Population designations in biomedical research: Limitations and perspectives

Gombault, Caroline; Grenet, Guillaume; Segurel, Laure; Duret, Laurent; Gueyffier, François; Cathébras, Pascal; Pontier, Dominique; Mainbourg, Sabine; Sanchez-Mazas, Alicia; Lega, Jean-Christophe

How to cite

GOMBAULT, Caroline et al. Population designations in biomedical research: Limitations and perspectives. In: HLA, 2022, vol. 101, n° 1, p. 3–15. doi: 10.1111/tan.14852

This publication URL: https://archive-ouverte.unige.ch/unige:177071

Publication DOI: 10.1111/tan.14852

Received: 20 May 2022 | Revised: 3 October 2022 | Accepted: 14 October 2022

DOI: 10.1111/tan.14852

REVIEW ARTICLE

HLA Immune Response Genetics WILEY

Population designations in biomedical research: Limitations and perspectives

Caroline Gombault[1] | **Guillaume Grenet**[1,2] | **Laure Segurel**[1] | **Laurent Duret**[1] | **François Gueyffier**[1,3] | **Pascal Cathébras**[4] | **Dominique Pontier**[1] | **Sabine Mainbourg**[1,5] | **Alicia Sanchez-Mazas**[6] | **Jean-Christophe Lega**[1,5]

[1]Laboratoire de Biométrie et Biologie Evolutive, Université Lyon 1, UMR CNRS 5558, Lyon, France

[2]Pole de Santé Publique, Hospices Civils de Lyon, Service Hospitalo-Universitaire de PharmacoToxicologie, Lyon, France

[3]Pôle de Santé Publique, Hospices Civils De Lyon, Lyon, France

[4]Service de Médecine Interne, Hôpital Nord, CHU de Saint-Etienne, Saint-Etienne, France

[5]Service de Médecine Interne et Pathologie Vasculaire, Hôpital Lyon Sud, Hospices Civils De Lyon, Lyon, France

[6]Laboratory of Anthropology, Genetics and Peopling history, Department of Genetics and Evolution, University of Geneva, Geneva, Switzerland

Correspondence
Jean-Christophe Lega, Hôpital Lyon Sud, Chemin du grand Revoyet, 69495 – Pierre Benite, Lyon, France.
Email: jean-christophe.lega@chu-lyon.fr

Funding information
Schweizerischer Nationalfonds zur Förderung der Wissenschaftlichen Forschung, Grant/Award Numbers: 1003A_144180, 310030_188820

In biomedical research, population differences are of central interest. Variations in the frequency and severity of diseases and in treatment effects among human subpopulation groups are common in many medical conditions. Unfortunately, the practices in terms of subpopulation labeling do not exhibit the level of rigor one would expect in biomedical research, especially when studying multifactorial diseases such as cancer or atherosclerosis. The reporting of population differences in clinical research is characterized by large disparities in practices, and fraught with methodological issues and inconsistencies. The actual designations such as "Black" or "Asian" refer to broad and heterogeneous groups, with a great discrepancy among countries. Moreover, the use of obsolete concepts such as "Caucasian" is unfortunate and imprecise. The use of adequate labeling to reflect the scientific hypothesis needs to be promoted. Furthermore, the use of "race/ethnicity" as a unique cause of human heterogeneity may distract from investigating other factors related to a medical condition, particularly if this label is employed as a proxy for cultural habits, diet, or environmental exposure. In addition, the wide range of opinions among researchers does not facilitate the attempts made for resolving this heterogeneity in labeling. "Race," "ethnicity," "ancestry," "geographical origin," and other similar concepts are saturated with meanings. Even if the feasibility of a global consensus on labeling seems difficult, geneticists, sociologists, anthropologists, and ethicists should help develop policies and practices for the biomedical field.

KEYWORDS
biomedical translational science, cultural diversity, ethnicity, genetic variation, racial groups

Alicia Sanchez-Mazas and Jean-Christophe Lega contributed equally to the study.

Following Raj Bhopal's example, some terms have been put in quotations "to alert the reader to the limited use of the word."

1 | INTRODUCTION

Population differences are of central interest in biomedical research. Variations in the frequency and severity of diseases among human population groups have been shown in many medical conditions,[1] including type 2 diabetes, obesity,[2] cardiovascular diseases,[3–5] prostate cancer,[6] transplantation,[7–9] asthma,[10] systemic lupus erythematosus,[11] and neonatal morbidities.[12] Similarly, a number of investigations have provided insights into differences observed after exposure or in response to a given medicine[2,5,13–16] as well as after tissue transplantation[9] across various populations. While many efforts are made to enhance our understanding of the mechanisms explaining differences in treatment effects, such as that of dipeptidyl peptidase-4 inhibitors in diabetes[17] and antihypertensive drugs,[5] most of the time it remains empirical. It is important to note that, since humans share over 99.9% of their genomes, the phenotypic differences between them are related to a 0.1% genome variation.[18]

Consequently, with the aim of developing individually-tailored healthcare strategies, the Food and Drug Administration (FDA) states that "Collecting data on race and/or ethnicity is critical to identifying population-specific signals."[19] Ramamoorthy et al. found that, in all the new molecular entities approved by the American agency between 2008 and 2013, 21% of them demonstrated population differences.[13] Along the same lines, the guidance "Ethnic Factors in the Acceptability of Foreign Clinical Data" of the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH E5) states that "any candidate medicine for global development should be characterized as ethnically sensitive or insensitive."[11] As a result, and in a holistic spirit, assigning an alleged affiliation to a particular population to participants in clinical studies is seen as a normal methodological choice, and the terminologies used to name the different human population groups are now being used extensively in health research.

Numerous systematic reviews on this topic have been performed, aiming to evaluate the use of these population denominations in biomedical research nowadays. A brief comparison of them reveals that the reporting of population differences in clinical research is characterized by a large disparity in practices, and fraught with methodological issues and inconsistencies. A systematic review of the use of "race/ethnicity" in three top medical journals reported 100 different population labels employed to describe study samples.[20] In another study reviewing a sample of 995 articles reporting "race/ethnicity," only 4.5% of them formally defined race/ethnicity and only 10.5% reported the method used to attribute the respective "racial/ethnic" designations.[21]

These findings point out the limitations concerning the practices in terms of population labeling in medical research. The aim of this narrative review is to explain the origins of such a haphazard approach to human diversity and its impact on biomedical research.

2 | ABOUT THE HIGH HETEROGENEITY IN POPULATION DESIGNATIONS

The dizzying array of terms can be partly explained by the origins of human classifications. Colonization first led to the advent of the folk idea of "race," when European armies started to encounter populations that differed from them in terms of both physical and cultural features.[22,23] For the purpose of standardizing "racial" classifications, modern human categories emerged from the work of European naturalists in the 18th Century,[24,25] leading to the labels "Mongoloid," "Negroid," or "Australoid," depending on the geographical area of origin.[26–28] In order to support the speculations of polygenists and advocates of "racial" ideology,[29,30] scientists began searching for evidence of natural racial distinctions, notably through craniometry, biology, anthropology, and medicine.[1,26,31,32] Some populations were then considered to be more vulnerable to illness and became commonly associated with it—for instance, diabetes was called "Judenkrankheit" (Jewish disease) in Germany at the beginning of the 20th Century.[33] Such racial classifications became ingrained in biomedical research, which slowly led to a common acceptance of "race" as a biological variable.[31,34–36]

It is also important to examine the origins of the human classifications currently used. Terminology has evolved since the classification systems proposed by naturalists.[37] Nowadays, population classification schemes are widely inspired by the legal and social classification systems of the countries in which the research is conducted.[30] Many biomedical scientists adopted official census categories in their trial designs and sampling strategies.[1,38,39] However, the definitions and the systems used to define population subgroups vary considerably from one country to another,[25,40,41] which partly explains the wide diversity of terms and approaches that can be observed in the international biomedical literature.[26,42]

The United States, heavyweight of international biomedical research, refers to the Office of Management and

Budget (OMB) Directive No. 15,[19] adopted in 1977 and revised in 1997.[37,43,44] The present day classification includes five "racial" groups based on geographical origin or skin color: "American Indian or Alaska Native," "Asian," "Black or African American," "Native Hawaiian or Other Pacific Islander," and "White," with a "Hispanic or Latino" category for ethnicity,[19] added in 1970.[40] Multiple choice is permitted, and allows a total of 126 possible combinations.[27,43] In the United Kingdom, the classification approach has evolved over time.[40,45] Currently, the Office for National Statistics (ONS) provides 18 options for England and Wales grouped into five main groups, and these are commonly found in the scientific literature—including, to cite a few examples, "White and Black Caribbean", "White and Asian", "Indian," "Black British," and "Arab."[29] In France, often regarded as an "ethnicity-blind" nation, there is no official demographic categorization scheme due to the almost complete prohibition of ethnic-based statistics.[40,41,46,47]

These three cases are sufficient to give an insight into why population denominations are far from being uniform on a global scale.

3 | THE ISSUES RAISED BY THE CURRENT STANDARDS

3.1 | Blurred concepts

The limitations concerning the use of "race/ethnicity" as variables in research are more and more acknowledged by scientists,[27,48] as well as the complexity of defining adequate labels for different human populations. The work of convincingly deconstructing the legitimacy of common designations, considered by some as reflecting a "biological naivety of researchers,"[49] has led to an increasing amount of publications in biomedical journals, resulting in numerous controversies and a passionate debate within the scientific community.[22]

The first issue concerns the breadth of potential interpretations regarding the concepts related to human diversity, since they are imprecise and lack clear-cut boundaries.[32,50] The most obvious concept attached to human diversity is "race." "Race" has no standard definition.[26] It is historically seen as a tool to subdivide the human species into groups of relatively homogenous people, sharing inborn physical characteristics that distinguish them from other groups, and which would reflect their "ancestry" and "geographical origins."[1,13,24] The change in mentalities led to a broadening of its scope to include social and cultural features,[10] making its differentiation with the concept of "ethnicity" unclear, as evidenced by the increasing use of the combined terminology "race/ethnicity."[1,37] "Ethnicity" is seen as a construct combining geographical origin and social dimensions,[44,51,52] encompassing a broad range of socially constructed characteristics derived from a shared history and a shared heritage, such as language, religion, habits, traditions, values, relational styles, and customs.[10,13,53–55] It is a multifaceted quality,[1,41] "achieved" but "not ascribed,"[34] a flexible cultural construct that can have biological implications.[23] It is increasingly employed, especially in Europe, as it is seen as more politically correct than "race."[1,24,29,34]

The terms "population" and "ancestry" appear to be less controversial.[56] However, their use is also subject to caution, given their lack of clear definition and their wide range of possible interpretations.[57] Their "neutral" appearance is a smokescreen: George Ellison reminds us that "population group" was the term used during apartheid in South Africa.[58] Some see the notion of "shared ancestry" as merely overlaying "racial" categories.[57] Additionally, although recent generations of ancestors are easy to identify, defining, and interpreting deeper genealogical information is highly complex.[57] In a study addressing the correlation between multiple self-identification measures and classifications based on ancestry informative markers (AIM) and HLA genes in the United States, the geographical ancestry of grandparents was more closely correlated with AIM and HLA-defined genetic ancestries than with "race/ethnicity," especially for subjects with African and Latin American ancestry.[59] Collecting information on the grandparents of donors may thus improve the chances of finding HLA-compatible donors in hematopoietic stem cell transplantation (HSCT), where HLA matches are essential. However, the correlation with the geographical origin of grandparents still remains imperfect for estimating deep genetic ancestry.[59] "Continent of origin" is also a very broad term that sounds neutral. Touted as the solution to old notions of "race" and "ethnicity," it refers to extremely large clusters, which in the end makes it useless for any proper study of human diversity. Continental population labels actually reinvigorate ideas about "major races," which would be "Africans" or "Europeans."[57]

Finally, the categories may change according to the set of reference populations used in a study, which are often poorly described in studies.[39,60] A high inconsistency regarding "racial" and "ethnic" classifications across papers as well as between countries have also been reported in a number of biomedical research fields, including pharmacogenomics.[61] The concept of "ancestry" is also ambiguous and encompasses distinct approaches, framed as "genetic ancestry": (i) methods to infer ancestral relationships between populations, (ii) methods that allow an individual genomes to be represented as a combination of reference

populations, and (iii) methods that an infer a degree of similarity between individuals.[60] However, only the third approach allows showing the continuous, that is, category-free, nature of genetic variation.[60]

3.2 | Imprecise designations

The current designations refer to broad, heterogeneous groups.[50] The epithet "Europeans" encompasses an extremely wide range of different phenotypes with, for instance, hair, eye, and skin color varying from relatively dark in Southern Europe to very light in the North.[62] Metabolic pathways and culture also vary considerably as demonstrated by the fact that the main allele (−13,910*T) conferring lactose tolerance varies from less than 10% to more than 95% in Europeans[63] and more than 200 languages (despite 24 official) are spoken in Europe.[64] Similarly, the term "African" is a crude simplification. More than 1.2 billion humans live in Africa, which is also the continent with the greatest amount of genetic variation on the planet.[65] Although other labels appear to be more accurate, they are not - the category "American Indians and Alaska Natives", also known as "Native Americans," actually refers to more than 500 federally recognized tribes, which are too different from each other to be grouped in one single "pan-Indian" group.[66] The breadths of these terms make them inadequate for their use in medical research.

Many used terms are simplistic, and represent, for most, obsolete terms.[31,35,37,38] The well-known "Caucasian" label, for instance, is an unfortunate relic of an old classification of humankind created by a German naturalist named Johann Friedrich Blumenbach in the 18th Century, who derived Europeans from the region of Caucasus.[1,7,24,56,67] The label "Black," while being a legacy of colonial anthropology encouraging the division of society by skin color,[7] is meaningless in terms of genetic ancestry - dark-skinned people are observed in many areas of the globe, from Africa to India, as well as in Southeast Asia and Australia.[68] The term "mixed" is in itself a nonsense. This category would be of major importance given the increasing rates of mixed sexual unions between distantly related people, particularly in countries such as the United States.[43] However, the number of potential combinations is so large that the willingness of capturing any "mixed" identity generates substantial reporting issues.[45] Grouping data of all "mixed" individuals in a single category has, in the end, the same effectiveness as having no data.[37,50,54,66] In addition, the term "trans-ethnic" is debated in genomics research because of its imprecision and ambiguity. This term encompasses heterogeneous studies including meta-analytic genome-wide association studies, fine mapping, and assessment of genetic architecture across populations, or calculation of polygenic risk scores. Kamariza et al. suggested the use of the terms "cross-ancestry," "multi-population" or "multi-ancestry" instead.[69]

These concepts are subjective, and therefore not easily assigned.[44,70] Penner et al. reported in 2008, from a sample of 12,686 American individuals, that 20% had been affiliated to at least two different "racial" categories by interviewers over a 19-year period.[71] According to distinct reporting formats (self-identification of "race/ethnicity" using single or multiple choices, geographical ancestry using single or multiple choices, classification by others), inconsistencies may occur between "race/ethnicity" and geographical ancestry as well as between self-identification and identification by others.[59] Most often, a lack of concordance between geographical ancestry and "race/ethnicity" reporting occurs when individuals acknowledge particular geographical ancestries but do not explicitly identify with a corresponding "racial/ethnic" group. Less common is "race/ethnicity" self-identification without reporting a corresponding geographical ancestry.[59]

Many dimensions must be taken into consideration when defining identity, and terminology fails to embrace them all—"Hispanic" invokes language, "American Indian" invokes ancestral groups, "Asian" invokes continental origin, "White" invokes skin color.[72] The United Nations stated about "ethnicity": "the subjective nature of the term [...] requires that information on ethnicity be acquired through self-declaration of a respondent".[41] This principle is sometimes circumvented by the use of name recognition software, such as the "Ethnicolr" program.[4,73] However, "ethnicity" reflects, above all, the feeling of belonging to a chosen community, which is subject to change.[27,29,30,32,33,50] Census categories, for instance, can take on new social or political meanings. An example concerns the restoration of Aboriginal people's pride in Canada which led to a major increase in the affiliations of individuals to this category in censuses at the end of the 20th Century.[40] Conversely, latent racism may influence self-identification in favor of a better socially-perceived group.[44] Since context influences categorization,[1,29] population designations are interchangeable.

The current use of population designations is ethically questionable.[27] Census categories were originally intended only for social and pragmatic reasons, but their widespread use in research unintentionally implies their adequacy for describing so-called "natural" patterns of diversity.[26,28,36,50,58] Even though the Nazi crimes of World War II largely tempered the previously facile acceptance of "race" as deeply and unalterably biogenetic,[37] and forced scientists to reexamine their thinking on "race," the 21st Century has surprisingly

witnessed a resurgence of "race" through genetics.[23,36,72] Relying on contemporary genomic findings, many researchers see "racial" categories as "reflecting genetic," argue that DNA variations between geographically distinct individuals confirm the existence of biological human "races,"[23,74] or believe that race is a valid and reliable proxy for the clustering of genetic diversity.[49] Some academics have found "new grist for the racial differences mill."[23] This renewed interest in clustering human beings genetically, and the attempts to link this clustering to biological outcomes, regrettably led to a normalization of the use of population designations in an uncritical manner in biomedical research.[28,33]

4 | THE NUMEROUS CONSEQUENCES OF THE USE OF "RACE/ETHNICITY" IN MEDICAL RESEARCH

The use of "race/ethnicity" as a cause of human heterogeneity can distract from investigating other factors that are really related to a medical condition. In 2013, Braun et al. observed that the majority of the literature studying lung function differences in "racial groups" ignored the importance of socioeconomic factors, despite their well-established influence on pulmonary conditions.[75] Yet, slavery and colonial history left a legacy of social and economic inequalities, inextricably linked to a greater exposure of immigrant populations to what Kreatsoulas and Anand called "the causes of the causes."[53] It refers to the social determinants of disease risk factors such as poor living conditions, low purchasing power, diminished access to healthcare, and health education.[22–24,76–78] Diet and exercise, smoking, cultural, and religious practices, compliance with medication, and other behavioral factors also result in health disparities.[4,50] To highlight the salience of considering environmental factors rather than relying on a self-defined "ancestry" designations, Cooper et al. compared rates of hypertension in several populations of West African "ancestry."[79] Their findings showed substantial differences in the prevalence of hypertension depending on location, rising from 16% for populations living in Nigeria and Cameroon to 26% in the Caribbean and 33% in metropolitan Chicago. Similarly, the prevalence of diabetes varies greatly among populations with comparable "ancestral background" but different environments.[33,806,26,29]

To prevent misunderstandings, we should make clear that claiming the inadequacy of the most common human classifications does not amount to asserting that biological and genetic differences among humans of various origins do not exist. The aim here is to shed light on the illogicality of trying to explain differences in disease between humans using such crude concepts and categories. The use of such classifications as shortcuts for the prevalence of a few rare genetic mutations such as that of sickle cell trait in populations of African, Southern European, and Caribbean origin, or of the BRCA gene in individuals self-identifying as "Ashkenazi Jews," may be useful in practice, if cautiously used.[6,26,29] However, as soon as we consider complex diseases such as asthma, cancer, diabetes, or cardiovascular diseases, they may become scientifically inadequate and dangerously misleading. Therefore, their use in research raises concerns, and the trivialization of their reporting in biomedical literature has serious consequences.

4.1 | The need for transnational collaborative research

International research relies on data portability and reproducibility.[48] In the reviews aforementioned, population labels are not treated as requiring further clarification in publications, which opens the door to every possible interpretation. The lack of explanation suggests that they are self-evident, that the reader will simply recognize and understand them. Yet, unfortunately, the latter will develop his own understanding of who falls into these designations and who does not. Catherine Lee notes, "when we (researchers) use prototype theory, we have a "broad picture" in our minds about what we think we are classifying. Different social groups have different prototypes in mind."[34] As long as investigators fail to render the labeling of population designations transparent to readers, it will prevent making valid comparisons across clinical trials, even those conducted in the same country.[27,50,81]

Additionally, notwithstanding that current global standards are based on the societal stratification of the United States and, hence, do not adequately reflect other nations' migration histories, non-American researchers tend to adopt these standards in their own countries,[19,26] which can result in misinterpretations. The term "Asian," for instance, generally refers to people from the Indian subcontinent in the United Kingdom, but to Far Eastern Asians in the United States.[1] Similarly, the term "White" is meant to include Scottish, New Zealander, Greek, Spanish, English, Canadian, Welsh, and Irish altogether, but also includes Iranians and Moroccans in the United States.[37] Communication issues can also arise from the fact that presumably similar groups can be exposed to completely different confounding factors according to the country. For example, Nazroo et al. showed that Caribbean immigrants have very different social and economic positions in the United Kingdom

and in the United States, impacting their health differently both in positive and negative ways.[77]

Accordingly, this haphazard use and reporting of human diversity is at odds with the fundamental requirements of transnational research, which are based on rigorous practices and valid data.[74,82]

In the domain of histocompatibility, such as HSCT where matching HLA genotypes between recipients and donors is essential, the chances of finding a compatible donor is expected to be greater among individuals sharing a greater level of genetics.[83] In some cases (e.g. in Brazil), it may correlate with government-defined broad racial categories such as "Black," "Mixed," and "White," because the latter reflect different amounts of African and European ancestry.[84] However, the use of such categories in transplantation may stigmatize the idea that graft success is race-dependent, which is not. Most HLA alleles encoding functional molecules at the protein level (known as alleles defined at the second-field level of resolution)[85] and hence, crucial in transplant matching, are shared across continents. This thus provides chances of finding compatible donors among distantly-related populations, although identifying a fully-matched donor at multiple HLA loci (ideally both alleles at five loci, known as 10/10 matching) remains challenging when no HLA-identical sibling donor is available.[86]

4.2 | Reinforcing stereotypes in medicine and biology

Many researchers claim that even if societal categories are ill-defined, they remain fairly correlated with economic factors and behavioral traits. But one should not minimize the potential social costs of this reasoning.[6,24,35,72] Human classification is by nature associated with stigmatization, marginalization, and prejudice, even though these concepts may be useful to address the effect of inequities on population health.[32,50,87] In research, it has been proven that "ancestry" or "origins" are poor proxies for sociocultural variables and, besides being a distraction from other relevant causes of disease,[27,32] their use emphasizes the alleged biological underpinnings of health inequalities.[58]

It has been demonstrated that the use of certain human categories in medical research influences the behavior of physicians, and reifies these categories by contributing to the perpetuation of healthcare disparities.[39,54] As a brief example, a study showed that patients with cystic fibrosis seen as being of "African origin" have a diagnostic disadvantage, because historically this affection was seen as being limited to "Europeans."[88] A similar situation was observed in women with osteoporosis.[48]

4.3 | The pitfalls of "racial-genetic determinism"

One of the main consequences of the trivialization of human classifications in research is the spread of "racial-genetic determinism"[28] or "genetic reductionism"[49,72] among researchers. Nancy Kieger denounces the "still dominant ahistorical and decontextualized biomedical and lifestyle theories of disease distribution, which reduce causes of disease to individuals' genetic constitution and 'personal tastes'".[78] Alan Goodman also warns us about "geneticization," which he defines as "the belief that most biology and behavior are located in the 'genes'".[32] Indeed, studies looking for "running genes" in Kenyan sprinters,[89] or linking the prevalence of diabetes in Pima Indians and asthma therapy efficiency in "Hispanic ascendance" to their retrospective degree of European admixture,[10] contribute to legitimizing simplistic conclusions on complex genetic and social processes.

Of course, some genetic variants do correlate with an individual's geographical origin, especially those which confer an adaptation to a specific local environment. It is the case for lactase persistence in pastoral populations,[90–92] and dark skin pigmentation which protects humans from sunburn and skin cancer, and also light skin pigmentation which facilitates the synthesis of vitamin D3 in northerly latitudes.[93,94] It is then commonplace to connect certain phenotypic features with specific areas of the globe. However, although some striking differences are observed concerning hair and eye color, the label "continental populations" has been proven to not be informative for other common heritable traits such as height,[95] which is subject to substantial differences between Maasai and Biaka people in sub-Saharan Africa, Swedish and Sicilians in Europe, and even new generations of Japanese compared to older generations.[22] The non-concordant patterns followed by genetic variations are so complex that attempting to graft them onto labeled delimited populations is misleading.[32,50,72,82] Alan Goodman said that "traits tend to vary independently of other traits" and that, consequently, human clustering schemes change according to the trait chosen: a classification based on sickle cell trait groups Africans, Greeks, and Turks together, while a classification based on lactase enzyme deficiency gathers southern Africans with Japanese and Native Americans.[32] The incidence of sickle cell traits in West Africa was about 4% in the Mandinka people, 14% in the Wolof people, rose to around 30% in the Fula people, but was also very high in India, with an incidence of 35% in the Oktar people.[75]

Surprisingly, the determining role of genetics for visible "skin deep" variations and marked population differences for a few rare traits led many biomedical scientists to apply the same reasoning to all diseases, thus

undervaluing the complexity of human genetic variation and of genotype–phenotype links.[20,28,35,49,96] This is highly problematic for research since, according to the WHO, the great majority of trials registered globally between 1999 and 2018 concern complex multigenic chronic affections such as diabetes, cancer, obesity, and hypertension. In these cases, the influence of genetic variants on incidence of disease or response to treatment is often associated with small effect sizes with low to moderate clinical consequences.[22,97,98]

This reasoning is also frequently found in pharmacology, especially in pharmacogenetics. Over the last decades, single nucleotide polymorphisms (SNPs) have been at the core of investigations, since they are known to have a great impact on pharmacokinetics and pharmacodynamics if located in some genes, specifically those encoding drug metabolizing cytochromes P450, phase II enzymes, and drug transporters. Factors influencing the emergence of these specific SNPs are still under deep investigation, and, ineluctably, "racial" considerations became center-stage, and prompted the FDA and other agencies to set up "population-specific recommendations" based on the genetic results observed in samples of individuals. For instance, the agency states that 7% to 10% of "Caucasians" and 3% to 8% of "Black/African Americans" are poor metabolizers for CYP2D6,[13] which can lead them to a higher exposure than others when taking certain medicines such as iloperidone. These low percentages show well that these phenomena are everything but "all-or-none."[14] Cytochromes P450 also displayed wide variations within populations.[99] In short, while identifying poor metabolizers is essential, such shortcuts have no place in pharmacogenetics, especially when a patient's safety is at stake.[14]

To conclude, it is widely acknowledged that all human beings do not have the exact same genetic makeup, and that genomics plays a great role in disease diagnosis and personalized treatments. However, commonly inferred genetic clusters, based on geographical arguments, are inaccurately represented by widely used "racial/ethnic" labels, and have been shown to inherently present a great and significant variance in drug-metabolizing profiles, on which target-oriented medicine largely relies on.[16]

5 | HOW CAN WE DURABLY IMPROVE PRACTICES IN RESEARCH?

5.1 | Enhance the involvement of biomedical journals to enforce "good practices" guidelines

Recommendations on how to handle population terminology in biomedical research have existed for many years.[27,82,100] At the beginning of their use, the FDA clearly stated that OMB categories, which now dominate scientific literature, were intended to ensure consistency in population data reporting, and specified that they should not be interpreted as being scientific or anthropological in nature.[19] Biomedical editors who are members of the International Committee of Medical Journal Editors (ICMJE) provided instructions to authors on the proper use of terminology for research on "race/ethnicity" in the fifth edition of their "Uniform Requirements for Manuscripts Submitted to Biomedical Journals" (renamed as "Recommendations for the Conduct, Reporting, Editing, and Publication of Scholarly Work in Medical Journals") including the justification and definition of "race/ethnicity" and the relevance of the use of such concepts in regards to the aim of the study.[34,37,101] In 2000, *Nature Genetics* published a policy stating that authors must explain why they make use of particular ethnic groups or populations, and how classification was achieved.[102] Unfortunately, these recommendations are rarely enforced by the reviewers' of biomedical journals or respected by the authors[21,24,39,82,103,104] including *Nature Genetics* itself.[72]

This global unheeding of guidance may be partly explained by the fact that calls for improving practices were made only in an intermittent and disjointed fashion, which consequently perpetuated ambiguity on this topic and made addressing change more challenging.[87] If all biomedical journals agree to strengthen the control of manuscripts' compliance with requirements on proper population labeling practices, improvements will surely be observed. Among the 486 reporting guidelines available on the Enhancing the QUAlity and Transparency Of health Research (EQUATOR) website, searching for the words "ethnicity" retrieved no result, and searching for 'race' gave three results but focused on specific situations (forensic context).[105]

5.2 | Changes in the mentalities of authors and their personal involvement

The profusion of "race/ethnicity" variables in biomedical research is partly rooted in the requirements established in 1993 by American regulatory bodies and funding agencies[19,106,107] to promote the inclusion of participants of all origins in study samples.[3,11,13,34,48,103,104,107–109] Reporting is, therefore, more frequent in the United States than in other countries[110] but, as mentioned before, is now accepted and expected internationally.[20] Beyond the fact that legal requirement is not a valid scientific argument,[38] the main concern here is that the use of "race/ethnicity" categories is often only required as a

"descriptive" tool meant to observe differences in inclusion or failures of randomization, but routinely drifts to an inappropriate "attributive" use of these labels as explanatory arguments.[26,39,50] Initially based on the well-intentioned goal of addressing underrepresentation of minorities in research, these requirements have unintendedly contributed to a further reifying of the census categories as genetic entities, and their continued use perpetuates the idea of them being adequate to explain patterns of human genetic variation.[74]

There is a clear contradiction in the research field: many studies showed that the limitations of common population designations is well-known and accepted by the research community, and, surprisingly, it is still extensively used and routinely reported in health research.[111] Many emphasize the need of reporting this variable for the sake of research, and affirm that "race" is the best proxy available for studying human diversity.[87] It is seen as useful, "in the meantime," until better surrogates are found.[26,54,72,97]

The lack of proper education concerning the question of human diversity is salient in biomedical education.[74] Behavior and genetic explanations are prioritized over discriminatory and environmental ones. Medical students traditionally learn that patients' histories begin with "this is a [insert patient's age, presumed race, sex] who presents with a chief complaint of…."—though they receive little or no training on these concepts.[27]

Some researchers, who explicitly defend the racial labeling of participants, qualify debates about the use of "race/ethnicity" in health as political correctness, denounce the "color-blind" approach to biomedical research as "[flying] in the face of clinical reality",[49,112] and claim that these practices are appropriate in the scientific area as soon as any value system is attached to it. Troy Duster reacts to the latter point by saying that "while the sentiment is admirable, this formulation constitutes a fundamentally flawed notion of a firewall between 'science' and 'politics'".[33] As an example, when the South African National Blood Service was blamed for classifying donors on the basis of "racial/ethnic" categories in order to estimate their risk of HIV infection, some clinicians found it regrettable to face political sensitivities for what they considered to be a risk-based approach founded on medical realities.[58] However, rather than blaming political correctness, underlining the role that apartheid and its socioeconomic consequences played in the differential spread of HIV could have explained more adequately the approach undertaken.[58] Lastly and briefly, other researchers are unfortunately reluctant to follow recommendations for the simple reason that it would require a revision of previous results, or a renunciation of the use of old databases built on OMB categories or on vague continental designations.[113]

In conclusion, practices regarding the labeling of subpopulations do not exhibit a level of rigor one would expect in biomedical research. Attempts of resolving these issues are not facilitated by the wide range of opinions among researchers, and, most of all, the quandary surrounding this topic. "Race," "ethnicity," "ancestry," "geographical origin," and other similar concepts are saturated with meanings and are increasingly considered as being non-acceptable due to their history of misuse and injustice - Cooper says that "race" is "the unwelcome guest in the disciplines of science."[30] The debate on whether reporting "race/ethnicity"-related outcomes is useful in research has persisted over many decades. Even if the feasibility of a global consensus regarding labeling is unclear,[41] sociologists, geneticists, anthropologists, and ethicists may help develop policies and practices for the biomedical field. For now, the priority for research publications are the following.

First, the use of outdated concepts such as "Caucasian" should be banned, while precise and adequate labeling (e.g., "African diaspora from the '15th Century to the present', 'grandparents born in Turkey') needs to be promoted according to the objectives and the setting of the study.[81] To reduce stereotypes in biomedical research, educational efforts are required during biological, medical, and pharmaceutical training to improve knowledge on human evolution, by including courses on human migration, demographic history, environmental adaptation, as well as genetic structure and diversity between individuals and populations. Multidisciplinary approaches are needed to propose new consensus[69] which would require the engagement of the communities concerned and international societies.[114]

Second, when it is legally required to ask patients to self-identify as being part of a particular population group in a study, it should, as much as possible, not appear in the study findings. Journal editors should promote adequate statistical methods (e.g., subgroup analysis in clinical trials[115,116]) (Figure 1) and develop new sets of standards in accordance with modern concepts. In 2021, the editors of the Journal of the American Medical Association (JAMA) acknowledged, as we do here, that "race" and "ethnicity" are social constructs without scientific or biological meaning and proposed guidance on how to report such designations in medical and science journals.[117] Among their recommendations, to provide explanations on the source and reasons of the classifications used and to prioritize the use of specific over collective terms (e.g., "underrepresented populations" rather than "minorities"; descriptions according to country or regional areas of origin rather than broad terms referring

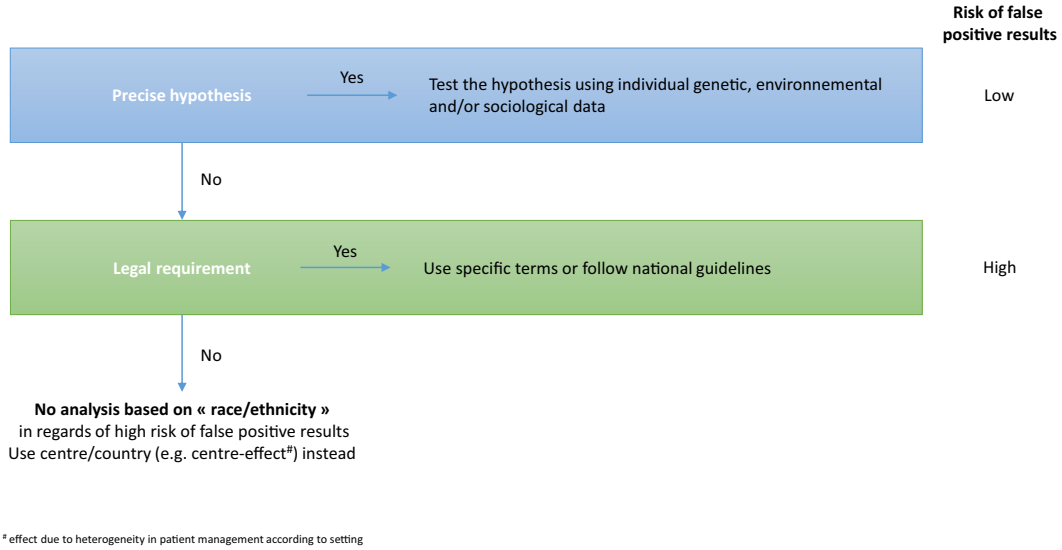

FIGURE 1 A proposal for the analysis and reporting of data from subpopulations in biomedical journals. An example of legal purpose may be the inclusion of groups that have been historically marginalized in randomized trials.19 The systematic search for associations between "race/ethnicity" and outcome/biomarker should be avoided due to the high risk of false positive results and/or confounding biases (e.g., sociological differences, environmental exposures)

to continental areas, etc.) represent an important step forward. However, other JAMA advices remain limited to the US context and are hardly exportable internationally: for example, the proper term of "race" is unaccepted in many countries like France,[118] and US government-defined descriptors like "Black" or "American Indian" are inadequate in non-US countries (e.g., "populations of African descent" is suggested for populations of African origin not living in Africa by the European HLA-NET network[7,38] and "First Nations" is used for indigenous people in Canada). A main challenge for scientific journals is thus still to adapt their guidance for population designations to an international context.

Finally, the reporting of patients' origins should not be tolerated if not justified, particularly if it is employed as a proxy for cultural habits, diet, or environmental exposure in regard to the high risk of false positive results (Figure 1). If all biomedical researchers adopt a critical view on how they handle and report human diversity, and invest themselves in adopting the right practices, real improvements can be made. However, for now sadly, the concepts of "race" and "ethnicity" remain deeply ingrained in biomedical research.

AUTHOR CONTRIBUTIONS

Caroline Gombault and Jean Christophe Lega designed this work. Caroline Gombault, Guillaume Grenet, Alicia Sanchez-Mazas, and Jean-Christophe Lega wrote the first draft of the manuscript. All authors revisited it critically and approved the final version.

ACKNOWLEDGMENT

The authors thank Véréna Landel for her valuable help in proofreading the manuscript.

FUNDING INFORMATION

This work was supported by the Swiss National Science Foundation (Grants Nos. 1003A_144180 and 310030_188820 to Alicia Sanchez-Mazasf).

CONFLICT OF INTEREST

The authors declare no conflicts of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Alicia Sanchez-Mazas https://orcid.org/0000-0002-7714-2432

Jean-Christophe Lega https://orcid.org/0000-0002-9398-2968

REFERENCES

1. Bhopal R. Glossary of terms relating to ethnicity and race: for reflection and debate. *J Epidemiol Community Health*. 2004; 58(6):441-445.
2. Osei-Assibey G, Adi Y, Kyrou I, Kumar S, Matyka K. Pharmacotherapy for overweight/obesity in ethnic minorities and white Caucasians: a systematic review and meta-analysis. *Diabetes Obes Metab*. 2011;13(5):385-393.

3. Sullivan LT, Randolph T, Merrill P, et al. Representation of black patients in randomized clinical trials of heart failure with reduced ejection fraction. *Am Heart J*. 2018;197:43-52.
4. Henderson SO, Haiman CA, Wilkens LR, Kolonel LN, Wan P, Pike MC. Established risk factors account for most of the racial differences in cardiovascular disease mortality. *PLoS One*. 2007;2(4):e377.
5. Brewster LM, Seedat YK. Why do hypertensive patients of African ancestry respond better to calcium blockers and diuretics than to ACE inhibitors and β-adrenergic blockers? A systematic review. *BMC Med*. 2013;11:141.
6. Rotimi CN, Jorde LB. Ancestry and disease in the age of genomic medicine. *N Engl J Med*. 2010;363(16):1551-1558.
7. Sanchez-Mazas A, Vidan-Jeras B, Nunes JM, et al. Strategies to work with HLA data in human populations for histocompatibility, clinical transplantation, epidemiology and population genetics: HLA-NET methodological recommendations. *Int J Immunogenet*. 2012;39(6):459-472. quiz 473–476.
8. Nunes JM, Buhler S, Roessli D, Sanchez-Mazas A, the HLA-net 2013 collaboration. The HLA-net GENE[RATE] pipeline for effective HLA data analysis and its application to 145 population samples from Europe and neighbouring areas. *Tissue Antigens*. 2014;83(5):307-323.
9. Morishima Y, Morishima S, Stevenson P, et al. Race and survival in unrelated hematopoietic cell transplantation. *Transplant Cell Ther*. 2022;28(7):357.e1-357.e6.
10. Ortega VE, Meyers DA. Pharmacogenetics: implications of race and ethnicity on defining genetic profiles for personalized medicine. *J Allergy Clin Immunol*. 2014;133(1):16-26.
11. Falasinnu T, Chaichian Y, Bass MB, Simard JF. The representation of gender and race/ethnic groups in randomized clinical trials of individuals with systemic lupus erythematosus. *Curr Rheumatol Rep*. 2018;20(4):20.
12. Janevic T, Zeitlin J, Auger N, et al. Association of Race/ethnicity with very preterm neonatal, et al. Association of Race/ethnicity with very preterm neonatal morbidities [internet]. *JAMA Pediatr*. 2018;172(11):30208467Accessed January 19, 2022. https://pubmed.ncbi.nlm.nih.gov/30208467/
13. Ramamoorthy A, Pacanowski M, Bull J, Zhang L. Racial/ethnic differences in drug disposition and response: review of recently approved drugs. *Clin Pharmacol Ther*. 2015;97(3):263-273.
14. McLeod HL. Pharmacogenetics: more than skin deep. *Nat Genet*. 2001;29(3):247-248.
15. Li JS, Baker-Smith CM, Smith PB, et al. Racial differences in blood pressure response to angiotensin-converting enzyme inhibitors in children: a meta-analysis. *Clin Pharmacol Ther*. 2008;84(3):315-319.
16. Wilson JF, Weale ME, Smith AC, et al. Population genetic structure of variable drug response. *Nat Genet*. 2001;29(3):265-269.
17. Kim YG, Hahn S, Oh TJ, Kwak SH, Park KS, Cho YM. Differences in the glucose-lowering efficacy of dipeptidyl peptidase-4 inhibitors between Asians and non-Asians: a systematic review and meta-analysis. *Diabetologia*. 2013;56(4):696-708.
18. Lander ES, Linton LM, Birren B, et al. Initial sequencing and analysis of the human genome. *Nature*. 2001;409(6822):860-921.
19. US Food and Drug Administration. *Collection of Race and Ethnicity Data in Clinical Trials [Internet]*. U.S. Food and Drug Administration; 2019;Accessed January 19, 2022. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/collection-race-and-ethnicity-data-clinical-trials
20. Kanakamedala P, Haga SB. Characterization of clinical study populations by race and ethnicity in biomedical literature. *Ethn Dis*. 2012;22(1):96-101.
21. Bokor-Billmann T, Langan EA, Billmann F. The reporting of race and/or ethnicity in the medical literature: a retrospective bibliometric analysis confirmed room for improvement. *J Clin Epidemiol*. 2020;119:1-6.
22. Cooper RS. Race and genomics. *N Engl J Med*. 2003;348:5-1170.
23. Smedley A, Smedley B. Race as biology is fiction, racism as a social problem is real: anthropological and historical perspectives on the social construction of race. *Am Psychol*. 2005;60(1):16-26.
24. Cathébras P. Pour en finir avec les «Caucasiens». Catégories raciales et ethniques dans la littérature médicale. *Rev Med Interne*. 2012;33(2):65-68.
25. Nobles M. History counts: a comparative analysis of racial/color categorization in US and Brazilian censuses. *Am J Public Health*. 2000;90(11):1738-1745.
26. Braun L, Fausto-Sterling A, Fullwiley D, et al. Racial categories in medical practice: how useful are they? *PLoS Med*. 2007;4(9):e271.
27. Winker MA. Measuring race and ethnicity: why and how? *JAMA*. 2004;292(13):1612-1614.
28. Gravlee CC. How race becomes biology: embodiment of social inequality. *Am J Phys Anthropol*. 2009;139(1):47-57.
29. Mathur R, Grundy E, Smeeth L. *Availability and Use of UK Based Ethnicity Data for Health Research - NCRM Eprints Repository [Internet]*. National Centre of Research Methods Accessed January 19, 2022. 2013 https://eprints.ncrm.ac.uk/id/eprint/3040/
30. Cooper RS. Race in biological and biomedical research. *Cold Spring Harb Perspect Med*. 2013;3(11):a008573.
31. Oppenheimer GM. Paradigm lost: race, ethnicity, and the search for a new population taxonomy. *Am J Public Health*. 2001;91(7):1049-1055.
32. Goodman AH. Why genes don't count (for racial differences in health). *Am J Public Health*. 2000;90(11):1699-1702.
33. Duster T. A post-genomic surprise. The molecular reinscription of race in science, law and medicine. *Br J Sociol*. 2015;66(1):1-27.
34. Lee C. "Race" and "ethnicity" in biomedical research: how do scientists construct and explain differences in health? *Soc Sci Med*. 2009;68(6):1183-1190.
35. Schwartz RS. Racial profiling in medical research. *N Engl J Med*. 2001;344(18):1392-1393.
36. Sankar P, Cho MK, Monahan K, Nowak K. Reporting race and ethnicity in genetics research: do journal recommendations or resources matter? *Sci Eng Ethics*. 2015;21(5):1353-1366.
37. Bhopal R, Donaldson L. White, European, Western, Caucasian, or what? Inappropriate labeling in research on race, ethnicity, and health. *Am J Public Health*. 1998;88(9):1303-1307.
38. Nunes JM, Buhler S, Sanchez-Mazas A. NO to obsolete definitions: YES to blanks. *Tissue Antigens*. 2014;83(2):119-120.

39. Ellison GTH, Smart A, Tutton R, Outram SM, Ashcroft R, Martin P. Racial categories in medicine: a failure of evidence-based practice? *PLoS Med*. 2007;4(9):e287.
40. Gagnon A, Piché V, Simon P. *IMISCOE Research Series - Social Statistics and Ethnic Diversity - IMISCOE Publications Catalogue [Internet]*. Springer; 2015.Accessed January 19, 2022. https://www.imiscoe.org/publications/library/2-imiscoe-research-series/83-social-statistics-and-ethnic-diversity.
41. Farkas L. *Report on Equality Data Based on Racial and Ethnic Origin [Internet]*. European Commission; 2017;Accessed January 19, 2022. https://ec.europa.eu/newsroom/just/items/112035.
42. United Nations Statistics Division - Demographic and Social Statistics [Internet]. Accessed January 19, 2022. https://unstats.un.org/unsd/demographic/sources/census/census3.htm
43. Perez AD, Hirschman C. The changing racial and ethnic composition of the US population: emerging American identities. *Popul Dev Rev*. 2009;35(1):1-51.
44. Gennaro S, Fantasia HC, Keshinover T, Garry D, Wilcox W, Uppal E. Racial and ethnic identity in nursing research. *Nurs Outlook*. 2013;61(3):174-180.
45. Aspinall PJ. What kind of mixed race/ethnicity data is needed for the 2020/21 global population census round: the cases of the UK, USA, and Canada. *Ethn Racial Stud*. 2018;41(11):1990-2008.
46. Inspection Générale des Affaires Sociales. Evolution des comités de protection des personnes (CPP) évaluant les projets de recherches impliquant la personne humaine, après la loi "Jardé" du 5 mars 2012 [Internet]. Accessed January 19, 2022. https://www.igas.gouv.fr/spip.php?article379
47. Guerrier G, Piot M-A, Resiere D, et al. Adequate clinical practice limited by the ethnic French taboo. *Lancet*. 2017;389(10085):2189-2190.
48. Megyesi MS, Hunt LM, Brody H. A critical review of racial/ethnic variables in osteoporosis and bone density research. *Osteoporos Int*. 2011;22(6):1669-1679.
49. Anderson W. Teaching "race" at medical school: social scientists on the margin. *Soc Stud Sci*. 2008;38(5):785-800.
50. Kaplan JB, Bennett T. Use of race and ethnicity in biomedical publication. *JAMA*. 2003;289(20):2709-2716.
51. European Medicines Agency. ICH E5 (R1) Ethnic factors in the acceptability of foreign clinical data [Internet]. 2018; Accessed January 19, 2022. https://www.ema.europa.eu/en/ich-e5-r1-ethnic-factors-acceptability-foreign-clinical-data
52. Rosendaal FR, Reitsma PH. Race. *J Thromb Haemost*. 2017;15(6):1049.
53. Kreatsoulas C, Anand SS. The impact of social determinants on cardiovascular disease. *Can J Cardiol*. 2010;26:8C-13C.
54. Burchard EG, Ziv E, Coyle N, et al. The importance of race and ethnic background in biomedical research and clinical practice. *N Engl J Med*. 2003;348(12):1170-1175.
55. Eisenhower A, Suyemoto K, Lucchese F, Canenguez K. "Which box should I check?": examining standard check box approaches to measuring race and ethnicity. *Health Serv Res*. 2014;49(3):1034-1055.
56. Fullwiley D. The biologistical construction of race: "admixture" technology and the new genetic medicine. *Soc Stud Sci*. 2008;38(5):695-735.
57. Fujimura JH, Rajagopalan R. Different differences: the use of 'genetic ancestry' versus race in biomedical human genetic research. *Soc Stud Sci*. 2011;41(1):5-30.
58. Ellison GTH. 'Population profiling' and public health risk: when and how should we use race/ethnicity? *Crit Public Health*. 2005;15(1):65-74.
59. Hollenbach JA, Saperstein A, Albrecht M, et al. Race, ethnicity and ancestry in unrelated transplant matching for the National Marrow Donor Program: a comparison of multiple forms of self-identification with genetics. *PLoS One*. 2015;10(8):e0135960.
60. Mathieson I, Scally A. What is ancestry? *PLoS Genet*. 2020;16(3):e1008624.
61. Zhang F, Finkelstein J. Inconsistency in race and ethnic classification in pharmacogenetics studies and its potential clinical implications. *Pharmgenomics Pers Med*. 2019;12:107-123.
62. Candille SI, Absher DM, Beleza S, et al. Genome-wide association studies of quantitatively measured skin, hair, and eye pigmentation in four European populations. *PLoS One*. 2012;7(10):e48294.
63. Itan Y, Powell A, Beaumont MA, Burger J, Thomas MG. The origins of lactase persistence in Europe. *PLoS Comput Biol*. 2009;5(8):e1000491.
64. Eberhard D, Simons G, Fenning C. *Languages of the World [Internet]*. SIL International; 2022Accessed January 19, 2022. https://www.ethnologue.com/.
65. 1000 Genomes Project Consortium, Auton A, Brooks LD, et al. A global reference for human genetic variation. *Nature*. 2015;526(7571):68-74.
66. Burhansstipanov L, Satter DE. Office of Management and Budget racial categories and implications for American Indians and Alaska natives. *Am J Public Health*. 2000;90(11):1720-1723.
67. Popejoy AB. Too many scientists still say Caucasian. *Nature*. 2021;596(7873):463.
68. Relethford JH. Hemispheric difference in human skin color. *Am J Phys Anthropol*. 1997;104(4):449-457.
69. Kamariza M, Crawford L, Jones D, Finucane H. Misuse of the term 'trans-ethnic' in genomics research. *Nat Genet*. 2021;53(11):1520-1521.
70. Hahn RA, Mulinare J, Teutsch SM. Inconsistencies in coding of race and ethnicity between birth and death in US infants. A new look at infant mortality, 1983 through 1985. *JAMA*. 1992;267(2):259-263.
71. Penner AM, Saperstein A. How social status shapes race. *Proc Natl Acad Sci U S A*. 2008;105(50):19628-19630.
72. Hunt LM, Megyesi MS. Genes, race and research ethics: who's minding the store? *J Med Ethics*. 2008;34(6):495-500.
73. Marschke G, Nunez A, Weinberg BA, Yu H. Last place? The intersection of ethnicity, gender, and race in biomedical. *AEA Pap Proc*. 2018;108(5):222-227.
74. Baer RD, Arteaga E, Dyer K, et al. Concepts of race and ethnicity among health researchers: patterns and implications. *Ethn Health*. 2013;18(2):211-225.
75. Braun L, Wolfgang M, Dickersin K. Defining race/ethnicity and explaining difference in research studies on lung function. *Eur Respir J*. 2013;41(6):1362-1370.
76. Beck AN, Finch BK, Lin S-F, Hummer RA, Masters RK. Racial disparities in self-rated health: trends, explanatory factors, and the changing role of socio-demographics. *Soc Sci Med*. 2014;104:163-177.

77. Nazroo J, Jackson J, Karlsen S, Torres M. The black diaspora and health inequalities in the US and England: does where you go and how you get there make a difference? *Sociol Health Illn*. 2007;29(6):811-830.
78. Krieger N. Living and dying at the crossroads: racism, embodiment, and why theory is essential for a public health of consequence. *Am J Public Health*. 2016;106(5):832-833.
79. Cooper RS, Wolf-Maier K, Luke A, et al. An international comparative study of blood pressure in populations of European vs. African descent. *BMC Med*. 2005;3(1):1-8.
80. King H, Rewers M. Global estimates for prevalence of diabetes mellitus and impaired glucose tolerance in adults. WHO Ad hoc diabetes reporting group. *Diabetes Care*. 1993;16(1):157-177.
81. Hall J. The unexamined "Caucasian". *Nat Genet*. 2004;36(6):541.
82. Sankar P, Cho MK, Mountain J. Race and ethnicity in genetic research. *Am J Med Genet A*. 2007;143(9):961-970.
83. Nunes K, Aguiar VRC, Silva M, et al. How ancestry influences the chances of finding unrelated donors: an investigation in admixed Brazilians [internet]. *Front Immunol*. 2020;11:584950. doi:10.3389/fimmu.2020.584950
84. Maróstica AS, Nunes K, Castelli EC, et al. How HLA diversity is apportioned: influence of selection and relevance to transplantation. *Philos Trans R Soc Lond, B, Biol Sci*. 2022;377(1852):20200420.
85. Robinson J, Barker DJ, Georgiou X, Cooper MA, Flicek P, Marsh SGE. IPD-IMGT/HLA Database. *Nucleic Acids Res*. 2020;48(D1):D948-D955.
86. Tiercy J-M. How to select the best available related or unrelated donor of hematopoietic stem cells? *Haematologica*. 2016;101(6):680-687.
87. Yudell M, Roberts D, DeSalle R, Tishkoff S. Taking race out of human genetics. *Science*. 2016;351(6273):564-565.
88. Stewart C, Pepper MS. Cystic fibrosis in the African diaspora. *Ann Am Thorac Soc*. 2017;14(1):1-7.
89. Holden C. Peering under the hood of Africa's runners. *Science*. 2004;305(5684):637-639.
90. Ségurel L, Bon C. On the evolution of lactase persistence in humans. *Annu Rev Genomics Hum Genet*. 2017;18:297-319.
91. Vicente M, Priehodová E, Diallo I, et al. Population history and genetic adaptation of the Fulani nomads: inferences from genome-wide data and the lactase persistence trait. *BMC Genomics*. 2019;20(1):915.
92. Gerbault P, Moret C, Currat M, Sanchez-Mazas A. Impact of selection and demography on the diffusion of lactase persistence. *PLoS One*. 2009;4(7):e6369.
93. Tang H, Barsh GS. Skin color variation in Africa. *Science*. 2017;358(6365):867-868.
94. Sulem P, Gudbjartsson DF, Stacey SN, et al. Genetic determinants of hair, eye and skin pigmentation in Europeans. *Nat Genet*. 2007;39(12):1443-1452.
95. Lanktree MB, Guo Y, Murtaza M, et al. Meta-analysis of dense Genecentric association studies reveals common and uncommon variants associated with height. *Am J Hum Genet*. 2011;88(1):6-18.
96. Goldstein DB. Common genetic variation and human traits. *N Engl J Med*. 2009;360(17):1696-1698.
97. Mountain JL, Risch N. Assessing genetic contributions to phenotypic differences among "racial" and "ethnic" groups. *Nat Genet*. 2004;36(11 Suppl):S48-S53.
98. 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(6):621-628.
99. Shah RR, Gaedigk A. Precision medicine: does ethnicity information complement genotype-based prescribing decisions? *Ther Adv Drug Saf*. 2018;9(1):45-62.
100. Bonham VL. Are Used in Biomedical Research. 2018;3.
101. ICMJE | Recommendations [Internet]. Accessed January 19, 2022. http://www.icmje.org/recommendations/
102. Census, race and science. *Nat Genet*. 2000;24(2):97-98.
103. Geller SE, Koch AR, Roesch P, Filut A, Hallgren E, Carnes M. The more things change, the more they stay the same: a study to evaluate compliance with inclusion and assessment of women and minorities in randomized controlled trials. *Acad Med*. 2018;93(4):630-635.
104. Brown M. Defining human differences in biomedicine. *PLoS Med*. 2007;4(9):e288.
105. The EQUATOR Network. Reporting guidelines [Internet]. [date unknown];Accessed January 19, 2022. https://www.equator-network.org/reporting-guidelines/
106. NIH Policy and Guidelines on The Inclusion of Women and Minorities as Subjects in Clinical Research | grants.nih.gov [Internet]. Accessed January 19, 2022. https://grants.nih.gov/policy/inclusion/women-and-minorities/guidelines.htm
107. Fontanarosa PB, Bauchner H. Race, ancestry, and medical research. *JAMA*. 2018;320(15):1539-1540.
108. Freeman BD, Kennedy CR, Frankel HL, et al. Ethical considerations in the collection of genetic data from critically ill patients: what do published studies reveal about potential directions for empirical ethics research? *Pharmacogenomics J*. 2010;10(2):77-85.
109. Anderson LM, Adeney KL, Shinn C, et al. Community coalition-driven interventions to reduce health disparities among racial and ethnic minority populations. *Cochrane Database Syst Rev*. 2015;6:CD009905.
110. Charrow A, Xia FD, Joyce C, Mostaghimi A. Diversity in dermatology clinical trials: a systematic review. *JAMA Dermatol*. 2017;153(2):193.
111. Fullerton SM, Yu J-H, Crouch J, Fryer-Edwards K, Burke W. Population description and its role in the interpretation of genetic association. *Hum Genet*. 2010;127(5):563-572.
112. Risch N, Burchard E, Ziv E, Tang H. Categorization of humans in biomedical research: genes, race and disease. *Genome Biol*. 2002;3(7):comment2007.
113. Outram SM, Ellison GTH. Improving the use of race and ethnicity in genetic research: a survey of instructions to authors in genetics journals. *Sci Ed*. 2006;29(3):78-81.
114. Lewis ACF, Molina SJ, Appelbaum PS, et al. Getting genetic ancestry right for science and society. *Science*. 2022;376(6590):250-252.
115. Burke JF, Sussman JB, Kent DM, Hayward RA. Three simple rules to ensure reasonably credible subgroup analyses. *BMJ*. 2015;351:h5651.
116. Tanniou J, van der Tweel I, Teerenstra S, Roes KCB. Subgroup analyses in confirmatory clinical trials: time

to be specific about their purposes. *BMC Med Res Methodol*. 2016;16(1):20.
117. Flanagin A, Frey T, Christiansen SL. AMA manual of style committee. Updated guidance on the reporting of race and ethnicity in medical and science journals. *JAMA*. 2021;326(7):621-627.
118. L'Assemblée supprime de la Constitution le mot «race» et interdit la « distinction de sexe » [Internet]. Le Monde.fr 2018;Accessed January 19, 2022. https://www.lemonde.fr/politique/article/2018/07/12/l-assemblee-supprime-dans-la-constitution-le-mot-race-et-interdit-la-distinction-de-sexe_5330615_823448.html

How to cite this article: Gombault C, Grenet G, Segurel L, et al. Population designations in biomedical research: Limitations and perspectives. *HLA*. 2023;101(1):3-15. doi:10.1111/tan.14852