Epistemological Pitfalls in the Proxy Theory of Race: The Case of Genomics-Based Medicine

The British Journal for the Philosophy of Science
https://doi.org/10.1086/727957

Joanna Karolina Malinowska
ORCID: 0000-0002-8330-5234

Davide Serpico
ORCID: 0000-0003-3997-3056

Abstract

In this article, we discuss epistemological limitations relating to the use of ethnoracial categories in biomedical research as devised by the Office of Management and Budget's institutional guidelines. We argue that the obligation to use ethnoracial categories in genomics research should be abandoned. First, we outline how conceptual imprecision in the definition of ethnoracial categories can generate epistemic uncertainty in medical research and practice. Second, we focus on the use of ethnoracial categories in medical genetics, particularly genomics-based precision medicine, where ethnoracial identity is understood as a proxy for medically relevant differences among individuals. Notably, extensive criticisms have been made already against the genetic interpretation of races, but, nonetheless, the concept of race remains a key element of contemporary genomics. This motivates us to explore possible reasons why such criticisms may have been ineffective in redirecting attention to other (non-race-based) ways of controlling for human variability. We contend that popular arguments against the idea that human races have a genetic basis, though convincing in many respects, are not sufficient to exclude the pragmatic use of race and ethnicity as proxies for genetic variability related to complex phenotypes. Finally, we provide two further arguments to support the idea that ethnoracial categories are unlikely to provide meaningful insights into medical genetics, which implies that even the interpretation of race as a useful tool to stratify disease risk is unwarranted.

Keywords: race, ethnicity, ethnoracial categories, reference class, genomics, P-medicine

Table of Contents:

1. Introduction

In principle, each person can be assigned to an infinite number of categories depending on the characteristics that are selected as relevant, such as age, height, sex, gender, eye colour, occupation, favourite breakfast cereals, and many others. As Venn notes:

> Every individual thing or event has an indefinite number of properties or attributes observable in it, and might therefore be considered as belonging to an indefinite number of different classes of things [. . .] This variety of classes to which the individual may be referred owing to his possession of a multiplicity of attributes, has an important bearing on the process of inference. (Venn [2019], p. 196)

In biomedical research, these categories are often called 'subgroups' or 'reference classes.' Constructing reference classes for research with human subjects ultimately relies on many evaluative decisions as well as on the historical, social, and institutional contexts in which researchers are embedded (Ludwig [2014], [2016]; Reydon and Ereshefsky [2022]). However, some suggest otherwise—that the choice of reference classes can be justified only by 'natural facts' (Boorse [1977]; Veit [2021]) or strictly epistemic purposes (Khalidi [2013]).

Scientists agree that the many categories into which people can be divided, such as their favourite actress or having a fringe, are irrelevant to most biomedical research. Some are recognized as significant, especially when they are associated with some biological markers or processes, such as ageing. However, there are also categories the use of which in biomedical

research is particularly problematic, for instance, those that involve socio-cultural factors. Some of these are categories of race and ethnicity.[1]

Ethnoracial categories are nowadays routinely used in research across epidemiology, pharmacogenomics, and genetics, for example, in randomised controlled trials (RCTs) and genome-wide association studies (GWAS). For instance, one of the aims of the National Institutes of Health (NIH) precision medicine initiative *All of Us* has been to collect data from groups that are typically underrepresented in medical research—including racial minorities (Cunningham [2018]). The motivation is that collecting data on non-white populations will help achieve a greater predictive capability and more reliable medical inferences: in other words, if we had more data about the race to which a given individual belongs, we would be able to make better predictions regarding their future health (Fatumo *et al*. [2022]). Notably, while some authors postulate that the use of ethnoracial categories in science can be epistemically beneficial (Spencer [2018a], [2018b], [2019]), in many clinical and genetic studies there is no explicit rationale for their use (Duello *et al*. [2021]; Malinowska and Żuradzki [2023a], [2023b]).

Although, for us, subtyping populations into smaller groups represents a fundamental step towards the implementation of more precise, individualized, and reliable medical decisions, we concur with the view that race and ethnicity are not good proxies for such analyses in genetics and genomics: in this sense, the use of ethnoracial categories as a heuristic concept to capture deeper, biological variability may bias empirical results and blind us to the actual symptoms presented by individual patients, their family illnesses, history, and more general social inequalities that affect health outcomes. In this article, we discuss the use of ethnoracial categories in biomedical research—particularly in medical genetics—and support the idea that such categories are unlikely to provide meaningful insights into human phenotypic and genetic variability.

In Section 2 we discuss the main reasons why ethnoracial categories play such a central role in biomedical research, which constitute what we call the 'proxy theory of race'. We delineate potential risks relating to the adoption of misleading reference classes, particularly

[1] In some countries, such as the US, the category of race is usually listed alongside ethnicity. In others, like the UK, the category of ethnicity replaces race (understood as referring to biologically relevant characteristics important in biomedical research). In genetics research, the two terms are often used interchangeably (see Duello *et al.* [2021]; Malinowska and Żuradzki [2023a]). We will refer to race and ethnicity or ethnoracial categories to denote both uses unless noted otherwise.

ethnoracial categories, and point out that the mere recommendation or requirement to collect data on racial/ethnic groups negatively affects research results.

In Section 3, we discuss how conceptual imprecision in the definition of ethnoracial categories can generate epistemic uncertainty in medical research and practice. We focus on two sources of uncertainty: first, how ethnoracial categories are conceptualized in institutional guidelines and, second, how single individuals (researchers, policymakers) tend to apply them in practice. Then, we outline the main philosophical perspectives on the use of ethnoracial categories that have emerged in response to the epistemic uncertainty above.

In Section 4, we turn to epistemological issues in the use of ethnoracial categories in medical genetics, a field where stratification biases are a pervasive, bothersome obstacle to the interpretation of the empirical data.[2] Within the vision of precision and personalized medicine (hereafter, P-medicine), contemporary genomics takes racial differences as an important factor to account for better and more precise medical inference and treatment. This is surprising given that several voices have already raised concerns about the genetic interpretation of ethnoracial categories. We explore potential reasons why such criticisms have been unable to disincentivize the use of such categories. Then, we provide additional arguments to support the idea that genetic data are unsuited to categorizing individuals in terms of stratified risk for complex diseases and to understanding 'racial differences' in medically relevant traits.

Finally, we emphasize that, for the reasons provided in Sections 3 and 4, the recommendation of institutional guidelines for the use of ethnoracial categories in medical genetics is epistemologically problematic.

Let us clarify that, in this article, we mostly focus on how the process of constructing reference classes can influence the course of research with human participants and its results in the context of medical genetics, specifically in the study of complex diseases. Indeed, accounting for the role of ethnoracial categories in every area of biomedicine would be beyond what a single article could possibly do (for some considerations on this point, see Section 3.1 below). To narrow down the discussion further, we focus on institutional guidelines adopted in the United States (US) for the use of ethnoracial categories. Yet, this problem is not unique to the US, as their regulations impact, for instance, the construction of reference classes in research conducted to receive approval from the Food and Drug Administration (FDA) for the

[2] Population stratification involves undetected heterogeneity in allele frequencies due to non-random mating and geographical isolation (Hellwege *et al*. [2018]; Lawson *et al*. [2020]).

distribution of certain products. For example, it has been established that there is a spillover of the US regulatory standards to the European Union (EU) and that significantly more EU than US pharmaceutical product labels report ethnoracial differences in drug responses (Mulinari *et al*. [2021]). We consider justifications and limitations for using ethnoracial categories in accordance with the recommendations of the FDA ([2016]) and the NIH ([2001]), which are both based on the classification provided by the Office of Management and Budget (OMB [1997]). Such institutions recommend that individuals self-identify their race and ethnicity, except in cases where, for example, an 'observer identification is more practical' when completing a death certificate.

2. Race as a Medically Relevant Proxy

Two main ideas are usually cited to support the use of ethnoracial categories in biomedicine. First, race is widely regarded as an important risk factor for a variety of diseases and conditions. To make sense of the statistical association between variation in ethnoracial identity and variation in medical conditions, scholars often consider race and ethnicity as a 'proxy' for other variables that are medically interesting or relevant for the susceptibility of a disease, such as physiological, genetic, or psychological characteristics. Race is thus used to correct stratification biases in clinical trials as well as genomics and pharmacogenomics studies (see footnote #2). This position usually goes in line with some form of biologization of this category, but the endorsement of a proxy theory of race does not automatically force us to accept some form of racial realism (for example, that the OMB racial categories are 'biologically real'). It implies only that medically relevant factors are distributed differently across human groups (such as those delineated by the OMB categorization) and that such distribution has some degree of consistency (see Section 3.1).

Second, gathering ethnoracial data and using it in research seems to be necessary, for example, to study social inequalities and, with the knowledge gained, reduce them over time (such a position usually interprets race and ethnicity as nonbiological, social kinds). Ethnoracial affiliations in this case represent a sort of idealization of the collective experience of racialized individuals. They are constructed to 'monitor equal access in housing, education, employment, and other areas, for populations that historically had experienced discrimination and differential treatment because of their race or ethnicity' (OMB [1997], p. 1). Thus, they serve

as a proxy in the analysis of how social inequalities affect people's health (Malinowska and Żuradzki [2023b]).

In certain countries, such as the US, collecting racial data is not only recommended but even required by some research funding agencies. For instance, since 2001, the NIH requires the use of racial categories to collect and report data in submissions for clinical trials, and the FDA has recommended it since 2005. The mandatory collection and use of racial data are expected to help combat social inequalities and lead to more reliable research results (for example, by reducing 'White Middle-Class American Male Discrimination' in science). But, as we show in the next sections, there is no scientific justification for using these categories as reference classes in genetic and genomic research. On the contrary, there are many reasons not to do so. One of our concerns is that automatically adding ethnoracial categories to the analysis only due to institutional guidelines increases the probability that one of these categories will be recognized as causally related to the studied intervention, while in fact, it is not. First, the result can be a statistical artefact. Second, the interpretation of the results may be completely wrong.

One of the methods of analysing data in this respect is subgroup analysis in which study samples are divided into classes of participants based on their shared characteristics. This allows researchers to understand how certain groups of people respond differently to certain interventions. In other words, it aims 'to explore whether there is evidence that the treatment difference depends on certain patient characteristics' (Pocock *et al*. [2002], p. 2917). However, in many cases, researchers analysing reference classes (denoting certain subgroups) overestimate their epistemic value. While a broad range of methods has been developed for exploratory as well as confirmatory subgroup analysis (Ondra *et al*. [2016]), these analyses still have limitations when it comes to recognizing causal relations between the analysed phenomena and the used reference classes (Rzepiński [2016], p. 88; Wallman and Williamson [2017]; Lin *et al*. [2019]). For instance, there are arguments that the results provided in the subgroup analysis have a low degree of reliability, since the formulation of hypotheses to distinguish the factors that differentiate reference classes often occurs only after completing the study (Cui *et al*. [2020]). Subgroup analysis is, in these cases, a research procedure used to obtain new hypotheses. However, there is no straightforward evidence behind this type of hypothesis 'independent' of the analysis of the subgroups that were the basis for their very formulation. In such a

situation, statistical differences underlying their articulation cannot be treated as evidence on their behalf (Rzepiński [2016], p. 90).

An example of such a situation is the famous ISIS-2 study critical analysis (intended to illustrate epistemological problems with the interpretation of its results), in which its authors chose to additionally divide population in the study into twelve subgroups according to the twelve astrological signs (Sleight [2000]; Peto [2011]). Before their intervention, analysed data indicated the general statistical benefit of taking aspirin over a placebo. After including the twelve reference classes, subgroup analysis results indicated that, for people born under the signs of Gemini and Libra, taking aspirin may not only not be beneficial but even have slight adverse effects. Because this is a fairly obvious situation, no one would take such results seriously. However, in many other cases, scientists far too readily accept the results of subgroup analysis (Sun *et al*. [2012]; Lin *et al*. [2019]; Cui *et al*. [2020]).

For categories the use of which appears to make sense—although it may not—this is particularly tricky because in this way it is possible to 'biologize' certain reference classes (that is, give them a biological significance). A classic example where the use of the category of race brought about unpalatable consequences is the case of BiDil (according to the manufacturer's description, the drug intended to 'treat heart failure in black patients'), which has been repeatedly described in the literature (Kahn [2012]; Pollock [2012]) and has had far-reaching negative social consequences. The design of the study (for example, lack of a 'non-Black' control group resulting from some administrative decisions) allowed its manufacturers to justify selling their product (BiDil) to a particular racialized group (African Americans). The FDA approval of the product for Blacks only contributed to the biologization of race—it began to function as evidence for the existence of biological differences between 'races' and inspired many companies and researchers to study these differences intensively. Moreover, it has also contributed to the development of so-called 'race marketing' (Sankar and Kahn [2005]; Sallaz [2010]; Crocket [2008]; Saha [2015]). In this way, it is easy to set off a spiral of flawed research, biased results, and capitalist (or more generally political) demand for more of it (see also Malinowska and Żuradzki [2023b]).

Notably, the adoption of misleading reference classes may have a variety of possibly fatal consequences, for example, errors in medical practice, such as misdiagnoses or erroneous exclusion of people from a given class from proper treatment or identifying certain patient characteristics as beneficial to therapy effects, while these traits do not actually affect the final

state of patients. In the next section, we reconstruct the main arguments why ethnoracial categories (especially as interpreted by the FDA and NIH) are extremely imprecise and vague concepts (and very lousy reference classes), which makes it difficult to conduct reliable research.

3. The Imprecise Construction of Ethnoracial Reference Classes

Let us start with a very basic and naïve question: how is it possible that the use of official classifications can lead to the imprecise construction of reference classes? After all, the classification is imposed by institutional guidelines, and therefore 'ready', that is, already constructed. However, when it comes to their conceptualization, such guidelines have been evaluated as imprecise (Meissner [2021]), diverse, and unreliable—their rationale is being challenged both ethically (Zack [2016]) and conceptually (Hochman [2021b]; Jackson [2022]; Winsberg [2022]). In this section, we demonstrate that ambiguities and conceptual imprecision in the definition of ethnoracial categories can generate major inconsistencies and uncertainty in medical studies. We focus on two major sources of uncertainty: first, how ethnoracial categories are conceptualized by institutions and, second, how single individuals (researchers, policymakers) tend to apply them in practice.

Let us start with how ethnoracial categories are conceptualized by institutions in the first place. Both the FDA and the NIH recommend the typology endorsed by the OMB and categorize study participants into at least five racial groups (American Indian or Alaska Native, Asian, Black, or African American, Native Hawaiian or Other Pacific Islander, and White) and two ethnic categories (Hispanic or Latino and Not Hispanic or Latino).[3] However, while the NIH defines racial categories as a mainly socio-political construct that 'should not be interpreted as anthropological in nature', the FDA ([2016]) recognizes races as populations whose representatives have common ancestors inhabiting specific geographical areas (in this sense, the category may be intended as a proxy for a genetic lineage) and whose differences in health with other populations may be additionally caused by external factors (see also Malinowska and Żuradzki [2023b]).

[3] The OMB classification differs substantially from other countries' censuses, such as those adopted in Brazil, the UK, and New Zealand (Valles [2016]). In addition to governmental institutions, scientific journals like JAMA (Flanagin *et al*. [2021]) and MDM (Zikmund-Fisher [2022]) develop their own recommendations on how to report ethnoracial categories in articles. Their interpretation of these categories is often in conflict with the FDA's perspective: JAMA and MDM interpret race and ethnicity as a strictly socio-cultural (or socio-political) construct. In this article, we focus on the OMB, FDA, and NIH recommendations due to their influence on biomedical studies.

Since in the main institutional guidelines race is, at times, recognized as a purely social construct and, in other cases, as a biologically justified category, reference to such a concept may have entirely different meanings across guidelines and articles on the very same subject (Huddart *et al*. [2019]; Byeon *et al*. [2021]; Malinowska and Żuradzki [2023a], [2023b]). Even in the document about the OMB standards for reporting race and ethnicity, there is an enigmatic (or pluralist, Jackson [2022]) provision that 'the racial and ethnic categories set forth in the standards should not be interpreted as being primarily biological or genetic in reference. Race and ethnicity may be thought of in terms of social and cultural characteristics as well as ancestry' (OMB [1997], p. 1). This formulation gives room for a wide variety of interpretations of what race is and how important the role of biological factors is in determining the clustering of humans into ethnoracial groups (while it is not a 'primarily biological or genetic' characteristic, there is no clear statement there that it is not biological at all, and it still may refer to ancestry).

Uncertainty is also related to the fact that researchers are often not sure what they refer to when they use ethnoracial categories. In a recent study, more than two-thirds of scientists (most were professional geneticists) were not confident in their ability to distinguish between the terms 'race,' 'ethnicity,' and 'ancestry' (Popejoy *et al*. [2020], p. 71). This problem is also visible in scientific publications. Text analysis of articles on COVID-19 (Malinowska and Żuradzki [2023a]) indicated that their authors used terms the 'race' and 'ethnicity' in at least five different ways (folk, demogeographic, socio-cultural, multileveled, and institutional). Moreover, while some of them conceptually divided races from ethnicities, others used the terms 'race' and 'ethnicity' interchangeably or referred to ethnicity (usually interpreted as a cultural category) in terms of genetic lineage. Finally, scientists use very different ethnoracial classifications depending, among other things, on their research goals, the available data, and the cultural contexts in which they operate (López *et al*. [2017]; Huddart *et al*. [2019]; Zhang and Finkelstein [2019]). Thus, ethnoracial categories resemble some mythical creatures—all have some hazy idea about them, but no one precisely knows what they are.

Such conceptual inaccuracies hinder scientific communication and prevent, for instance, comparative analyses of different studies. Yet, researchers sometimes treat their country's racial/ethnic categories as universal, objective, and scientifically well-grounded. For instance, some researchers who examined the heritability of IQ across 'racial' or ethnic groups in a meta-analysis (published in the journal *Intelligence* in 2020) were surprised to learn that, although they did not confine their search exclusively to the US, most of their samples came

from there. Although they attempted to explain this in different ways (homogeneity, lack of biometric research elsewhere, etc.), they failed to mention an easier explanation: implicitly, they used the US racial categorization based on the OMB policy directive, which is not a universally accepted categorization (Davenport [2020]). This example shows that some researchers are only interested in ethnoracial classifications accepted in their social environment (compare with Spencer [2018a], [2018b]) and may not realize that ethnoracial ontologies are fashioned conventionally—or even in legal terms—rather than discovered through scientific investigation (Winther and Kaplan [2013]). They may also not realize that theoretical assumptions (including the fundamental classifications they are using) are influencing their research results under the principle that 'what you put in is what you get out'.

Overall, institutional classifications, like the OMB's, involve a high degree of arbitrariness and conventionality: they are built on folk racial beliefs and reasoning (based on people's appearance and shaped by socio-political conditions) popular in specific countries, such as the US (Haslanger [2019]; Hochman [2021b]; Winsberg [2022]). These folk beliefs are supposed to correspond to the so-called continental populations (when it comes to the biological interpretation of 'race') or to some universal experiences (when it comes to its socio-cultural interpretation). Moreover, people's ethnoracial identities are context-dependent constructs, and there are also great differences in reporting these identities by research subjects. Thus, the OMB categorization lays claim to objectivity, but, in fact, it is strictly limited to the US context, representing only one of many possible existing ethnoracial classifications (Ludwig [2019]; Hochman [2021b]), and its use is greatly influenced by the researchers' beliefs. That leads to significant differences in gathering, reporting, and interpreting ethnoracial data between scientists.

3.1 Perspectives on the Use of Ethnoracial Categories in Biomedicine

At the crossroads of the epistemic issues discussed above, the philosophical literature has matured two main conceptual perspectives on the use of ethnoracial categories in science: 'conservationism' and 'eliminativism' (Mallon [2006]; James and Burgos [2022]). The former covers various theoretical perspectives: those who point to some form of racial realism (biological or socio-cultural) and those who allow ethnoracial categories to be used in research merely as a useful conceptual tool (some kind of racial pragmatism).

Let us first consider the two forms of racial realism. Briefly, biological realism on race is based on the assumption that it is possible to design a stable and scientifically meaningful taxonomic system dividing people into five populations that correspond to the five human races from the OMB classification (Spencer [2018a], [2018b], [2019]). Socio-cultural realism on race assumes that while biological human races do not exist, races understood as some social kind (socio-cultural constructs) exist and that ethnoracial categories should be used in science to describe the experiences of racialized people (for example, the experience of racism).

Yet, there is an increasing number of criticisms against realist positions, which can be categorized as various forms of antirealism and eliminativism. These also include the argument that biological realism is based on serious logical fallacies (Hochman [2013], [2021b]). When it comes to racial social realism (also known as social constructionism), some scholars argue, for example, that most of its versions are imprecise and even compatible with racial hereditarianism (Hochman [2022]). Moreover, while the use of ethnoracial categories in science is often aimed at decreasing social inequalities, in reality, such practices may have an opposite effect: they may reinforce prejudicial racial beliefs (Hochman [2019], [2021a]). Racialized individuals may have different life experiences, different access to education, and may live in different environments: unifying them into coarse-grained categories not only leads to unreliable research but also to stereotyping and essentializing race and ethnicity (Malinowska [2021]; Meissner [2021]; see also Malinowska [2016]). In biomedical terms, there seems to be no scientific justification for homogenizing the experiences of, for instance, Black people who have lived in the US for generations (those living in poor neighbourhoods as well as the wealthiest residents of, for example, Atlanta) and recent migrants to the States (Kuzawa and Sweet [2009]; Valles [2012]; Kalewold [2020])—not to mention recent refugees from some African countries, who are stuck in the forests or closed centres for illegal migrants on the Polish-Belarusian border, if we expand the OMB classification beyond the US. However, if that is the case, how then should we study health inequalities resulting from racism?

According to many scholars, medical practitioners, and institutional guidelines, regardless of their ontological status, socially interpreted categories of race and ethnicity are useful proxies for other social determinants of health that are related to racial injustice, such as socioeconomic status and access to healthcare systems (NIH [2001]; FDA [2016]; Centers for Disease Control and Prevention [2021]). As noted by Valles ([2021a]), for instance: "race-associated risk offers a summary measure of features of how patients, on average, interact with their

social worlds". In fact, there is evidence, from various epidemiological sources, of inequalities in morbidity and mortality between racially defined populations (Devakumar *et al.* [2022]; Gravlee [2009]; Krieger [2021]; Selvarajah *et al.* [2022]). Yet, these disparities consist of many different factors that require detailed identification and to which not all people with a particular racial identity are exposed (Malinowska and Żuradzki [2023b]; Meissner [2021]; Valles [2021a]).

We fully agree with the view that there are many pathways through which racism (and other forms of discrimination) affects people's health (Abubakar *et al.* [2022]; Devakumar *et al.* [2022]; Krieger [2019], [2021]; Selvarajah *et al.* [2022]; Shannon *et al.* [2022]; Sullivan [2013]; Valles [2021a], [2021b]; Williams *et al.* [2016]), and that these pathways need to be carefully recognized and studied. However, we also share concerns that interpreting, collecting, and using ethnoracial data—especially as recommended by the FDA and the NIH—is not as precise and reliable enough to accomplish that goal (on the same point, see Malinowska and Żuradzki [2023a], [2023b]; Meissner [2021]). Among other factors, this is due to the fact that racialization processes are highly dynamic and contextual (Chellappoo and Baedke [2023]), and can thus affect one's health on many different levels (Malinowska and Żuradzki [2023b]), which cannot satisfactorily be accounted for by heterogeneous and coarse-grained categories like race and ethnicity. These categories seem to us unable to reflect such complexity, and therefore their application to study the effects of racialization on health can only have limited utility. For instance, they can be reasonably used for specific research questions (such as psychological analysis of the impact of the US residents' racial identities on their psychological and physiological wellbeing), but they do not cover all groups that may experience racialization and racism, nor do they recognize the full spectrum of racialization processes and their potential impact on health.

For instance, a growing amount of research on racism is directed at characteristics such as accent, religion, and nationality (Krivonos [2023]; Lewicki [2023]; Narkowicz [2023]; Rzepnikowska [2023]; Tereshchenko *et al.* [2019]). Those who experience these forms of racialization are not only people of color, but also whites, who are not "white enough" in certain contexts (Kalmar [2022], [2023]), especially to other whites. One of such groups are Eastern Europeans. In particular, those who migrate from east to west commonly experience systemic exclusion and exploitation, as well as racial aggression: marginalization, insults, harassment, threats, and physical assaults (Krivonos [2023]; Lewicki [2023]; Narkowicz [2023];

Rzepnikowska [2023]). In studies where race and ethnicity are the only proxies for social inequality, similar forms of racialization tend to remain invisible. Thus, researchers should be encouraged to critically assess whether ethnoracial classifications operationalize well enough the processes of racialization they plan to study, as well as to report socioeconomic status and deprivation using disadvantage indexes or other measures that are better suited for their scientific goals (Malinowska and Żuradzki [2023a], [2023b]). However, a thorough discussion of this issue is beyond the scope of our article.

Our main focus, here, is biological racial pragmatism (also known as racial pragmatist naturalism), which more closely aligns with the proxy theory of race, especially in the context of genomics. In its common form, racial pragmatism only implies that geographically based polymorphisms that contribute to disease incidence can be examined with the use of the OMB racial classification (or other ethnoracial typologies) purely for pragmatic reasons and without entering a discussion about the existence of human races (Rosenberg *et al*. [2002]; compare with Jackson [2022]).

Although racial pragmatism is quite common in science, it is increasingly apparent that any ethnoracial classifications are not reflective of global genetic variation (Huddart *et al*. [2019]; Popejoy *et al*. [2020]; Jackson [2022]). Moreover, the use of self-reported racial affiliations in genetics research is simply misleading and unjustified (Kuzawa and Gravlee [2016]; Borrell *et al*. [2021]). In many cases, culturally constructed ethnoracial identities are difficult to track, as they are dependent on many factors such as political, historical, economic, or ecological contexts (Meissner [2021]; Ludwig [2019]). A person ascribed to one class from a political perspective can be classified into another from the perspective of the prevailing cultural stereotypes. Sometimes, depending on legal regulations (or even methods for filling in medical or legal documentation), one can be assigned to different subgroups in one place or another, or (independently from the official institutional classifications) change their ethnoracial identity, for example, due to migration (Keskinen and Andreassen [2017]; Grill [2018]). In other words, while one cannot deliberately choose their genetic makeup, one can choose or change their ethnoracial identity. There is also a growing number of people who identify themselves as multiracial and those who do not even know that they have ancestors from a few different populations.

Yet, while the above criticisms of the use of self-described ethnoracial data in genetics and genomics are usually well received among scholars, their application in these fields was

not disincentivized but is rather routinely recommended by institutional guidelines. Moreover, the proxy theory of race is still valued for handling stratification issues, which are thought to limit the "portability" of genomics findings and reduce the predictive capability of PRSs across populations (Johnston and Matthews [2022]; Matthews [2021]). For many scholars, such issues suggest the need for a higher inclusion of ethnoracial minorities in genetics research (e.g., Boem [2022]; Fatumo *et al.* [2022]; Martin *et al.* [2019]; Palk *et al*. [2019]). On this view, the obligation to use such categories in genomics would ameliorate healthcare disparities.[4]

In the next section, we focus on the use of the OMB racial classification in medical genetics. We examine in more detail the available arguments against a genetic interpretation of the concepts of race and ethnicity and explore the possible reasons why such arguments may have been ineffective in redirecting attention to other (non-race-based) ways for controlling for human variability. We then provide two further—and hopefully more effective—arguments to support the idea that ethnoracial categories are unlikely to provide meaningful insights in biomedical studies, which implies that even their interpretation as a mere pragmatic tool is unsound.

4. Unreliable Applications of Ethnoracial Categories: The Case of Medical Genetics

One of the most problematic uses of the OMB racial classification comes with the adoption of race in genetics and genomics, where races are usually seen 'as a result of human migration with genetic isolation leading to the development of distinct populations that share DNA as the result of common descent' (Duello *et al*. [2021]; see also Dobzhansky [1937], p. 138). As we mentioned in Section 2, the main argument justifying the application of ethnoracial categories in this context is that they are allegedly a reliable indicator of 'something biologically real' (Spencer [2019], p. 76) and can serve as a pragmatic proxy in biomedical research to predict

[4] The use of the category of race in genetic and genomic research has declined in recent years (Byeon *et al*. [2021]; Malinowska and Żuradzki [2023a]), for example, due to its replacement by the category of ethnicity understood as a proxy for a genetic line or the category of ancestry—a disturbing phenomenon because not only does it not solve the problem with the biologization of the category of race, but it may additionally lead to the biologization of the category of ethnicity. Moreover, this does not acknowledge that ancestry is a substantially different concept to race and ethnicity. While the former represents the genetic origin of one's population (and it is thus potentially a better predictor for genetic polymorphisms), ethnicity and race are 'identities' that are self or socially ascribed (Borrell *et al*. [2021]) based on superficial characteristics. As such, these categories originate from the secular processes of racialisation (Hochman [2017], [2019]; Malinowska and Żuradzki [2023b])—which is not reducible to geographical and genetic parameters.

phenotypic variability based on genetic differences (for example, variability in disease aetiology and response to treatment).

Although many scholars and scientific societies criticized the genetic interpretation of ethnoracial categories, there seem to be durable beliefs, among scientists and policymakers, that genetic variation follows some consistent patterns that map 'ethnoracial divisions' as understood by the OMB. Such beliefs may seem to justify the pragmatic use of race and ethnicity as a proxy for genetic and phenotypic variability.

In what follows, we analyse existing popular arguments against the biologization of race and ethnicity in genetics and genomics. Although such arguments are convincing in many respects, we examine potential reasons why they may have been ineffective in redirecting attention to other (non-race-based) ways for controlling for human variability. This would explain why racial pragmatism and the proxy theory persist and race and ethnicity remain key concepts in contemporary genomics despite decades of controversies (Bliss [2020]; Duello *et al*. [2021]). We then defend two further—and hopefully more effective—arguments to support the idea that ethnoracial categories are unlikely to provide meaningful insights in genetic and genomic studies, which implies that even their interpretation as a mere pragmatic tool is unsound.

4.1 Why Previous Criticisms May Be Ineffective

A widely discussed criticism of the genetic interpretation of ethnoracial categories connects to Richard Lewontin's ([1972]) influential argument that, considering any single genetic locus, there is on average more genetic variation 'within' human racial groups than 'between' them. As a 2001 *Nature* editorial (cited in Edwards [2003], p. 798) explains:

> This means that two random individuals from any one group are almost as different as any two random individuals from the entire world. Although it may be easy to observe distinct external differences between groups of people, it is more difficult to distinguish such groups genetically, since most genetic variation is found within all groups. (Edwards [2003], p. 798)

Lewontin's reasoning was well received by many scholars and scientific societies (Barbujani *et al*. [1997]; Rosenberg *et al*. [2002]) and became an important foundation for the idea that human variability is continuously distributed across human populations and it is thus

impossible to make clear-cut distinctions of any sort.[5] This point was recently reiterated by the statement of the American Society of Human Genetics (ASHG [2018], p. 636) denouncing the use of genetic data in racial suprematism discourses:

> The study of human genetics challenges the traditional concept of different races of humans as biologically separate and distinct. [. . .] Most human genetic variation is distributed as a gradient, so distinct boundaries between population groups cannot be accurately assigned. [. . .] There is considerable genetic overlap among members of different populations. Such patterns of genome variation are explained by patterns of migration and mixing of different populations throughout human history. In this way, genetics exposes the concept of 'racial purity' as scientifically meaningless. (ASHG [2018], p. 636)

As early as 2003, Edward criticized Lewontin's argument and described it as a 'fallacy' (which is, in fact, known as 'Lewontin's Fallacy'). From this view, considering variation at one single locus does not do justice to the fact that 'most of the information that distinguishes populations is hidden in the correlation structure of the data and not simply in the variation of the individual factors' (Edwards [2003], p. 798). In other words, Edwards argued, if we consider 'correlations among many loci' through principal component analysis, for instance, the correlations reveal legitimate ways to categorize people into racial groups.

Interestingly, the recognition that some correlational structure in the distribution of genetic risk across 'races' is accepted widely. The premise in the opening of the ASHG statement ([2018], p. 636), for instance, says:

> 'Although there are clear observable correlations between variation in the human genome and how individuals identify by race', the study of human genetics challenges the traditional concept of different races of humans as biologically separate and distinct. (our emphasis) (ASHG [2018], p. 636)

This recognition could already represent a potential reason why some scholars may agree that human variability is continuously distributed across populations (and is thus impossible to make clear-cut distinctions among 'races') but still try to identify associations between variation at the genetic level and the membership of different ethnoracial groups.

[5] Notably, in ethical terms, this is a very desirable conception of human variability. Similar considerations are made in discussions on the distinction between mental health and pathology, particularly in the context of autism and ADHD (Koi [2021]). Recently, Burt ([2022], p. 7) made the similar point that 'these populations are abstractions from an underlying continuum of genetic relatedness and *should not* be thought of as genetically distinct subpopulations'.

To discuss a second potential reason, we need to consider another type of criticism of the genetic interpretation of races, namely, criticism of racial realism.

Hochman ([2021b]) has recently reconsidered Edwards' argument and conceded the possibility of clustering individuals into groups based on genetic data (in particular, he points to various examples of empirical evidence that many different classifications are possible from anthropometric and genetic data; see Rosenberg *et al*. [2002]; Tang *et al*. [2005]; Witherspoon *et al*. [2007]). However, he argues that such clusters do not correspond to 'conventional races'. This way, Hochman's argument directly tackles racial realism and questions the validity of race categories on the basis that not every stable and convenient scientific classification is equally valid:

> it is possible to create a stable classification system that does not pick out valid scientific categories. [. . .] It would be possible to create a stable classification system of fruit, based on the color of its skin when ripe. This system would lump strawberries with pomegranates, limes with watermelons, star fruit with quince. This classification system would be fairly stable, but it would not be of any taxonomic significance whatsoever. [. . .] If the ability to design a stable classification system were sufficient to vindicate that system as scientifically meaningful, then it would be legitimate to classify fruit based on color. [. . .] Valid scientific taxonomies need more than mere stability. They need to be meaningful in the context of the relevant science. (Hochman [2021b], pp. 78–79).

Hochman also argues that there is no privileged way to cluster humans into races: rather, depending on how we interpret correlational data (for example, what sampling scheme and resolution), different classifications may arise.

Although we find Hochman's analysis compelling, there is an important issue with these arguments: they are effective against racial realism, but they may be not strong enough to discourage a more 'pragmatic' use of races.[6]

We are afraid that this pragmatic use can 'bypass' arguments against racial realism by appealing to a non-ontological view of races, which would be exactly what a proxy theory of race needs. Indeed, one may agree with Hochman that conventional ethnoracial categories are not 'valid', meaning that they are not 'natural' divisions or that, in other words, conventional classifications are unable to 'carve nature at its joints'. One may also agree that any

[6] Another recent criticism of racial realism can be found in Winsberg ([2022]). Also in this case, the argument is extremely convincing, but similar concerns apply as to the case of Hochman's arguments (see the main text).

categorization of humans is doomed to involve conventional factors or be imprecise for various reasons (as we explain in Section 3). And yet, one can think that 'some' ethnoracial categorization could be 'useful' for addressing stratification issues in biomedicine and address questions such as: is there any pattern in the distribution of genetic variants that may help us handle human variability and population stratification? How to revise racial categories (or construe new ones) to track down genetic variants that constitute the basis of complex diseases?

To summarize, the argument that no clear-cut distinctions exist among races, as well as arguments against racial realism, may represent a weak basis to prevent the use of ethnoracial categories in genomics. For instance, Spencer ([2018a], p. 1031) states that 'the quantity of genetic differentiation among human continental populations is irrelevant to whether the genetic differentiation is important to medicine'. It is probably based on this type of reasoning, we suggest, that ethnoracial taxonomies are often considered 'sufficiently stable' classification systems to serve some scientific purpose, especially in genetics and genomics. Indeed, much of the use and institutional guidelines of races in biomedicine often rely on pragmatic concerns that bypass entirely questions about their 'reality' or questions about the 'conceptual precision' of such categories.

In other words, if we stick to the point that ethnoracial categories should not be used simply 'because races do not exist', we are unable to explain why such categories are still widely used to stratify genetic variability even though most scientists agree already (at least since the 1950s' UNESCO statement (Brattain [2007])) that races are highly idealized types. We know already that the categories of race and ethnicity are often constructed based on the assumption that, at some stage of the evolutionary history of humankind, there were geographically separated races, their representatives were 'racially pure', and had 'ideal' genotypes lacking the admixture of any 'other race' genes (Zack [2016]). And we know already that this assumption is incorrect: contemporary genomics teaches us that there is no such thing as a 'typical Asian DNA' or 'typical African DNA' (Bliss [2020]). And still, there are durable beliefs that some reliable patterns exist. Although 'race' is an imprecise concept, and ethnoracial categories involve important idealizations, the defenders of racial pragmatism may still think that race is precisely the type of concept that a pragmatic proxy theory would need to handle the distribution of genetic risk: not a perfect concept, but a 'useful' one. In the next section, we argue that this conclusion would be epistemologically misleading.

4.2 Ethnoracial Categories Have Limited Heuristic Power

In this section, we aim to provide two arguments against the pragmatic use of ethnoracial categories in medical genetics. We want to make the case that, beyond ontological concerns, not every stable scientific classification is equally reliable or epistemically beneficial. In particular, the classification of humans into races based on genetic data might be able to predict some regularities (ranging from regularities in skin colour, for instance, or access to healthcare systems) but be unable to generalize prediction consistently beyond such factors. In contrast to Spencer and the advocates of the proxy theory of race, this makes ethnoracial categories unreliable and thus substantially reduces their practical utility, especially if they are construed via the analysis of genetic diversity. Moreover, we argue that contemporary biology provides strong reasons to believe that, when it comes to complex diseases, the variability patterns captured by ethnoracial categories at the phenotypic level (if any) do not have a linear correspondence with variability patterns at the genotype level. This makes the use of such categories in genetics and genomics very problematic even as a heuristic concept, including their use to stratify disease risk.

The first of our arguments concerns the limited inferential power of ethnoracial categories as regards the distribution of various independent alleles (starting from knowledge about the distribution of one of them) that are supposed to be statistically associated with each other in the determination of complex diseases. To approach the problem, let us outline a typical inference that is being made from what we know about single-gene diseases and polygenic traits. Remarkably, at first sight, such an inference is not completely unwarranted and is rather grounded in classical models in quantitative genetics.

Rare Mendelian diseases—which are related to variation in a single or a few genes—appear to be unequally distributed across human populations (for example, Tay-Sachs disease is prevalent in some Jewish lineages, sickle-cell anaemia in African ‘Blacks’; for a ‘good example’ concerning differences in frequencies of lactase persistence alleles, see Spencer [2018a], p. 1028). This testifies that some medically relevant genetic characteristics are more likely to be found in some ethnoracial groups for both contingent historical reasons (for example, geographical isolation) and evolutionary reasons (some diseases increase fitness in certain environmental contexts, for example, sickle-cell anaemia protects against malaria).

When it comes to complex traits (for example, cancer, diabetes, hypertension, asthma, obesity, major depression, intelligence, and personality disorders), quantitative genetics

models predict that the same as Mendelian diseases might be true for multifactorial and polygenic diseases because variation in genotypes is distributed differently across populations for the very same reasons that single genes are distributed that way—that is, geographical and evolutionary reasons. Notably, this should not be taken as something that is established on 'empirical grounds'—in fact, for most traits, we do not know yet what the polymorphisms are that might constitute the polygenic basis of complex traits (this is a question that genetic studies, particularly GWAS, try to address). Rather, this is a general principle that would follow from such models (for example, Fisher [1918]; Mather [1941], [1943]), according to which, variation in complex traits is due to many (hundreds or thousands) alleles that are normally distributed in any given population.

The idea, in short, is that some populations or groups (and thus races) may just happen to have more of those polymorphisms that constitute the polygenic basis of a complex disease so that an individual from such groups would carry more of such polymorphisms. So, based on classical models, there are apparently good reasons to think that some correlational structure would follow a geographical distribution of genetic variability, and such a correlational structure may match ethnoracial divisions.

This is another crucial piece in understanding the current use of races in medical genetics: although one may agree that there is no clear-cut distinction among races (see Section 4.1), the data testify that some alleles are distributed differently across groups; nothing, then, prevents also thinking that 'sets' of alleles, relating to polygenic diseases, are distributed in the very same way. The inferential potential of ethnoracial categories based on the distribution of single genetic variants is, for instance, the main idea put forward by Burchard *et al*. ([2003]; see Spencer [2018a], p. 1028) to defend the notion of race as a useful one in clinical practice (for example, diagnosis and treatment) as well as in research (for example, GWAS and studies on drug efficacy). As explained in Section 2, this is essentially the main tenet of the proxy theory of race in the context of contemporary genetics and genomics.[7]

[7] A similar position is discussed by Spencer ([2018a]), who ends up defending a stronger position than Burchard *et al*. ([2003]) according to which the OMB categorisation correspond to real continental populations. If by 'real' Spencer means that races exist because they are epistemically useful, he is clearly wrong: as the natural-kind debate demonstrated extensively, scientific categories can be useful for epistemic practices such as induction, generalisation, and prediction regardless of their 'reality'. For instance, the categorisation of fruits based on their colour (examined by Hochman ([2021b]), see Section 4.1) might have some inferential power or allow for some generalisations, though their utility would depend on our practical aims. We may instead read Spencer's position more charitably, in the sense that 'reality' means 'natural kindness', that is, the property of a scientific category

Years ago, Graves ([2002]) provided the basic structure of a counterargument to this idea, though his discussion regards independent alleles related to single-gene diseases. He argued against the view that the distribution of single alleles should be consistent with the distribution of *n* other alleles. According to Graves:

> the frequency of disease-associated alleles could be claimed to be 'racially' differentiated at a particular locus, but certainly not at all such loci. We can examine this problem using a binomial equation: if we allow the rates to be higher or lower in 'blacks' versus 'whites' at some particular frequencies p and q, then the probability that all such independent loci would be at higher frequency in 'blacks' would be p^n, where n is the number of loci in question. It is easy to show that the probability that all such loci (and their phenotypes) are always higher in 'blacks' rapidly approaches zero as the number of loci increases. (Graves [2002], p. 154)

In contemporary genomics, this type of generalization issue is even more crucial than it was in the early 2000s: indeed, recent genetic methods such as GWAS and PRSs calculation focus on complex traits relating to hundreds or thousands of alleles, rather than so-called 'simple ones' relating to a single or a few genes (for a discussion of these two kinds of traits, see Serpico 2020). It is in this type of study that, nowadays, ethnoracial categories would play a role in accounting for between-individual differences at the genetic and phenotypic levels.

If we translate Graves' argument in the context of polygenic traits, we can easily see that ethnoracial categories may have limited inferential power and little epistemic utility: even if some alleles are more likely to be found in some ethnoracial groups, knowledge about the distribution of one allele does not allow the making of any reliable inference regarding the distribution of other genetic factors. So, it might be true that ethnoracial variation is statically associated with genetic variation for single-gene conditions, but is probably false in the case of 'sets of many genes' (and the complex diseases that would be associated with them).

In other words, although some geographical patterns can be observed in the distribution of single genetic variants (and thus in single-gene conditions), we cannot expect that the distribution of an allele associated with a given polygenic disease will coincide with the

to enhance epistemic practices such as induction, generalisation, and prediction (this 'epistemic interpretation' is more evident in Spencer ([2012]), where he uses the notion of 'genuine kinds'). Of course, epistemic properties alone do not make a category 'ontologically real' but they can, in principle, make it scientifically useful. Yet, we agree with the opinion that the question of the reality of racial categories should not be reduced to a question of their utility (for example, Winsberg [2022], pp. 20–22).

distribution of another (independent) allele that relates to the same disease. For instance, we cannot expect that the allele *xyz*, more likely to be found, say, in South Asian populations, will be stably associated with the allele *pqr* and thus more likely to be found in South Asian populations—unless the two alleles are physically located on the same chromosome and are thus in linkage disequilibrium (LD). The reason is that the constant flow and inter-reproduction of *Hominidae* over the globe, over the last many thousands of years, generated a distribution of human genetic variants that is chaotic, to say the least. The unpredictability skyrockets if we consider more than just two of the many alleles that are predicted to be related to variation in a complex disease, none of which is individually necessary for developing it.

When it comes to the use of a few ethnoracial categories as reference classes, what sort of practical utility could these categories have if even this sort of basic inference (from the distribution of one allele to the distribution of other two, three, four…) is problematic? If in real individuals there is no consistency in how genes tend to be associated together, how can we reliably infer the genetic characteristics of an individual based on their (supposed) ethnoracial identity? If we are right, we cannot expect that many such alleles will be stably associated with each other in such a way as to form 'clusters of variation' that correspond to a few conventional races. At most, as has been mentioned above, certain alleles can be more frequent in people coming from similar geographical areas, but there is no evidence that any given allele is stably associated with another relevant allele (or many others) in such a way that we can make a reliable inference from knowledge about one genetic marker to others.

Some critics may be still convinced that, for genealogical or geographical reasons, it could be expected that some correlations do exist between variations in sets of alleles and ethnoracial categories, and this should be enough to make a medical inference based on such correlations. If so, genetic data on ethnoracial categories would enable us to make inferences from genetic to phenotypic variation, helping us predict the properties of a given individual based on the average properties of the group they belongs to and thus enhance prevention and diagnosis—which, within the picture delineated by P-medicine, represents a key epistemological goal.

This leads us to our second criticism of the idea that race can serve as a proxy for genetic and phenotypic variation. We aim to argue that the very idea that ethnoracial groups (as usually construed) consistently relate to disease-associated genetic variants requires us to assume some sort of genotype-phenotype (G-P) linearity, which is an untenable assumption. The key idea is

that there is no typical 'racial biology' (African, Asian, American, European, etc.) that, starting from some genetic pattern, is conserved 'all the way up to the phenotype'. At the same time, the inferential direction does not hold the other way around, either: there is no reliable inference that, starting from the phenotype at any level, allows us to capture genetic patterns reliably. In other words, there is no linear correspondence between genetic factors, medically relevant features, and race-typical (superficial) traits that would allow reliably clustering humans into a small number of races.

Biologically speaking, the only way to accept a co-occurrence between genetic factors and higher-level traits (physical, physiological, etc.) is to accept some form of G-P linearity. In this view, the effect of a single or many genes would be observable at various phenotypic levels, all the way up to clinical symptoms, behaviour, and physical characteristics. The same, of course, would be true in the other direction: phenotypes analysed at any level (from superficial traits like skin pigmentation to medically relevant traits at the immunological and endocrine levels, for instance) would be consistently associated with genetic characteristics. However, it is well established that there is no such G-P relationship in complex systems like those of humans. Let us clarify why.

Ever since the late nineteenth century, geneticists have known that linearity can be observed in Mendelian experiments or in very rare conditions where variation in single genes is associated with variation in single traits through major biochemical pathways (Mendel [1866]; Morgan *et al*. [1915]). As regards complex traits, early models in quantitative genetics (Falconer [1965]; Fisher [1918]) were designed in such a way that a linear G-P map was assumed by default (indeed, such models tried to make sense of the continuous distribution of complex traits in terms of simple Mendelian inheritance). These are, however, very idealized contexts and models, respectively, that do not apply to the majority of traits, where the interaction between lower and higher levels of organization involves more complex relationships.[8]

It is, in fact, possible that a genetic factor has little or no observable phenotypic effects in some individuals due to some protective factors at higher levels (neuroendocrine, immunological, and molecular features that can mediate genetic expression), like in the case of imperfect penetrance of Mendelian genes (Chen *et al*. [2016]; Cooper *et al*. [2013]; Katsanis [2016];

[8] On linearity in experimental and Mendelian contexts, see DiFrisco and Jaeger [2019]; Kendler [2006]; Griffiths and Stotz [2013]; Ratner [2004]; Rheinberger *et al*. [2015]; Serpico [2020]. On linearity in quantitative genetics models, see Koi [2021]; Huang and MacKay [2016)]; Koi [2021]; Nelson *et al*. [2013]; Serpico & Petrolini [2023].

Lynch [2021]), or due to gene-environment interactions that generate heterogeneity in phenotypic outcomes (treatment response, disease manifestation). In the case of viral infections, for instance, the level of symptoms may be of little relevance for predicting variation at the genetic level whereby there are protective factors at the immunological level that prevent the manifestation of symptoms or the development of a disease.

Why is this relevant to our problem? The non-linearity of biological systems suggests that humans can differ from each other at many levels of organization, none of which is straightforwardly related to the others. This makes it very difficult to identify the proper level according to which we should categorize humans in biomedical research or assign individuals to a reference class: in the analysis of a certain disease, for instance, the relevant level to stratify risk could be the immunological one, but for another disease, it could be the endocrine level. The problem is all the more crucial when we categorize individuals through superficial, coarse-grained parameters such as age, gender, IQ, and socioeconomic status. Ethnoracial categories are a paradigmatic case since they cluster individuals into groups based on self-identified ethnoracial identity and on traits that are probably irrelevant to the aetiology of (most) diseases.

The take-home message is that contemporary biology provides sufficient reasons to think that what we observe at the genetic level may not correspond to what we observe at higher levels of organization. On the contrary, it probably does not. So, if a given population of individuals is divided into subgroups based on their skin pigmentation or self-identified ethnoracial identity, it is very unlikely that this classification will match the stratification of the population in terms of relevant variables, namely, the biological and genetic variables that increase the liability to a given disease. Here, we concur with the view of Mills ([1998]) and Winsberg ([2022]) according to which the folk ontology of race is murky, and, in this sense, folk races do not supervene on any clear set of biological or psychological properties or ‘essence’. But we want to emphasize a more specific, often neglected point. Race-typical features picked up by the OMB classification involve a complex mix of physical and social characteristics. While it is possible that humans can be divided into subgroups based on one (or the combination of a few) of such features, speaking biologically (in terms of the multi-level complexity of human organisms), there is little chance that such features are consistently associated in such a way that variation at the genetic level matches variation at higher levels (physical, immunological, etc.).

This connects to the use of ethnoracial categories to control for stratification issues. As we noted in Section 3.1, many scholars call for a higher inclusion of ethnoracial minorities in genetics research in order to increase the generalizability of genetic findings across different populations, including group-specific PRSs. However, developing PRSs for specific ethnoracial groups will not necessarily solve the portability problem: as Kaplan and Fuellerton ([2022]) argued, the analysis of non-European ancestry populations may be necessary, but not sufficient, for achieving more generalizable polygenic prediction (note, for instance, that the PRSs predictive accuracy can differ substantially also across groups of similar ancestry, see Mostafavi *et al.* [2020]). The portability problem does in fact depend on a number of issues, among them the fact that the effects of alleles on the phenotype are mediated by context-relative interactions between genetic, epigenetic, and environmental variation (Burt [2022]; Janssens [2019]; Wang *et al.* [2022]), which—crucially—can be distributed differently among individuals that are assigned to the same ethnoracial categories through classical methods (e.g., self-identification). If our argument holds, ethnoracial groups cannot be expected to be internally homogeneous for every biological and environmental factor that is relevant to a given pathological phenotype, which suggests that ethnoracial categories cannot account for the clustering of human variability at the relevant medical levels.

5. Conclusions

The problem of categorizing individuals into groups that are biologically meaningful and clinically relevant is both an epistemic aim of contemporary P-medicine and a precondition for its success. In this article, we have discussed the major conceptual and epistemological limitations in the use of ethnoracial categories as reference classes in biomedical research and practice, particularly in medical genetics.

In Section 2, we outlined the role of ethnoracial categories as a proxy for medically relevant differences among individuals and introduced the potential risks of construing misleading reference classes.

In Section 3, we explained that socially defined groups like ethnoracial categories are often conceptualized according to institutional guidelines (for example, the FDA and NIH ones endorsing the OMB categorization) and folk beliefs. How people are categorized directly influences the results obtained throughout the process of designing a study, analysing its results, formulating conclusions, and providing applications in diagnostic algorithms and practical

guidelines for institutions, medical practitioners, and the patients themselves. Thus, rampant ambiguities in the categories of race and ethnicity bring about substantial epistemic and ethical uncertainty in biomedicine. For instance, they leave the door open to the misleading impression that some functional rationale stands behind the categorization of individuals into ethnoracial groups or, in other words, that these categories have some biological or medical significance. This is a very simple illustration of how the top-down requirement to use ethnoracial classifications in science can potentially distort its results.

In Section 4, we thus argued that the use of ethnoracial categories is especially problematic in medical genetics. As has been explained, previous criticisms of the genetic interpretation of race and ethnicity did not disincentivize the use of ethnoracial categories in contemporary genetics. We have suggested that this depended, at least in part, on the fact that such criticisms target the ontological reality of races but leave the door open to their use as a heuristic, pragmatic tool to handle human variability at the genetic level. Indeed, the 'imprecision' of ethnoracial categories is not enough to disregard their many uses. More generally, scholars agree that some concepts and categories may not be descriptively accurate or 'real' but could still be part of scientific practice if sufficiently stable to allow for useful generalizations. This is, of course, unless their utility is argued to be compromised.

So, we have provided two further arguments to show that ethnoracial categories provide little epistemic benefit in medical genetics and cannot serve as valuable heuristic tools to account for medically relevant phenotypic differences among individuals. Our analysis of the use of ethnoracial classifications in genetic studies reveals that scarce attention has been given to the multi-level biological complexity of individual differences and their distribution worldwide: if human diversity is to be taken into account seriously, it is crucial to construe reference classes that are both biologically (not just genetically) and socially sensitive.

Notably, our reasoning does not discount the fact that disease incidence can be linked to certain polymorphisms that are distributed differently in different geographical areas or subgroups: what they do indicate is that conventional ethnoracial categorizations cannot unequivocally correlate 'races' and 'ethnicities' with these polymorphisms, especially when it comes to complex phenotypes related to many alleles. In this regard, we agree with the voices that, in genetics and genomics, using ethnoracial classifications as proxies for genetic differences between certain populations (if necessary, that is, if the research methodology requires using such proxies) should be replaced, for instance, by application of more precise demo-geographic

classifications (as, by now, though it remains only a rough generalization, geography seems to be the strongest indicator of genetic variation among human populations; see Huddart *et al.* [2019], p. 1258). Moreover, our arguments support the view that human variation at the genetic level should not be confused with biological, socially mediated variability (Duello *et al.* [2021]; Gravlee [2009]; Kaplan [2010]; Valles [2016]). Thus, it is about time for institutions such as the FDA and the NIH to rethink the recommendation/requirement to collect and use ethnoracial data and classifications in genetics and genomics research.

Acknowledgements

Joanna Karolina Malinowska was supported by a grant from the National Science Centre, Poland, no. UMO-2020/39/D/HS1/00636. Davide Serpico's research has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No 805498).

Joanna Karolina Malinowska (corresponding author)
Adam Mickiewicz University, Faculty of Philosophy, ul. Szamarzewskiego 89C, 60-568 Poznań, Poland
Email: malinowska@amu.edu.pl

Davide Serpico
Department of Economics and Management, University of Trento, Via Vigilio Inama 5, 38122, Trento, Italy
Interdisciplinary Centre for Ethics & Institute of Philosophy, Jagiellonian University, Grodzka 52, 31-044 Kraków, Poland

References

Abubakar, I., Gram, L., Lasoye, S., Achiume, E. T., Becares, L., Bola, G. K., ... & Devakumar, D. [2022]: 'Confronting the Consequences of Racism, Xenophobia, and Discrimination on Health and Health-Care Systems', *The Lancet*, **400**(10368), pp. 2137-2146.

American Society of Human Genetics (ASHG) [2018]: 'ASHG Denounces Attempts to Link Genetics and Racial Supremacy', *The American Journal of Human Genetics,* **103***,* p. 636.

Barbujani, G., Magagni, A., Minch, E. and Cavalli-Sforza, L. L. [1997]: 'An Apportionment of Human DNA Diversity', *Proceedings of the National Academy of Sciences,* **94**(9), pp. 4516–19.

Bliss, C. [2020]: 'Conceptualizing Race in the Genomic Age', *Hastings Center Report,* **50**, pp. S15–S22.

Boem, F. [2022]: 'Racism After the End of the Race: A Brief Epistemological Viewpoint on Genomic Studies and Racism', *Africa e Mediterraneo*, **96**, pp. 1-22.

Boorse, C. [1977]: 'Health as a Theoretical Concept', *Philosophy of Science*, **44**(4), pp. 542–73.

Borrell, L. N., Elhawary, J. R., Fuentes-Afflick, E., et al. [2021]: 'Race and Genetic Ancestry in Medicine—A Time for Reckoning with Racism', *New England Journal of Medicine*, **384**(5), pp. 474–80.

Brattain, M. [2007]: 'Race, Racism, and Antiracism: UNESCO and the politics of presenting Science to the postwar Public', *The American Historical Review,* **112**(5), pp. 1386–1413.

Burchard, E., Ziv, E., Coyle, N., Gomez, S., L., Tang, H., Karter, A., J., Mountain, J., L., Pérez-Stable, E., J., Sheppard, D., Rish, N. [2003]: 'The Importance of Race and Ethnic Background in Biomedical Research and Clinical Practice', *The New England Journal of Medicine,* **348**(12), pp. 1170–75.

Burt, C. H. [2022]: 'Challenging the Utility of Polygenic Scores for Social Science: Environmental Confounding, Downward Causation, and Unknown Biology', *Behavioral and Brain Sciences,* pp. 1–36.

Byeon, Y. J. J., Islamaj, R., Yeganova, L., Wilbur, W. J., Lu, Z., Brody, L. C. and Bonham, V. L. [2021]: 'Evolving Use of Ancestry, Ethnicity, and Race in Genetics Research—A Survey Spanning Seven Decades' *The American Journal of Human Genetics*, **108**(12), 2215–23.

Centers for Disease Control and Prevention. [2021]: 'COVID-19 Hospitalization and Death by Race/Ethnicity', available at https://www.cdc.gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-race-ethnicity.html

Chellappoo, A. & Baedke, J. [2023]: 'Where the Social Meets the Biological: New Ontologies of Biosocial Race', *Synthese*, **201**(1), pp. 14.

Chen, R., Shi, L., Hakenberg, J., Naughton, B., et al. [2016]: 'Analysis of 589,306 Genomes Identifies Individuals Resilient to Severe Mendelian childhood Diseases', *Nature Biotechnology,* **34**(5), pp. 531–38.

Cooper, D. N., Krawczak, M., Polychronakos, C., Tyler-Smith, C. and Kehrer-Sawatzki, H. [2013]: 'Where genotype is Not Predictive of Phenotype: Towards an Understanding of the Molecular Basis of Reduced Penetrance in Human Inherited Disease', *Human Genetics,* **132**(10), pp. 1077–1130.

Crockett, D. [2008]: 'Marketing Blackness: How Advertisers Use Race to Sell Products', *Journal of Consumer Culture,* **8**(2), pp. 245–68.

Cui, L., Xu, T. and Zhang, L. [2020]: 'Issues Related to Subgroup Analyses and Use of Intensive Stratification', in N., Ting, J. C. Cappelleri, & S. Ho (*eds*.), *Design and Analysis of Subgroups with Biopharmaceutical Applications*, Cham, Switzerland: Springer, pp. 3–17.

Cunningham, P. W. [2018]: 'The Health 202: NIH wants 1 Million Americans to Contribute to New Pool of Gene Data', *The Washington Post*, available at <https://www.washingtonpost.com/news/powerpost/paloma/the-health-202/2018/01/16/the-health-202-nih-wants-1-million-americans-to-contribute-to-new-pool-of-gene-data/5a5ba45a30fb0469e8840135/>.

Davenport, L. [2020]: 'The Fluidity of Racial Classifications', *Annual Review of Political Science*, **23**, pp. 221–40.

Devakumar, D., Selvarajah, S., Abubakar, I., Kim, S. S., McKee, M., Sabharwal, N. S., ... & Achiume, E. T. [2022]: 'Racism, Xenophobia, Discrimination, and the Determination of Health', *The Lancet*, **400**(10368), pp. 2097-2108.

DiFrisco, J. and Jaeger, J. [2019]: 'Beyond Networks: Mechanism and Process in Evo-Devo', *Biology & Philosophy,* **34**(6), p. 54.

Dobzhansky, T. [1937]: *Genetics and the Origin of Species*, New York: Columbia University Press.

Duello, T. M., Rivedal, S., Wickland, C. and Weller, A. [2021]: 'Race and Genetics Versus "Race" in Genetics: A Systematic Review of the Use of African Ancestry in Genetic Studies', *Evolution, Medicine, and Public Health*, **9**(1), pp. 232–45.

Edwards, A. W. [2003]: 'Human Genetic Diversity: Lewontin's Fallacy', *BioEssays,* **25**(8), pp. 798–801.

Falconer, D. S. [1965]: 'The Inheritance of Liability to Certain Diseases, Estimated From the Incidence Among Relatives', *Annals of Human Genetics,* **29**(1), pp. 51–76.

Fatumo, S., Chikowore, T., Choudhury, A., Ayub, M., Martin, A.R., Kuchenbaecker, K. [2022]: 'A Roadmap to Increase Diversity in Genomic Studies', *Nature Medicine,* **28**(2), pp. 243–50.

Fisher, R. A. [1918]: 'The Correlation Between Relatives on the Supposition of Mendelian Inheritance', *Transactions of the Royal Society of Edinburgh,* **52**(2), pp. 399–433.

Flanagin, A., Frey, T., Christiansen, S. L. and AMA Manual of Style Committee. [2021]: 'Updated Guidance on the Reporting of Race and Ethnicity in Medical and Science Journals', *JAMA*, **326**(7), pp. 621–27.

Food and Drug Administration (FDA) [2016]: *Collection of Race Ethnicity Data in Clinical Trials. Guidance for Industry and Food and Drug Administration Staff*, available at <https://www.fda.gov/media/75453/download>.

Gravlee, C. C. [2009]: 'How Race Becomes Biology: Embodiment of Social Inequality', *American Journal of Physical Anthropology*, **139**(1), pp. 47-57.

Graves, J. L. [2002]: 'Scylla and Charybdis: Adaptationism, Reductionism, and the Fallacy of Equating Race with Disease', in L.S. Parker and R. A. Ankeny (*eds*.), *Mutating Concepts, Evolving Disciplines: Genetics, Medicine, and Society*, Dordrecht: Springer, pp. 143–64.

Griffiths, P. and Stotz, K. [2013]: *Genetics and Philosophy: An Introduction*. Cambridge, UK: Cambridge University Press.

Grill, J. [2018]: '"In England, they don't call you black!" Migrating Racialisations and the Production of Roma Difference across Europe', *Journal of Ethnic and Migration Studies*, **44**(7), pp. 1136–55.

Haslanger, S. [2019]: 'Tracing the Sociopolitical Reality of Race', in J. Glasgow, S. Haslanger, C. Jeffers, Q. Spencer, *What Is Race? Four philosophical views*, Oxford: Oxford University Press, pp. 4–37.

Hellwege, J. N., Keaton, J. M., Giri, A., Gao, X., Velez Edwards, D. R. and Edwards, T. L. [2018]: 'Population Stratification in Genetic Association Studies', *Current Protocols in Human Genetics,* **95**(1), pp. 1–22.

Hochman, A. [2013]: 'Racial discrimination: How not to do it', *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, **44**(3), pp. 278-286.

Hochman, A. [2017]: 'Replacing Race: Interactive Constructionism About Racialized Groups', *Ergo: An Open Access Journal of Philosophy*, **4.**

Hochman, A. [2019]: 'Racialization: A Defense of the Concept', *Ethnic and Racial Studies*, **42**(8), pp. 1245–62.

Hochman, A. [2021a]: 'Against the Reification of Race in Bioethics: Anti-Racism Without Racial Realism', *The American Journal of Bioethics*, **21**(2), pp. 88–90.

Hochman, A. [2021b]: 'Racial Classification Without Race: Edwards' Fallacy', in L. Lorusso and R.G. Winther (*eds*.), *Remapping Race in a Global Context*, New York: Routledge, pp. 74–91.

Hochman, A. [2022]: 'Has Social Constructionism About Race Outlived its Usefulness? Perspectives From a Race Skeptic', *Biology & Philosophy*, **37**(6), p. 48.

Huang, W., & Mackay, T. F. [2016]: 'The genetic architecture of quantitative traits cannot be inferred from variance component analysis', *PLoS Genetics*, **12**(11), e1006421.

Huddart, R., Fohner, A.E., Whirl-Carrillo, M., Wojcik, G.L., Gignoux, C. R., Popejoy, A.B. and Klein, T. E. [2019]: 'Standardized Biogeographic Grouping System for Annotating Populations in Pharmacogenetic Research', *Clinical Pharmacology & Therapeutics,* **105**(5), pp. 1256–62.

Jackson Jr, J. P. [2022]: 'The Perils of Polysemy: Racial Realism in the Real World', *Philosophy, Theory, and Practice in Biology*, **14**.

James, M. and Burgos, A. [2022]: 'Race,' in Edward N. Zalta (*ed*.), *The Stanford Encyclopedia of Philosophy*, available at <https://plato.stanford.edu/archives/spr2022/entries/race/>.

Janssens, A. C. J. [2019]: 'Validity of Polygenic Risk Scores: Are We Measuring What We Think We Are?', *Human Molecular Genetics*, **28**(R2), pp. R143-R150.

Kahn, J. [2012]: *Race in a Bottle: The Story of BiDil and Racialized Medicine in a Post-Genomic Age*, New York: Columbia University Press.

Kalewold, K. H. [2020]: 'Race and Medicine in Light of the New Mechanistic Philosophy of Science', *Biology & Philosophy*, **35**(4), pp. 1–22.

Kalmar, I. [2022]: *White but Not Quite: Central Europe's Illiberal Revolt*, Bristol: Bristol University Press.

Kalmar, I. [2023]: 'Race, Racialisation, and the East of the European Union: An Introduction', *Journal of Ethnic and Migration Studies*, **49**(6), pp. 1465-1480.

Kaplan, J. M. [2010]: 'When Socially Determined Categories Make Biological Realities: Understanding Black/White Health Disparities in the US', *The Monist,* **93**(2), pp. 281–97.

Katsanis, N. [2016]: 'The Continuum of Causality in Human Genetic Disorders', *Genome Biology,* **17**(1), p. 233.

Kendler, K. S. [2006]: 'Reflections on the Relationship Between Psychiatric Genetics and Psychiatric Nosology', *American Journal of Psychiatry,* **163**(7), pp. 1138–46.

Keskinen, S. and Andreassen, R. [2017]: 'Developing Theoretical Perspectives on Racialisation and Migration', *Nordic Journal of Migration Research*, **7**(2), p. 64.

Khalidi, M. A. [2013} *Natural Categories and Human Kinds: Classification in the Natural and Social Sciences*, Cambridge: Cambridge University Press.

Koi, P. [2021]: 'Genetics on the Neurodiversity Spectrum: Genetic, Phenotypic and Endophenotypic Continua in Autism and ADHD', *Studies in History and Philosophy of Science Part A*, **89**, pp. 52–62.

Krieger, N. [2019]: 'Measures of Racism, Sexism, Heterosexism, and Gender Binarism for Health Equity Research: From Structural Injustice to Embodied Harm-An Ecosocial Analysis', *Annual Review of Public Health*, **41**, pp. 37-62.

Krieger, N. [2021]: *Ecosocial Theory, Embodied Truths, and the People's Health*, Oxford: Oxford University Press.

Krivonos, D. [2023]: 'Racial Capitalism and the Production of Difference in Helsinki and Warsaw', *Journal of Ethnic and Migration Studies*, **49**(6), pp. 1500–1516.

Kuzawa, C. W. and Gravlee, C. C. [2016]: 'Beyond Genetic Race: Biocultural Insights Into the Causes of Racial Health Disparities', in M. K. Zuckerman and D. L. Martin (*eds*.), *New Directions in Biocultural Anthropology*, New Jersey: Wiley, pp. 89–105.

Kuzawa, C. W. and Sweet, E. [2009]: 'Epigenetics and the Embodiment of Race: Developmental Origins of US Racial Disparities in Cardiovascular Health', *American Journal of Human Biology: The Official Journal of the Human Biology Association*, **21**(1), pp. 2–15.

Lawson, D.J., Davies, N.M., Haworth, S., Ashraf, B., Howe, L., Crawford, A., Hemani, G., Smith, G.D., Timpson, N.J. [2020]: 'Is Population Structure in the Genetic Biobank Era Irrelevant, a Challenge, or an Opportunity?', *Human Genetics,* **139**(1), pp. 23–41.

Lewicki, A. [2023]: 'East–West Inequalities and the Ambiguous Racialisation of 'Eastern Europeans'', *Journal of Ethnic and Migration Studies*, **49**(6), pp. 1481-1499.

Lewontin, R. C. [1972]: 'The Apportionment of Human Diversity', in T. Dobzhansky, M. K. Hecht, and W. C. Steere (*eds.*), *Evolutionary Biology: Volume* 6, New York, NY: Springer US, pp. 381–98.

Lin, J., Bunn, V. and Liu, R. [2019]: 'Practical Considerations for Subgroups Quantification, Selection and Adaptive Enrichment in Confirmatory Trials', *Statistics in Biopharmaceutical Research*, **11**(4), pp. 407-418.

López, M. M., Bevans, M., Wehrlen, L., Yang, L. and Wallen, G. R. [2017]: 'Discrepancies in Race and Ethnicity Documentation: A Potential Barrier in Identifying Racial and Ethnic Disparities', *Journal of Racial and Ethnic Health Disparities,* **4**(5), pp. 812–18.

Ludwig, D. [2014]: 'Disagreement in Scientific Ontologies', *Journal for General Philosophy of Science,* **45**, pp. 119–31.

Ludwig, D. [2016]: 'Ontological Choices and the Value-Free Ideal', *Erkenntnis,* **81**, pp. 1253–72.

Ludwig, D. [2019]: 'How Race Travels: Relating Local and Global Ontologies of Race', *Philosophical Studies,* **176**(10), pp. 2729–50.

Lynch, K. E. [2021]: 'The meaning of "cause" in genetics', C*old Spring Harbor Perspectives in Medicine,* **11**(9), p. a040519.

Malinowska, J. K. [2016]: 'Cultural Neuroscience and the Category of Race: The Case of the Other-Race Effect', *Synthese*, **193**, pp. 3865–87.

Malinowska, J. K. [2021]: ‘Methodological and Ethical Risks Associated with the Epistemic Unification of Tribe Members’, *The American Journal of Bioethics*, **21**(10), pp. 32–34.

Malinowska, J. K. and Żuradzki, T. [2023a]: ‘Reductionist Methodology and the Ambiguity of the Categories of Race and Ethnicity in Biomedical Research: An Exploratory Study of Recent Evidence’, *Medicine, Health Care and Philosophy*, **26**(1), pp. 55–68.

Malinowska, J. K. and Żuradzki, T. [2023b]: ‘Towards the Multileveled and Processual Conceptualisation of Racialised Individuals in Biomedical Research’, *Synthese*, **201**(1), p. 11.

Mallon, R. [2006]: ‘“Race”: Normative, Not Metaphysical or Semantic’, *Ethics*, **116**(3), pp. 525-51.

Martin, A. R., Kanai, M., Kamatani, Y., Okada, Y., Neale, B. M., & Daly, M. J. [2019]: 'Clinical Use of Current Polygenic Risk Scores May Exacerbate Health Disparities', *Nature Genetics*, **51**(4), pp. 584-591.

Mather, K. [1941]: ‘Variation and Selection of Polygenic Characters’, *Journal of Genetics,* **41**(1), pp. 159–93.

Mather, K. [1943]: ‘Polygenic Inheritance and Natural Selection’, *Biological Reviews,* **18**(1), pp. 32–64.

Meissner, S. N. [2021]: ‘How Does the Consideration of Indigenous Identities in the US Complicate Conversations About Tracking Folk Racial Categories in Epidemiologic Research?’, *Synthese*, **198(**10), pp. 2439–62.

Mendel, G. [1866]: Versuche über Pflanzen-Hybriden. Verhandlungen des Naturforschenden Vereines in Brünn, IV. [Experiments on Plant Hybrids. New Translation with Commentary. Mueller-Wille, S., Hall, K. Dostal, O., Radick, G., and Martin, R., (2020).]

Mills, C. W. [1998]: *Blackness Visible: Essays on Philosophy and Race*, Ithaca and London: Cornell University Press.

Morgan, T. H., Sturtevant, A. H., Muller, H. J. and Bridges, C. B. [1915]: *The Mechanism of Mendelian Heredity*, New York: Henry Holt and Company.

Mostafavi, H., Harpak, A., Agarwal, I., et al. [2020]: 'Variable Prediction Accuracy of Polygenic Scores Within an Ancestry Group', *Elife*, **9**, e48376.

Mulinari, S., Vilhelmsson, A., Ozieranski, P. and Bredström, A. [2021]: ‘Is There Evidence for the Racialization of Pharmaceutical Regulation? Systematic Comparison of New Drugs Approved Over Five Years in the USA and the EU’, *Social Science & Medicine*, **280**, p. 114049.

Narkowicz, K. [2023]: 'White Enough, Not White Enough: Racism and Racialisation Among Poles in the UK', *Journal of Ethnic and Migration Studies*, **49**(6), pp. 1534–1551.

National Institutes of Health (NIH) [2001]: *NIH Policy and Guidelines on The Inclusion of Women and Minorities as Subjects in Clinical Research (October 2001),* available at <https://grants.nih.gov/grants/guide/notice-files/NOT-OD-02-001.html>.

Nelson, R. M., Pettersson, M. E. and Carlborg, O. [2013]: ‘A Century After Fisher: Time for a New Paradigm in Quantitative Genetics’, *Trends in Genetics,* **29**(12), pp. 669–76.

OMB. [1997]: The Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity (OMB Statistical Policy Directive No. 15), https://orwh.od.nih.gov/toolkit/other-relevant-federal-policies/OMB-standards.

Ondra, T., Dmitrienko, A., Friede, T., Graf, A., Miller, F., Stallard, N. and Posch, M. [2016]: 'Methods for Identification and Confirmation of Targeted Subgroups in Clinical Trials: A Systematic Review', *Journal of Biopharmaceutical Statistics*, **26**(1), pp. 99–119.

Palk, A. C., Dalvie, S., De Vries, J., et al. [2019]: 'Potential Use of Clinical Polygenic Risk Scores in Psychiatry–Ethical Implications and Communicating High Polygenic Risk', *Philosophy, Ethics, and Humanities in Medicine*, **14**(1), pp. 1-12.

Peto, R. [2011]: 'Current Misconception 3: That Subgroup-Specific Trial Mortality Results Often Provide a Good Basis for Individualising Patient Care', *British Journal of Cancer*, **104**(7), pp. 1057–58.

Pocock, S. J., Assmann, S. E., Enos, L. E. and Kasten, L. E. [2002]: 'Subgroup Analysis, Covariate Adjustment and Baseline Comparisons in Clinical Trial Reporting: Current Practice and Problems', *Statistics in Medicine*, **21**(19), pp. 2917–30.

Pollock, A. [2012]: *Medicating Race: Heart Disease and Durable Preoccupations with Difference*, Durham: Duke University Press.

Popejoy, A. B., Crooks, K. R., Fullerton, S. M., Hindorff, L. A., Hooker, G. W., Koenig, B. A., … and Diversity Working Group. [2020]: 'Clinical Genetics Lacks Standard Definitions and Protocols for the Collection and Use of Diversity Measures', *The American Journal of Human Genetics,* **107**(1), pp. 72–82.

Ratner, C. [2004]: 'Genes and Psychology in the News', *New Ideas in Psychology,* **22**(1), pp. 29–47.

Reydon, T. A. and Ereshefsky, M. [2022]: 'How to Incorporate Non-Epistemic Values into a Theory of Classification', *European Journal for Philosophy of Science*, **12**(1), pp. 1–28.

Rheinberger, H. J., Müller-Wille, S. and Meunier, R. [2015]: 'Gene', in E. N. Zalta (*ed.*), *The Stanford Encyclopedia of Philosophy*, available at <https://plato.stanford.edu/archives/spr2015/entries/gene>.

Rosenberg, N. A., Pritchard, J. K., Weber, J. L., Cann, H. M., Kidd, K. K., Zhivotovsky, L. A. and Feldman, M. W. [2002]: 'Genetic Structure of Human Populations', *Science,* **298**(5602), pp. 2381–85.

Rzepiński, T. [2016]: 'Problemy Epistemologiczne Analizy Podgrup Randomizowanych Badań Klinicznych' [Epistemological Problems of Subgroup Analyses of Randomised Controlled Trials], *Filozofia Nauki*, **24**(4), pp. 73–94.

Rzepnikowska, A. [2023]: 'Racialisation of Polish Migrants in the UK and in Spain (Catalonia)', *Journal of Ethnic and Migration Studies,* **49**(6), pp. 1517-1533.

Saha, A. [2015]: 'The Marketing of Race in Cultural Production', in K. Oakley and J. O'Connor (*eds.*), *The Routledge Companion to the Cultural Industries*, New York: Routledge, pp. 528–37.

Sallaz, J. J. [2010]: 'Talking Race, Marketing Culture: The Racial Habitus in and out of Apartheid', *Social Problems*, **57**(2), pp. 294–314.

Sankar, P. and Kahn, J. [2005]: 'BiDil: Race Medicine or Race Marketing? Using Race to Gain a Commercial Advantage Does Not Advance the Goal of Eliminating Racial/Ethnic Disparities in Health Care', *Health Affairs,* **24**(Suppl1), pp. W5–455.

Selvarajah, S., Maioli, S. C., Abi Deivanayagam, T., de Morais Sato, P., Devakumar, D., Kim, S. S., ... & Paradies, Y. [2022]: 'Racism, Xenophobia, and Discrimination: Mapping Pathways to Health Outcomes', *The Lancet,* **400**(10368), pp. 2109-2124.

Serpico, D. [2020]: ‘Beyond Quantitative and Qualitative Traits: Three Telling Cases in the Life Sciences’, *Biology & Philosophy,* **35***(*3), pp. 1–26.

Serpico, D. & Petrolini, V. [2023]: 'Crossing the Threshold: An Epigenetic Alternative to Dimensional Accounts of Mental Disorders', *The British Journal for the Philosophy of Science*.

Shannon, G., Morgan, R., Zeinali, Z., Brady, L., Couto, M. T., Devakumar, D., ... & Muraya, [2022]: 'Intersectional Insights into Racism and Health: Not Just a Question of Identity', *The Lancet*, **400(**10368), pp. 2125-2136.

Sleight, P. [2000]: ‘Debate: Subgroup Analyses in Clinical Trials: Fun to Look At – But Don’t Believe Them!’, *Trials*, **1**(1), pp. 1–3.

Spencer, Q. [2018a]: ‘A Racial Classification for Medical Genetics’, *Philosophical Studies,* **175**(5), pp. 1013–37.

Spencer, Q. [2018b]: ‘Are Folk Races Like Dingoes, Dimes, or Dodos?’, in G. Rosen, A. Byrne, J. Cohen, E. Harman, & S. Shiffrin (*eds*.), *The Norton Introduction to Philosophy*, New York: W.W. Norton & Company Inc, pp. 571–81.

Spencer, Q. [2019]: ‘How to be a Biological Racial Realist’, in J. Glasgow, S. Haslanger, C. Jeffers, Q. Spencer, *What Is Race? Four philosophical views*, Oxford: Oxford University Press, pp. 73–110.

Sullivan, S. [2013]: 'Inheriting Racist Disparities in Health: Epigenetics and the Transgenerational Effects of White Racism', *Critical Philosophy of Race*, **1**(2), pp. 190-218.

Sun, X., Briel, M., Busse, J. W., You, J. J., et al. [2012]: ‘Credibility of Claims of Subgroup Effects in Randomised Controlled Trials: Systematic Review’, *BMJ*, **344**.

Tang, H., Quertermous, T., Rodriguez, B., Kardia, S.L.R., Zhu, X., Brown, A., Pankow J.S., Province, M.A., Hunt, S.C., Boerwinkle, E., Schork, N.J., Rish, N.J. [2005]: ‘Genetic Structure, Self-Identified Race/Ethnicity, and Confounding in Case-Control Association Studies’, *The American Journal of Human Genetics,* **76**(2), pp. 268–275.

Tereshchenko, A., Bradbury, A., & Archer, [2019]: 'Eastern European Migrants’ Experiences of Racism in English Schools: Positions of Marginal Whiteness and Linguistic Otherness', *Whiteness and Education*, **4**(1), pp. 53-71.

Valles, S. A. [2012]: ‘Heterogeneity of Risk Within Racial Groups, a Challenge for Public Health Programs’, *Preventive Medicine*, **55**(5), pp. 405–8.

Valles, S. A. [2016]: ‘Race in Medicine’, in M. Solomon, J.R. Simon and H. Kincaid (*eds*.), *The Routledge Companion to Philosophy of Medicine*, Yew York: Routledge, pp. 433–45.

Valles, S. A. [2021a]: 'Why Race and Ethnicity Are Not Like Other Risk Factors: Applying Structural Competency and Epistemic Humility in the Covid-19 Pandemic', *Philosophy of Medicine*, **2**(1), pp. 1-8.

Valles, S. A. [2021b]: 'A Pluralistic and Socially Responsible Philosophy of Epidemiology Field Should Actively Engage with Social Determinants of Health and Health Disparities', *Synthese*, **198(**Suppl 10), pp. 2589-2611.

Veit, W. [2021]: ‘Biological Normativity: A New Hope for Naturalism?’, *Medicine, Health Care and Philosophy*, **24**(2), pp. 291–301.

Venn, J. [2019]: *The Logic of Chance*. Norderstedt: BoD–Books on Demand.

Wallmann, C. and Williamson, J. [2017]: 'Four Approaches to the Reference Class Problem', in G. Hofer-Szabó and L. Wroński (*eds*.), *Making it Formally Explicit*, Cham: Springer, pp. 61–81.

Wang, Y., Tsuo, K., Kanai, M., et al. [2022]: 'Challenges and Opportunities for Developing More Generalizable Polygenic Risk Scores', *Annual Review of Biomedical Data Science*, **5**.

Williams, D. R., Priest, N., & Anderson, N. B. [2016]: 'Understanding Associations Among Race, Socioeconomic Status, and Health: Patterns and Prospects', *Health Psychology*, **35**(4), pp. 407.

Winsberg, E. [2022]: 'Putting Races on the Ontological Map: A Close Look at Spencer's "New Biologism" of Race', *Biology & Philosophy*, **37**(6), pp. 1–25.

Winther, R. G. and Kaplan, J. M. [2013]: 'Ontologies and Politics of Biogenomic "Race"', *Theoria*, **60**(136), pp. 54–80.

Witherspoon, D. J., Wooding, S., Rogers, A. R., Marchani, E. E., Watkins, W.S., Batzer, M.A., Jorde, L.B. [2007]: 'Genetic Similarities Within and Between Human Populations', *Genetics,* **176**(1), pp. 351–59.

Zack, N. [2016]: 'Philosophical and Social Implications of Race', in L. May (*ed*.), *Applied Ethics*, New York: Routledge, pp. 247–56.

Zhang, F. and Finkelstein, J. [2019]: 'Inconsistency in Race and Ethnic Classification in Pharmacogenetics Studies and its Potential Clinical Implications', *Pharmacogenomics and Personalized Medicine,* **12**, p. 107.

Zikmund-Fisher, B. J. [2022]: 'Toward Transparent Demographic Analyses: Statement on the Use and Reporting of Classification Variables Presented as Measuring Individual Characteristics Such as Race, Ethnicity, Indigeneity, National Origin, Gender, Sexual Orientation, or Socioeconomic Status', *Medical Decision Making*, **42**(3), pp. 277–79.