Racialising genetic risk: assumptions, realities, and recommendations

Jessica P Cerdeña, Vanessa Grubbs, Amy L Non

Lancet 2022; 400: 2147–54

Yale School of Medicine, Yale University, New Haven, CT, USA (J P Cerdeña PhD); **Department of Anthropology, University of Connecticut, Storrs, CT, USA** (J P Cerdeña); **Department of Ambulatory and Preventive Medicine, Alameda Health System, Oakland, CA, USA** (V Grubbs MD); **Department of Anthropology, University of California, San Diego, CA, USA** (A L Non PhD)

Correspondence to: Assoc Prof Amy Non, Department of Anthropology, University of California, San Diego, CA 92093, USA **alnon@ucsd.edu**

Introduction

Scientists and clinicians wield the immense power of defining reality and producing facts.[1] Although no person can truly claim objectivity, scholars enjoy the authority of expertise and cultural capital, a combination that provides them with near-deified credence. When doctors and scientists, such as Carolus Linnaeus (a Swedish naturalist) and Johan Blumenbach (a German doctor and anthropologist), attempted to taxonomise the world during the Enlightenment era by consolidating their observations using the travel logs of European colonisers,[2] their proposals for biologically distinct human varieties, or hierarchised races, became entrenched as knowledge and legitimised ongoing practices of imperialism.[3] Racist logic preceded the invention of biological, racial types.

Although human genomic data show continual rather than clustered genetic variation, and contemporary scientists and clinicians have stopped using such explicit scientific racism, harmful race-based practices persist in biomedical and clinical research, often using putatively precise terms, such as ethnicity and ancestry. For at least 30 years, scholars have debated the value and meaning of these terms in biomedicine,[4–6] yet researchers and clinicians continue to misunderstand and misuse them. In this Viewpoint, we assess common and problematic assumptions in genomics research and clinical practice and provide recommendations for researchers, clinicians, funders, and academic journals in response to frequent assumptions that occur during study design, data analysis, and peer review. Our aim is to promote race-conscious medicine and increase theoretical and analytical rigour.[7]

Race and ethnicity

There is a flawed assumption that race and ethnicity can be used interchangeably in genetic and medical studies and are both markers of complex disease risk. The reality is that both race and ethnicity are sociopolitical terms, and neither term describes fixed biological or genetic characteristics of a population

Although many clinicians and researchers continue to use race as a biological classification, scholarly consensus considers race a sociopolitical invention used to hierarchise humans according to the aims of the groups in power.[8] Ethnicity (although rarely defined and often interchanged with race) commonly refers to cultural, socioeconomic, religious, linguistic, and political qualities of groups that establish cohesion and order through membership, rather than their population genetics.[8] Similarly to race, ethnicity is socially constructed, with dynamic boundaries that change depending on places, times, and contexts. For instance, census categories for Black individuals in the USA and Māori people in New Zealand previously specified blood quanta (eg, "Mulatto" or "Quadroon" and "half-caste").[9–11] The US census now distinguishes race (eg, White or Black) from ethnicity, yet provides only one ethnicity category: Hispanic or Latino. The majority of Latinx people in the USA, however, consider the terms Hispanic or Latino to be either a race or both a race and ethnicity.[12] Similarly, Jewish identity, although functionally a religion, has been constructed as both a race and an ethnicity.[13] In the wake of Nazism, European people avoided the term race entirely, favouring use of the term ethnicity. Instead, however, scholars have used terms such as "culture", "migration background", or "country of origin" when describing minoritised and immigrant groups.[14,15] Self-identified race and ethnicity often differ from assigned race and ethnicity, highlighting the limitations of their uses in biomedicine.[16] The combination of racial and ethnic terms in biomedical research reflects their common interchange by the public, which uses both terms to classify groups of people from non-dominant social strata.[17]

The use of race or ethnicity analogously in research and clinical care derives from recommendations to collect and report data for these sociopolitically established groups to assess disease risk without appropriate guidance on how to analyse and interpret these data. For example, federal guidelines in the USA recommend collection of data on minoritised populations to record "cultural and behavioral attitudes, beliefs, lifestyle patterns, diet, environmental living conditions" and guidelines in the European Union recommend collection of data on minoritised populations to address "discrimination".[18–21] However, these guidelines do not provide specific instructions as to how to use these data responsibly, treating race and ethnic origin as risk markers rather than as risk factors for disease.

The assumption that race and ethnicity are markers of disease risk in clinical practice comes from historical efforts to pathologise minoritised populations based on laboratory findings outside of the range typically found in White populations (figure). For example, benign ethnic neutropenia describes a condition of so-called defective granulocyte release from otherwise typical bone marrow found occasionally in some populations (eg, African Americans, Yemenite Jewish people, people from Ethiopia, and some Arab people). The ethnicity-based label of this phenotype reinforces the assumption that ethnicity is causally related to disease, and reifies the idea that any phenotype that is different from the

	Flawed assumptions	Realities	Problematic examples	Evidence-based alternatives
Race-based and ethnicity-based diagnosis and medication doses	Race can accurately approximate genetic risk or medication dose	Both race and ethnicity are sociopolitical inventions, and neither describes fixed biological or genetic characteristics of a population	Pathologising healthy states (eg, benign ethnic neutropenia)	Clinicians interpret normal findings based on clinical presentation rather than comparison with race or ethnicity standard
				Researchers and journal editors continue to collect data on race and ethnicity over time and ensure definitions and rationale for including race or ethnicity in analyses are clearly defined
			The kidney donor risk index uses Black race rather than *APOL1* genotype to estimate donor kidney viability	Clinicians test for specific risk alleles according to clinical indication regardless of patient race
			Warfarin recommendations are for low doses in Asian patients, intermediate in White patients, and high levels in Black patients	Clinicians adjust medication doses according to therapeutic goals, regardless of self-reported race or ancestry
Race-based and ancestry-based clinical algorithms and genetic screening	Self-reported ancestry accurately proxies increased risk for pathogenic genotypes	Ancestry, race, and ethnicity are not accurate markers of allele-carrier risk or risk of complex, non-mendelian diseases (eg, diseases resulting from complex polygenic and environmental exposures and interactions)	Race (and potentially genetic ancestry) is incorporated into spirometry reference equations so that Black patients have to show more lung damage than patients of other races to be classified as having abnormal results	Researchers and clinicians assess relevant social and environmental exposures, measure genetic risk factors directly, and do not include race or genetic ancestry in reference equations
			Genetic screening panels for alleles common in Ashkenazi Jewish people arenot tested in other groups of people that also have these alleles at a high frequency (eg, French Canadian and Creole groups)	Clinicians use familial risk or universal screening models to guide testing and counselling to avoid over-screening some populations and under-screening other populations
Structural racism in clinical research	Inclusion of a race variable in research is a sufficient proxy for structural racism	Race does not accurately approximate individual experiences of structural racism	A study concludes that genetic differences explain racial disparities after they have controlled for race	Funders prioritise empirical research of sociostructural contributors to health inequities Researchers directly measure the health effects of racist and reparative policies rather than presuming that individuals of the same racial group experience racialisation in the same way Journal editors ensure race or social experience are not used as a broad proxy of structural racism or racist policy
Race-based genomic analysis	Because race is a fundamental risk factor for complex diseases, analyses should be stratified by race a priori	Stratifying analyses by race, ethnicity, or continental ancestry introduces bias that reinforces essentialist notions of biological race	Analysis of genome-wide association studies within racial groups before testing associations within the global sample	Researchers stratify genome-wide association study populations by disease characteristics (eg, progressing rapidly *vs* progressing slowly) rather than a priori by race or ethnicity
	Standard genomics panels adequately capture global genetic diversity	Genomic panels designed primarily for European people lead to ascertainment bias when assessing people who are not European	Use of standard microarrays designed without inclusion of diverse populations	Researchers include whole-exome or whole-genome analyses (or comprehensive microarrays) to represent rare variants around the world

Figure: **Clinical and research examples of flawed assumptions in clinical domains with evidence-based alternatives**

phenotype typically seen in White populations is a disease.[22] In another example, the Kidney Donor Risk Index downgrades the viability of kidneys from Black donors due to the assumption of the *APOL1* double-variant genotype, a condition only found in 13% of the Black population in the USA, and also in up to 5% of many Latinx and Native American populations, who are rarely screened.[23,24] Similarly, prescribers and pharmacists consider race and ethnicity when deciding on warfarin dose, despite the fact that the genetic markers that influence pharmacokinetic enzymes are not found exclusively in any specific race and have mostly been studied in European people and White Americans.[25]

We recommend that researchers continually collect data on race and ethnicity, specify whether they are self-identified or assigned, and clarify the specific rationale for their use in marking the sociostructural determinants of health inequities, rather than as proxies for potentially pathogenic alleles.[19–29]

Genetic ancestry

There is a flawed assumption that genetic ancestry is more precise than race or ethnicity, so we should use it in clinical algorithms and biomedical research.[30] The reality is that ancestry has many of the same issues as race and ethnicity, and is not useful as a marker of complex, non-mendelian diseases (ie, diseases resulting from complex polygenic and environmental exposures and interactions)

Ancestry is a vague term. Although ancestry typically refers to the geographical regions in which the biological ancestors of an individual lived (eg, Balkan, referring to the peninsula), scholars and members of the public might use it in a combination of geopolitical (eg, Vietnamese, referring to the nation-state established in 1976), cultural, or linguistic senses (eg, South American cultures and dialects).[8] Ancestry also depends on timescale; a medical family history refers to the health of the parents and grandparents of an individual, whereas ancestry beyond approximately 100 000 years traces all humans back to Africa. Individuals often have ancestry from multiple regions, and knowledge of those regions can be circumscribed (eg, southern Liberia) or broad (eg, all of Africa). These varied ways in which ancestry is defined—or self-determined—make it a problematic proxy for predicting who might have pathogenic alleles.[31,32]

Contemporary assessments of genetic ancestry, or ancestry estimates inferred from informative markers in the genome, come from living reference populations classified by race, ethnicity, or nationality.[33] Geneticists treat these living reference populations as ancestors, reflecting ingrained assumptions regarding racialised typologies and their persistence for an extended period of time.[34] Early iterations of DNA testing grouped ancestors by continental geography. As reference samples and genomic coverage increased, precision according to current geopolitical boundaries also increased. Some scholars have argued that genetic ancestry is a more precise and scientific alternative to race,[30] but most ancestry studies use continental regions (eg, European ancestry), which proxy racial categories and reflect imposed discontinuities that do not necessarily show the gradation of human genetic variation.[35,36] Because of the diversity in Africa in particular, an estimate of African genetic ancestry—or even West African ancestry—is not predictive of any specific pathogenic allele.

Furthermore, ancestry is often accompanied by cultural and psychosocial experiences shared by a particular racialised group, such as discrimination. Thus, a genetic ancestry estimate on its own, even when statistically linked to a disease, is not sufficient evidence of a genetic contribution to a particular non-mendelian disease or racial phenotype. Ancestry does not have the objectivity and precision scholars usually assign it and the term instead obscures the racial essentialism (ie, the view that people from different racial categories have fundamentally different biological properties) that is integral to its estimation.

There are some clinicians and researchers who argue that race might proxy ancestry in clinical algorithms (figure), but racial self-identification varies widely with assessments of ancestry. The 23andMe research team found that although the mean proportion of African ancestry in Black Americans is estimated to be 73%, those with at least 28% African ancestry tend to self-identify as African American, and about 2% of Black Americans have less than 2% African ancestry.[37] Furthermore, the percentage of European contribution to African American genetic samples across the USA has been shown to vary from 3·5% in the isolated Gullah-speaking Sea Islanders from South Carolina to 35% in Seattle.[38] Furthermore, in a large empirical study, self-reported ethnicity was shown to be a flawed indicator of carrier status for genetic markers of commonly screened diseases.[39] Specifically, 9% of individuals had more than half of their genetic ancestry from a population inconsistent with their self-reported ethnicity, and for seven of the 16 examined conditions, most people with carrier status were from a population other than the one included in the current screening guideline.[39] The practice of inferring disease risk from race, ethnicity, or ancestry can contribute to health-care inequities by encouraging racial stereotyping, stratified care, and misclassification of disease risk.[7]

We recommend that clinicians should evaluate disease risk based on clinical history, hypothesis-based and allele-specific genetic testing, and environmental exposures.[40] They should avoid using ancestry as a risk factor in clinical algorithms or calculations, as it does not accurately proxy genetic or social risk of disease (figure). Researchers should provide definitions of ancestry terms in clinical research and clear, competing hypotheses to justify the role of ancestry in study design and analyses.

Race as a proxy for genetics

There is a flawed assumption that race is a useful proxy for genetics because there is an increased prevalence of sickle cell disease in Black people and an increased prevalence of Tay-Sachs disease in Ashkenazi Jewish people. The reality is that although some diseases that come from single gene mutations (ie, mendelian diseases) are more common in geographical regions with histories of genetic bottlenecks (eg, a reduction in genetic diversity due to a substantial reduction in population size) or historical selective pressures (eg, evolutionary forces that favour reproduction of specific phenotypes over others in particular environmental conditions) than in the rest of the world, these diseases are not exclusive to specific racial or ethnic groups and are not relevant for most medical conditions

See Online for appendix

Some mendelian diseases that occur at high frequencies in particular regional groups, such as Tay-Sachs disease, thalassemia, and sickle cell disease, correspond to geographical areas that do not exclusively pattern by continent or race.[41,42] In other words, these diseases are not found exclusively in specific continents or racial groups. However, the majority of human global variation occurs at neutral loci and is due to random drift, serial founder events, and restricted gene flow imposed by distance and natural barriers, such as oceans and mountain ranges.[43] This evolutionary history has led to the development of a human genetic structure in which differences between individuals within a population explain 93–95% of genetic variation and differences between continental groups explain 3–5% of genetic variation.[44] Ancestry-based selection events in response to environmental pressures—including the adaptive evolution that resulted in increased prevalence of the allele that caused sickle cell disease in malaria-exposed populations—are rare. Claims that this directional selection in ancestral populations occurred commonly, and therefore explains racial differences in complex diseases, are misguided and often the result of oversimplified understandings of human evolution.

Many studies use genome-wide measures to calculate an approximate estimate of African or Indigenous ancestry, and test this ancestry estimate as a direct risk factor for disease.[45–47] If the genetic ancestry estimate is associated with disease, researchers conclude that a genetic difference between races is affecting disease disparities. Although phenotypic traits associated with race, including skin colour and hair texture, have a genetic basis,[48,49] most of these traits show continuous variation, influenced by dozens or even hundreds of alleles, and are thus polygenic.[50] That some genes regulate racialised features does not mean those same genes contribute to or are linked to genes that increase disease risk. Human traits are non-concordant; genes controlling different traits are not necessarily inherited together. For example, even if a shared evolutionary history has contributed to more people of African descent carrying some alleles (eg, for dark skin) than other groups of people, the high genetic variation inside and outside of Africa, thousands of years of gene flow with other groups, and scarce evidence for truly race-specific pathogenic alleles—particularly in complex diseases—mean the presence of dark skin alleles cannot predict alleles for specific diseases (appendix p 1).

Genetic ancestry estimates are always conflated with other—usually unmeasured—sociocultural or environmental factors, making it impossible to disentangle their effects on disease. For example, Tang and colleagues,[51] in a 2006 case-control association study, claimed a non-significant positive association between African genetic ancestry and high blood pressure in Black Americans. However, when Non and colleagues[52] reanalysed these same data in 2012, but accounted for a basic measure of social experience (ie, years of education), the genetic ancestry effect was reduced, showing that environmental exposures that are linked to race confounded the originally observed effects of ancestry. This confounding of genetic and environmental effects that can lead to spurious associations with genetic ancestry could be a more common occurrence than previously thought, as few studies of racial disparities include social data.[53] However, when studies do find persistent ancestry effects after adjustment for usually basic social or environmental data, they might still be affected by residual confounding because of superficial measures of a complex environment across the life course.[54]

We recommend that clinicians should test patients with symptoms or family histories that are suggestive of a genetic disorder for the corresponding genetic markers, regardless of their phenotypic appearance or self-reported race, ethnicity, or ancestry (figure). Researchers should include measures of racism and the social environment (eg, everyday discrimination, educational attainment, and stress exposures) in assessments of racialised health disparities to avoid essentialising racial differences (ie, characterising them as fundamentally distinct) by not measuring confounding social factors.

Race as surrogate measure for racism

There is a flawed assumption that, in genetic analyses, controlling for race as a variable can account for the contribution of structural racism to disease. The reality is that because of varied experiences of racialisation and enforcement of structurally racist policies, race does not accurately approximate individual experiences of structural racism

Structural or systemic racism broadly refers to "the totality of ways in which societies foster [racial] discrimination, via mutually reinforcing [inequitable] systems".[55] It underlies all dimensions of society, including historical, cultural, institutional, and interpersonal dynamics. Racist policies—including discriminatory mortgage lending, law and immigration enforcement, and health care—reinforce structural racism. Structural racism is not represented by

measures of racism at the individual level (eg, self-perceived discrimination), but instead represents systemic forces that influence health at a population level.

Some researchers consider race to be a useful surrogate for structural racism and support its continued use in research.[56] Structural racism, evidenced through segregation and inequities in employment and education, can also contribute to inequities in toxic environmental exposures (eg, air pollution or lead in water), health-care access, and health-care quality, which can increase both risk and progression of many diseases (eg, cancer and kidney disease).[57] Lifetime experiences of adversity and oppression can induce epigenetic modifications in genes involved in multiple physiologic systems.[58,59] Research has provided an increasing evidence base for the mechanisms by which structural racism mediates health inequities.[60] These mechanisms include a pathway through which increased stress can activate the hypothalamic-pituitary-adrenal axis, which can increase vascular tension and impair regulation of the inflammatory response.[58] Increased vascular tension contributes to hypertension and sleep disorders, and high amounts of inflammation can increase risk for cardiovascular disease.[57] Racial differences in COVID-19 outcomes in the USA and the UK are also influenced by structural inequalities influencing exposures and restricting access to equitable care,[61,62] but are often assumed to be genetic in cause.[63]

However, because not all members of the same racialised group have the same experiences of racialisation—for example, due to phenotypic differences or differences in social environment—the use of race as a measure of racism is inappropriate.[64] Although many in the global medical research community have emphasised the importance of directly measuring structural factors,[65,66] most genetic studies of complex disease still do not regularly include basic sociopolitical variables, such as income, health insurance, or nativity. The studies that do include measures of the social environment often use race-neutral or individualistic variables, such as educational attainment. In a 2018 systematic review of concepts of structural racism in the 50 most high-impact public health journals, only four (16%) of 25 research articles considered structural racism to be a main concept.[67] A 2021 systematic review showed that, despite an increase in mentions of racism in clinical and public health literature in the past 30 years, more than 90–96% of these publications were commentaries, viewpoints, or letters rather than empirical studies.[68]

Measuring structural racism is a substantial, complex, and underfunded research task. Research based in the USA has improved the operationalisation of composite assessments of structural racism at the state and local levels.[69–73] For example, a study of state-level structural racism that incorporated judicial, education, and employment inequities found that increased structural racism was associated with myocardial infarction in Black Americans.[74] Future studies should consider the use of latent class analyses to construct multidimensional models of structural racism and analyses of racist or reparative policies to evaluate associations between structural racism and health outcomes. Although current funding priorities favour genomics research and precision medicine globally,[75] funding equity is necessary to develop comprehensive, valid, replicable, and theoretically sound ways to operationalise structural racism.

We recommend that researchers should measure the effects of racist policies rather than controlling for race. Funders should prioritise empirical research assessing the sociostructural contributors to health inequities.

Race stratification in genetic studies

There is a flawed assumption that when designing a genome-wide association study (GWAS) of disease, stratifying a standard genome panel sample a priori by race or ethnicity helps to detect meaningful genetic differences between populations. The reality is that stratifying analyses by race, ethnicity, or continental ancestry can introduce bias that reinforces essentialist notions of biological race (ie, a belief in fundamentally and intrinsically distinct biological groups). Even after controlling for population stratification, biased samples and standard genomic panels lead to ascertainment bias when assessing non-European populations

GWAS researchers consistently assume that race is a fundamental risk factor for disease at the beginning of study design, and often separately analyse genomic data within each racial group before testing for associations within the global sample.[47] Thus, stratified study designs can be biased in favour of race-specific effects when shared variants across the entire dataset might be more relevant to many common diseases. When analysing populations of various ancestries, population stratification should be controlled for to avoid spurious associations. Principal component analysis of genomic data can characterise geographical and genetic gradients that come from environmental and reproductive isolation and genetic drift in human history. These analyses remove the need for reference populations and reduce confounding in admixed populations, in which purported ancestral differences might be associated with a phenotype in the absence of a causal genetic pathway.[76,77]

We also emphasise that commonly used genome panels do not represent the broad genetic diversity across geographical space because of the limitations of genetic testing in low-resource settings. Additionally, 78% of GWAS participants are European, greatly restricting the scope of human genetic diversity represented, especially considering that the majority of this diversity exists in Africa.[78] Although whole-genome sequencing (WGS) can overcome ascertainment bias, it has not been affordable for most studies to date. In the absence of WGS, genetic researchers could be missing important rare variants and their analyses could be confounded by differences in patterns of linkage

disequilibrium between causal genetic variants and tagging single nucleotide polymorphisms that vary across geographical regions. Including globally diverse geographical populations does not mean that these groups represent genetically discrete units; large samples from locations throughout the world is the best way to represent the complex, gradated nature of genetic variation in humans.[79] Furthermore, different physical and social environments that are experienced by members of various racial or ethnic groups might influence gene–environment interactions, which could alter disease associations across various populations.

We recommend that researchers should stratify analyses of GWAS study populations by disease characteristics (eg, progressing rapidly *vs* progressing slowly) rather than a priori by race, ethnicity, or continental ancestry. Population stratification should be adjusted via principal component analysis or similar approaches, such as multidimensional scaling or mixed models. If possible, they should include whole-exome or WGS in genomic studies, or comprehensive microarrays designed for diverse samples, to ensure generalisability, confidence in results, and equity.[80]

Conclusion

Ongoing discussions regarding the use of the terms race, ethnicity, and ancestry in biomedicine reflect the persistent misunderstanding of the definitions of these terms and their true associations with pathogenic genetic variants. In fact, structural racism—or the way policies established through legacies of slavery and European colonisation encourage ongoing racialised oppression in all parts of society—produces environmental but non-uniform distribution of inequities. These inequities become embodied in the individual, influencing hormonal activity, epigenetics, and gene expression alongside reduced health-care access and health-care quality to produce health inequity. Misuse and conflation of the terms race, ethnicity, and ancestry are restricting progress in understanding health disparities. These terms are all socially constructed concepts with no fixed biological meaning, although variable experiences of racialisation can produce health inequities via physiological responses to racial oppression. Reliance on race, ethnicity, and ancestry as surrogates for pathogenic alleles of complex disease in biomedical research and clinical practice risks identification of spurious associations, misdiagnosis of disease risks, and missed opportunities to fund research that can potentially identify the true causes of health disparities. Practicing race-conscious medicine by emphasising racism rather than poorly defined sociopolitical categories can reveal the underlying causes of racialised health inequities and the appropriate targets of intervention.

Contributors

VG, JPC, and ALN conceptualised this Viewpoint. JPC and ALN wrote the original draft and VG, JPC, and ALN reviewed and edited the manuscript.

Declaration of interests

ALN has received funding from the University of California, San Diego Altman Clinical and Translational Research Institute Pilot Award and the Mother-Milk-Infant Center of Research Excellence Pilot Award. VG has been paid honoraria by Kaiser Permanente, Akebia Pharmaceuticals, and the University of Illinois. JPC has received funding from the National Institutes of Health Medical Scientist Training Program Grant (T32GM136651) and the Robert Wood Johnson Foundation Health Policy Research Scholars Program.

Acknowledgments

We thank Clarence C Gravlee and Naomi N Bishop for providing comments on an early draft of the manuscript.

References

1 Latour B. Science in action: how to follow scientists and engineers through society. Cambridge, MA: Harvard University Press, 1988.
2 Müller-Wille S. Race and history: comments from an epistemological point of view. *Sci Technol Human Values* 2014; **39:** 597–606.
3 Bhopal R. The beautiful skull and Blumenbach's errors: the birth of the scientific concept of race. *BMJ* 2007; **335:** 1308–09.
4 Kaufman JS, Cooper RS. Commentary: considerations for use of racial/ethnic classification in etiologic research. *Am J Epidemiol* 2001; **154:** 291–98.
5 Kittles RA, Weiss KM. Race, ancestry, and genes: implications for defining disease risk. *Annu Rev Genomics Hum Genet* 2003; **4:** 33–67.
6 Montagu A. Man's most dangerous myth: the fallacy of race, 6th edition. Lanham, MA: AltaMira Press, 1997.
7 Cerdeña JP, Plaisime MV, Tsai J. From race-based to race-conscious medicine: how anti-racist uprisings call us to act. *Lancet* 2020; **396:** 1125–28.
8 Race E. The use of racial, ethnic, and ancestral categories in human genetics research. *Am J Hum Genet* 2005; **77:** 519–32.
9 Kukutai T. Building ethnic boundaries in New Zealand: representations of Maori identity in the census. In: Axelsson P, Sköld P, eds. Indigenous peoples and demography: the complex relation between identity and statistics. New York, NY: Berghahn Books, 2011.
10 U.S. Census Bureau. US Census index of questions. 2021. https://www.census.gov/history/www/through_the_decades/index_of_questions/1930_1.html (accessed Nov 14, 2022).
11 Ministry of Social Development. Ethnicity measures, intermarriage and social policy. 2004. https://www.msd.govt.nz/about-msd-and-our-work/publications-resources/journals-and-magazines/social-policy-journal/spj23/ethnicity-measures-intermarriage- and- social- policy-pages109-140.html (accessed Nov 9, 2022).
12 Gonzalez-Barrera A, Lopez MH. Is being Hispanic a matter of race, ethnicity or both? 2015. https://www.pewresearch.org/fact-tank/2015/06/15/is-being-hispanic-a-matter-of-race-ethnicity-or-both/ (accessed Jan 29, 2021).
13 Korelitz S. The Menorah Idea: from religion to culture, from race to ethnicity. *Am Jew Hist* 1997; **85:** 75–100.
14 Bhopal R. Race and ethnicity: responsible use from epidemiological and public health perspectives. *J Law Med Ethics* 2006; **34:** 500–07.
15 Moffitt U, Juang LP, Syed M. Intersectionality and youth identity development research in Europe. *Front Psychol* 2020; **11:** 78.
16 Kressin NR, Chang BH, Hendricks A, Kazis LE. Agreement between administrative data and patients' self-reports of race/ethnicity. *Am J Public Health* 2003; **93:** 1734–39.
17 Cerdeña JP, Grubbs V, Non AL. Genomic supremacy: the harm of conflating genetic ancestry and race. *Hum Genomics* 2022; **16:** 18.
18 Haynes MA, Smedley BD. The unequal burden of cancer: an assessment of NIH research and programs for ethnic minorities and the medically underserved. Washington, DC: National Academy of Sciences, 1999.
19 National Heart, Lung, and Blood Institute. Recommendations on the use and reporting of race, ethnicity, and ancestry in the NHLBI Trans-Omics for Precision Medicine (TOPMed) program. 2021. https://topmed.nhlbi.nih.gov/guidelines-use-and-reporting-race-ethnicity-and-ancestry-topmed (accessed March 14, 2022).

20 Yudell M, Roberts D, DeSalle R, Tishkoff S. NIH must confront the use of race in science. *Science* 2020; **369:** 1313–14.

21 High Level Group on Non-discrimination, Equality and Diversity: Subgroup on equality data. Guidance note on the collection and use of equality data based on racial or ethnic origin. 2021. https://fra.europa.eu/sites/default/files/fra_uploads/ec-july-2018-guidelines-equality-data-collection.pdf (accessed July 20, 2022).

22 Merz LE, Achebe M. When non-Whiteness becomes a condition. *Blood* 2021; **137:** 13–15.

23 Nadkarni GN, Gignoux CR, Sorokin EP, et al. Worldwide frequencies of *APOL1* renal risk variants. *N Engl J Med* 2018; **379:** 2571–72.

24 Cerdeña JP, Tsai J, Grubbs V. *APOL1*, Black race, and kidney disease: turning attention to structural racism. *Am J Kidney Dis* 2021; **77:** 857–60.

25 Kaye JB, Schultz LE, Steiner HE, Kittles RA, Cavallari LH, Karnes JH. Warfarin pharmacogenomics in diverse populations. *Pharmacotherapy* 2017; **37:** 1150–63.

26 Boyd RW, Lindo EG, Weeks LD, McLemore MR. On racism: a new standard for publishing on racial health inequities. 2020. https://www.healthaffairs.org/do/10.1377/forefront.20200630.939347/ (accessed July 23, 2020).

27 Caulfield T, Fullerton SM, Ali-Khan SE, et al. Race and ancestry in biomedical research: exploring the challenges. *Genome Med* 2009; **1:** 8.

28 Brothers KB, Bennett RL, Cho MK. Taking an antiracist posture in scientific publications in human genetics and genomics. *Genet Med* 2021; **23:** 1004–07.

29 United Nations Human Rights Office of the High Commissioner. A human rights-based approach to data: leaving no one behind in the 2030 agenda for sustainable development. 2018. https://www.ohchr.org/sites/default/files/Documents/Issues/HRIndicators/GuidanceNoteonApproachtoData.pdf (accessed July 20, 2022).

30 Borrell LN, Elhawary JR, Fuentes-Afflick E, et al. Race and genetic ancestry in medicine—a time for reckoning with racism. *N Engl J Med* 2021; **384:** 474–80.

31 Fleskes RE, Bader AC, Tsosie KS, Wagner JK, Claw KG, Garrison NA. Ethical guidance in human paleogenomics: new and ongoing perspectives. *Annu Rev Genomics Hum Genet* 2022; **23:** 627–52.

32 Jackson L, Tsosie KS, Fox K. Changing the wrapping won't fix genetic-racism package. *Nature* 2021; **597:** 475.

33 Roberts D. Fatal invention: how science, politics, and big business re-create race in the twenty-first century. New York, NY: New Press, 2011.

34 Fullwiley D. DNA and our twenty-first-century ancestors. 2021. https://www.bostonreview.net/articles/duana-fullwiley-dna-and-our-twenty-first-century-ancestors/ (accessed Feb 9, 2021).

35 Tsai J, Cerdeña JP, Khazanchi R, et al. There is no 'African American Physiology': the fallacy of racial essentialism. *J Intern Med* 2020; **288:** 368–70.

36 Lewis ACF, Molina SJ, Appelbaum PS, et al. Getting genetic ancestry right for science and society. *Science* 2022; **376:** 250–52.

37 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* 2015; **96:** 37–53.

38 Mersha TB, Abebe T. Self-reported race/ethnicity in the age of genomic research: its potential impact on understanding health disparities. *Hum Genomics* 2015; **9:** 1.

39 Kaseniit KE, Haque IS, Goldberg JD, Shulman LP, Muzzey D. Genetic ancestry analysis on >93,000 individuals undergoing expanded carrier screening reveals limitations of ethnicity-based medical guidelines. *Genet Med* 2020; **22:** 1694–702.

40 Shim JK, Ackerman SL, Darling KW, Hiatt RA, Lee SSJ. Race and ancestry in the age of inclusion: technique and meaning in post-genomic science. *J Health Soc Behav* 2014; **55:** 504–18.

41 Penman BS, Gupta S, Buckee CO. The emergence and maintenance of sickle cell hotspots in the Mediterranean. *Infect Genet Evol* 2012; **12:** 1543–50.

42 Ashorobi D, Ramsey A, Yarrarapu SNS, Bhatt R. Sickle cell trait. Treasure Island, FL: StatPearls Publishing, 2021.

43 Handley LJL, Manica A, Goudet J, Balloux F. Going the distance: human population genetics in a clinal world. *Trends Genet* 2007; **23:** 432–39.

44 Li JZ, Absher DM, Tang H, et al. Worldwide human relationships inferred from genome-wide patterns of variation. *Science* 2008; **319:** 1100–04.

45 Pereira FDSCF, Guimarães RM, Lucidi AR, Brum DG, Paiva CLA, Alvarenga RMP. A systematic literature review on the European, African and Amerindian genetic ancestry components on Brazilian health outcomes. *Sci Rep* 2019; **9:** 8874.

46 Flores C, Ma SF, Pino-Yanes M, et al. African ancestry is associated with asthma risk in African Americans. *PLoS One* 2012; **7:** e26807.

47 Peterson RE, Kuchenbaecker K, Walters RK, et al. Genome-wide association studies in ancestrally diverse populations: opportunities, methods, pitfalls, and recommendations. *Cell* 2019; **179:** 589–603.

48 Bonilla C, Boxill LA, Donald SAM, et al. The *8818G* allele of the agouti signaling protein (ASIP) gene is ancestral and is associated with darker skin color in African Americans. *Hum Genet* 2005; **116:** 402–06.

49 Liu F, Visser M, Duffy DL, et al. Genetics of skin color variation in Europeans: genome-wide association studies with functional follow-up. *Hum Genet* 2015; **134:** 823–35.

50 American Association of Biological Anthropologists. AABA statement on race and racism. 2019. https://physanth.org/about/position-statements/aapa-statement-race-and-racism-2019/ (accessed March 25, 2020).

51 Tang H, Jorgenson E, Gadde M, et al. Racial admixture and its impact on BMI and blood pressure in African and Mexican Americans. *Hum Genet* 2006; **119:** 624–33.

52 Non AL, Gravlee CC, Mulligan CJ. Education, genetic ancestry, and blood pressure in African Americans and Whites. *Am J Public Health* 2012; **102:** 1559–65.

53 Braun L, Wolfgang M, Dickersin K. Defining race/ethnicity and explaining difference in research studies on lung function. *Eur Respir J* 2013; **41:** 1362–70.

54 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:** 621–28.

55 Bailey ZD, Krieger N, Agénor M, Graves J, Linos N, Bassett MT. Structural racism and health inequities in the USA: evidence and interventions. *Lancet* 2017; **389:** 1453–63.

56 Ioannidis JPA, Powe NR, Yancy C. Recalibrating the use of race in medical research. *JAMA* 2021; **325:** 623–24.

57 Williams DR, Mohammed SA. Racism and health I: pathways and scientific evidence. *Am Behav Sci* 2013; **57:** 1152–73.

58 Kuzawa CW, Sweet E. Epigenetics and the embodiment of race: developmental origins of US racial disparities in cardiovascular health. *Am J Hum Biol* 2009; **21:** 2–15.

59 Non AL. Social epigenomics: are we at an impasse? *Epigenomics* 2021; **13:** 1747–59.

60 Gravlee CC. How race becomes biology: embodiment of social inequality. *Am J Phys Anthropol* 2009; **139:** 47–57.

61 Mathur R, Rentsch CT, Morton CE, et al. Ethnic differences in SARS-CoV-2 infection and COVID-19-related hospitalisation, intensive care unit admission, and death in 17 million adults in England: an observational cohort study using the OpenSAFELY platform. *Lancet* 2021; **397:** 1711–24.

62 Tan SB, deSouza P, Raifman M. Structural racism and COVID-19 in the USA: a county-level empirical analysis. *J Racial Ethn Health Disparities* 2022; **9:** 236–46.

63 Gravlee C. Racism, not genetics, explains why Black Americans are dying of COVID-19. 2020. https://blogs.scientificamerican.com/voices/racism-not-genetics-explains-why-black-americans-are-dying-of-covid-19/ (accessed July 26, 2022).

64 Jones DS. Moving beyond race-based medicine. *Ann Intern Med* 2021; **174:** 1745–46.

65 Durie M. Understanding health and illness: research at the interface between science and indigenous knowledge. *Int J Epidemiol* 2004; **33:** 1138–43.

66 Link BG, Phelan J. Social conditions as fundamental causes of disease. *J Health Soc Behav* 1995; **35:** 80–94.

67 Hardeman RR, Murphy KA, Karbeah J, Kozhimannil KB. Naming institutionalized racism in the public health literature: a systematic literature review. *Public Health Rep* 2018; **133:** 240–49.

68 Krieger N, Boyd RW, De Maio F, Maybank A. Medicine's privileged gatekeepers: producing harmful ignorance about racism and health. 2021. https://www.healthaffairs.org/do/10.1377/hblog20210415.305480/full (accessed Aug 10, 2021).

69 Chantarat T, Van Riper DC, Hardeman RR. The intricacy of structural racism measurement: a pilot development of a latent-class multidimensional measure. *EClinicalMedicine* 2021; **40:** 101092.

70 Chantarat T, Van Riper DC, Hardeman RR. Multidimensional structural racism predicts birth outcomes for Black and White Minnesotans. *Health Serv Res* 2022; **57:** 448–57.

71 Hardeman RR, Homan PA, Chantarat T, Davis BA, Brown TH. Improving the measurement of structural racism to achieve antiracist health policy. *Health Aff (Millwood)* 2022; **41:** 179–86.

72 Gordon SH, Sommers BD, Wilson IB, Trivedi AN. Effects of medicaid expansion on postpartum coverage and outpatient utilization. *Health Aff (Millwood)* 2020; **39:** 77–84.

73 Agénor M, Perkins C, Stamoulis C, et al. Developing a database of structural racism-related state laws for health equity research and practice in the United States. *Public Health Rep* 2021; **136:** 428–40.

74 Lukachko A, Hatzenbuehler ML, Keyes KM. Structural racism and myocardial infarction in the United States. *Soc Sci Med* 2014; **103:** 42–50.

75 Carnethon MR, Kershaw KN, Kandula NR. Disparities research, disparities researchers, and health equity. *JAMA* 2020; **323:** 211–12.

76 Lawson DJ, Davies NM, Haworth S, et al. Is population structure in the genetic biobank era irrelevant, a challenge, or an opportunity? *Hum Genet* 2020; **139:** 23–41.

77 Reich D, Price AL, Patterson N. Principal component analysis of genetic data. *Nat Genet* 2008; **40:** 491–92.

78 Sirugo G, Williams SM, Tishkoff SA. The missing diversity in human genetic studies. *Cell* 2019; **177:** 26–31.

79 Epstein S. Inclusion: the politics of difference in medical research. Chicago, IL: University of Chicago Press, 2009.

80 Varma T, Jones CP, Oladele C, Miller J. Diversity in clinical research: public health and social justice imperatives. *J Med Ethics* 2022; published online April 15. https://doi.org/10.1136/medethics-2021-108068