meas* [Internet]. 2003 Jun 2 [cited 2019 Sep 14];63(3):502–15. Available from: <http://journals.sagepub.com/doi/10.1177/0013164403063003010>
23. Helms J. Black and White racial identity: Theory, research, and practice. Greenwood Press; 1990.

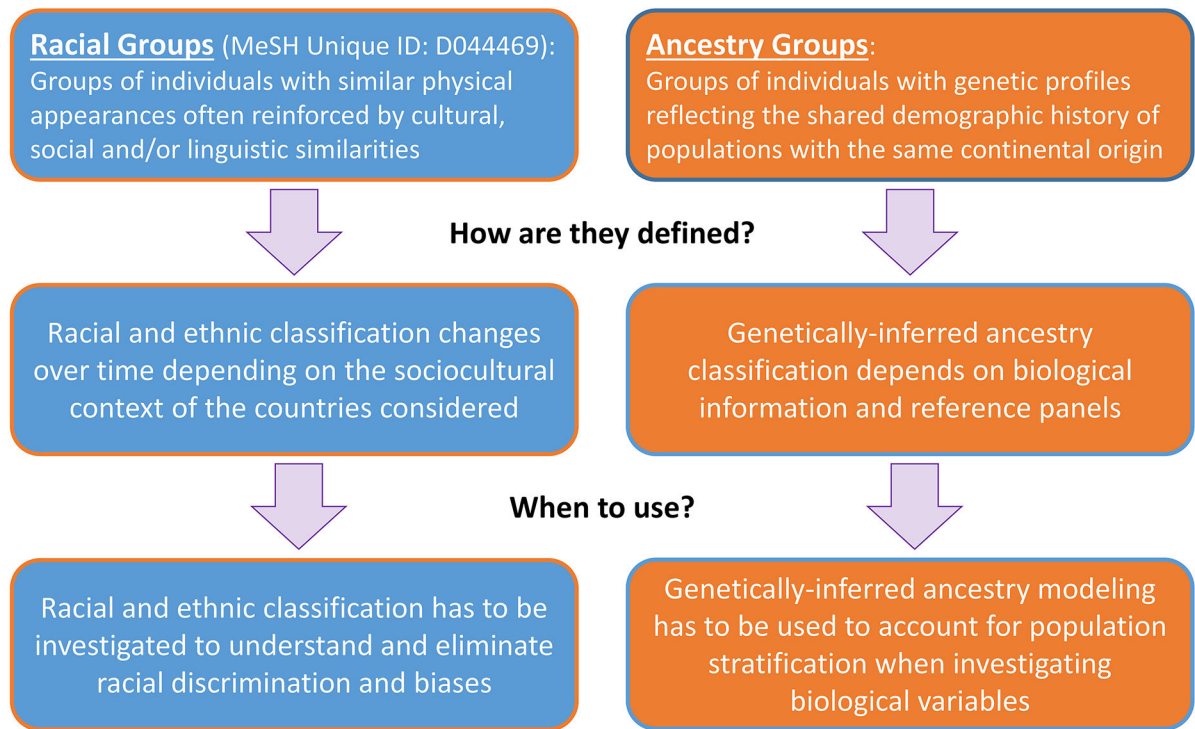


Figure 1.
Differentiating between race and ancestry. MeSH, Medical Subject Headings.