



Perspective

Population Pharmacogenomics for Health Equity

I. King Jordan ^{1,*} , Shivam Sharma ¹ and Leonardo Mariño-Ramírez ² 

¹ School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA 30332, USA; shivamsharma13@gatech.edu

² National Institute on Minority Health and Health Disparities, National Institutes of Health, Bethesda, MD 20892, USA; marino@nih.gov

* Correspondence: king.jordan@biology.gatech.edu; Tel.: +1-404-385-2224

Abstract: Health equity means the opportunity for all people and populations to attain optimal health, and it requires intentional efforts to promote fairness in patient treatments and outcomes. Pharmacogenomic variants are genetic differences associated with how patients respond to medications, and their presence can inform treatment decisions. In this perspective, we contend that the study of pharmacogenomic variation within and between human populations—population pharmacogenomics—can and should be leveraged in support of health equity. The key observation in support of this contention is that racial and ethnic groups exhibit pronounced differences in the frequencies of numerous pharmacogenomic variants, with direct implications for clinical practice. The use of race and ethnicity to stratify pharmacogenomic risk provides a means to avoid potential harm caused by biases introduced when treatment regimens do not consider genetic differences between population groups, particularly when majority group genetic profiles are assumed to hold for minority groups. We focus on the mitigation of adverse drug reactions as an area where population pharmacogenomics can have a direct and immediate impact on public health.

Keywords: pharmacogenomics; health equity; health disparities; race; ethnicity; ancestry; population; adverse drug reactions



Citation: Jordan, I.K.; Sharma, S.; Mariño-Ramírez, L. Population Pharmacogenomics for Health Equity. *Genes* **2023**, *14*, 1840. <https://doi.org/10.3390/genes14101840>

Academic Editor: Zhuo (Shawn) Shao

Received: 28 July 2023

Revised: 19 September 2023

Accepted: 20 September 2023

Published: 22 September 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Population Pharmacogenomics Can and Should Be Leveraged for Health Equity—Improved Health Outcomes for All People Everywhere

The United States (US) National Institute on Minority Health and Health Disparities (NIMHD) defines health equity as the assurance that “all individuals or populations have optimal opportunities to attain the best health possible” [1]. The converse of health equity is health disparities, which are defined by NIMHD as differences in health outcomes that affect disadvantaged populations, and the elimination of health disparities is a key component of health equity [2]. It should be stressed, however, that the health equity paradigm goes beyond just eliminating health disparities to include more aspirational goals of achieving optimal health for all populations.

NIMHD states that the application of a health equity lens to biomedical research requires intentional efforts to “promote fairness, opportunity, quality, and social justice in treatments and outcomes” [1]. This perspective article is focused on how population pharmacogenomics research can be deployed as one such intentional effort in support of health equity. We define population pharmacogenomics as the study of pharmacogenetic variation within and between populations, where pharmacogenetic variation refers to genetic differences associated with the response to medications, and populations are delineated as groups of people that share common characteristics, such as race, ethnicity, or ancestry. We explore how population pharmacogenomics research can and should be leveraged to facilitate health equity through a focus on medical treatments that consider the genetic variation of socially disadvantaged racial and ethnic groups that experience health

disparities. We highlight examples from our own research on the relationship between race, ethnicity, ancestry, and pharmacogenomic variation in cosmopolitan countries of the global north, the US and the United Kingdom (UK), and in the South American country of Colombia, to illuminate the potential of pharmacogenomics to promote health equity.

2. Race, Ethnicity, and Pharmacogenomic Variation

Socially Defined Racial and Ethnic Groups Exhibit Differences in the Frequencies of Pharmacogenomic Variants, with Direct Implications for Health Equity

The US National Academies of Science Engineering and Medicine (NASEM) recently released a comprehensive and detailed report on the use of race, ethnicity, and ancestry as population descriptors in genomics research [3]. The overarching recommendation of this report was that researchers should carefully consider and justify how and why they use race and ethnicity as population labels in their work, a recommendation with which we fully agree. The report emphasizes that race and ethnicity are social concepts, which are nevertheless often used as surrogates for describing human genetic differences. The use of race and ethnicity in this way can be problematic in the sense that it may reinforce the erroneous notion of discrete human groups with innate differences, while failing to capture complex patterns of global genetic variation.

Our use of race and ethnicity population descriptors in this perspective, and the broader relevance of these group labels to pharmacogenomics, is rooted in both empirical and pragmatic considerations. With respect to the empirical justification for the use of race and ethnicity descriptors in pharmacogenomic research, we recognize that racial and ethnic groups are socially constructed. By this, we mean that the definitions of these groups, the boundaries around the groups which in turn define their composition, are determined by human actors. In the case of the US, the UK, and Colombia, government bureaucrats decide on and enumerate group labels and their definitions [4–7]. Racial and ethnic groups differ across space and time, depending on the demographic composition of the societies in which they are used. The US uses a system that combines racial groups, defined by shared ancestral origins, and ethnic groups defined by shared culture. The UK defines ethnic groups based on individuals' national origins, and Colombia uses the term *etnia* (Spanish for ethnicity) for ethnic minority groups and *sin pertenencia étnica* (Spanish for no ethnicity) for the majority population, which shows European and American admixture. In the US, the definitions of racial and ethnic groups have changed twenty times since they were first used in the late 18th-century census, and similar changes have occurred over time in the UK and Colombia [8].

Nevertheless, the social and genetic dimensions of race and ethnicity are not mutually exclusive, and they can both be important for health outcomes. In particular, there is abundant evidence that socially defined racial and ethnic groups can show pronounced differences in the allele frequencies of genetic variants, including groups of variants that collectively influence polygenic traits, many of which may have clinical relevance [9–14]. It should be emphasized that these are differences of quantity rather than differences of kind; alleles differ in relative frequencies between groups with almost no alleles found universally present in one group and universally absent in another group [15–17]. Allele frequency differences between groups are a consequence of the fact that individuals' racial and ethnic identities are tied to their ancestral origins, which are in turn correlated with genetic differences. This fact is underscored by the NASEM report's definition of race, ethnicity, and ancestry as “descent-associated” groups composed of members who share characteristics based on common origins [3]. Groups with ancestors that evolved in different parts of the world were once reproductively isolated, and they may remain so owing to cultural differences, and thus accumulated genetic differences over time [18–20]. The resulting differences in the allele frequencies of pharmacogenomic variants among racial and ethnic groups have direct relevance to medical treatment decisions.

As a more pragmatic consideration, race and ethnicity are readily available for use by clinicians, whereas patient genetic data is substantially harder to come by. Later in this

perspective, we discuss how direct genetic assays capture pharmacogenomic variation far more reliably than race and ethnicity, which are poor proxies for individual-level genetic variants, but patients from socially disadvantaged groups are less likely to have access to these kinds of data than patients from majority groups. Finally, the use of race and ethnicity in pharmacogenomic research and clinical applications can help to avoid the kind of bias that arises when drug development and treatment decisions are made based on research conducted primarily on members of the majority group. If there are meaningful differences in the frequencies of pharmacogenomic variants between race and ethnicity groups, treatment strategies tailored to the majority group have the potential to do real harm to members of minority racial and ethnic groups. We believe that consideration of race and ethnicity when making treatment decisions, as we elaborate on below, can help to avoid these kinds of biases and their downstream consequences.

Here, we briefly review examples of pharmacogenomic differences among racial and ethnic groups from our own work in the US, the UK, and Colombia [21–25]. Most recently, we used population biobanks to interrogate the relationship between race, ethnicity, and pharmacogenomic variation in the US and the UK [25]. Socially defined race and ethnicity were found to be highly correlated with genome-wide patterns of pharmacogenomic variation, in support of the relevance of these categories for patient stratification. The cohorts for this study were taken from the US All of Us Research Program (*All of Us*; $n = 65,120$; Figure 1A) and the UK Biobank (*UKB*; $n = 31,396$; Figure 1B). Machine learning classifiers were used to predict biobank participants' self-identified race and ethnicity (SIRE) based on patterns of variation for 6311 pharmacogenomic variants in *All of Us* (Figure 1C) and 5966 pharmacogenomic variants in *UKB* (Figure 1D). Pharmacogenomic variation was found to predict participants' SIRE with 92.1% accuracy for *All of Us* and 94.3% accuracy for *UKB*. The highest group-specific prediction accuracy was 99.0% for the White group in *UKB*, and the lowest group-specific prediction accuracy was 92.9% for the Hispanic group in *All of Us*. We also found numerous individual pharmacogenomic variants with large allele frequency differences between racial and ethnic groups in these cohorts, consistent with previous studies.

In 2020, we analyzed a cohort of 8628 participants from the US Health and Retirement Study (*HRS*) to compare the ability of SIRE versus genetic ancestry (GA) to stratify pharmacogenomic variants [22]. We hypothesized that GA would provide greater resolution for pharmacogenomic variant stratification compared to SIRE, given that SIRE is socially defined whereas GA is a characteristic of the genome, which can be characterized objectively and with precision. Nevertheless, GA was found to predict SIRE with >96% accuracy, and thus GA provided a negligible increase in resolution for pharmacogenomic variant stratification. This study also confirmed the long-held notion that the majority of human genetic variation is found within rather than between racial and ethnic groups, a fact that is often taken to argue against the relevance of race and ethnicity to human genetic variation. Despite this pattern of pharmacogenomic variance partitioning, numerous pharmacogenomic variants were found to show significant frequency differences among SIRE groups, and genome-wide patterns of pharmacogenomic variation were highly concordant with SIRE. This study underscored the clinical relevance of SIRE for the stratification of pharmacogenomic risk among groups, while also showing far less utility of SIRE as a proxy for individual-level pharmacogenomic stratification. We explore the distinction between individual-level versus population-level pharmacogenomic risk stratification further in the following section on precision medicine versus precision public health.

In 2019, data from the ChocoGen research project was used to investigate population pharmacogenomics in the Colombian populations of Antioquia and Chocó in support of health equity in Colombia [21]. Antioquia and Chocó are neighboring departments (states) with distinct ancestry profiles: European followed by American ancestry in Antioquia compared to primarily African ancestry in Chocó. We found numerous pharmacogenomic variants with highly divergent allele frequencies between these two Colombian populations, with differences that were correlated to their distinct ancestral backgrounds. We used

these findings, working with local partners in Colombia, to develop and validate allele-specific PCR assays to test patients for population-specific pharmacogenomic variants. This approach allowed our Colombian partners to focus local resources on the populations where they are more likely to yield a return on investment, thereby serving as an example of how population pharmacogenomics can be leveraged to support precision medicine in resource-limited settings.

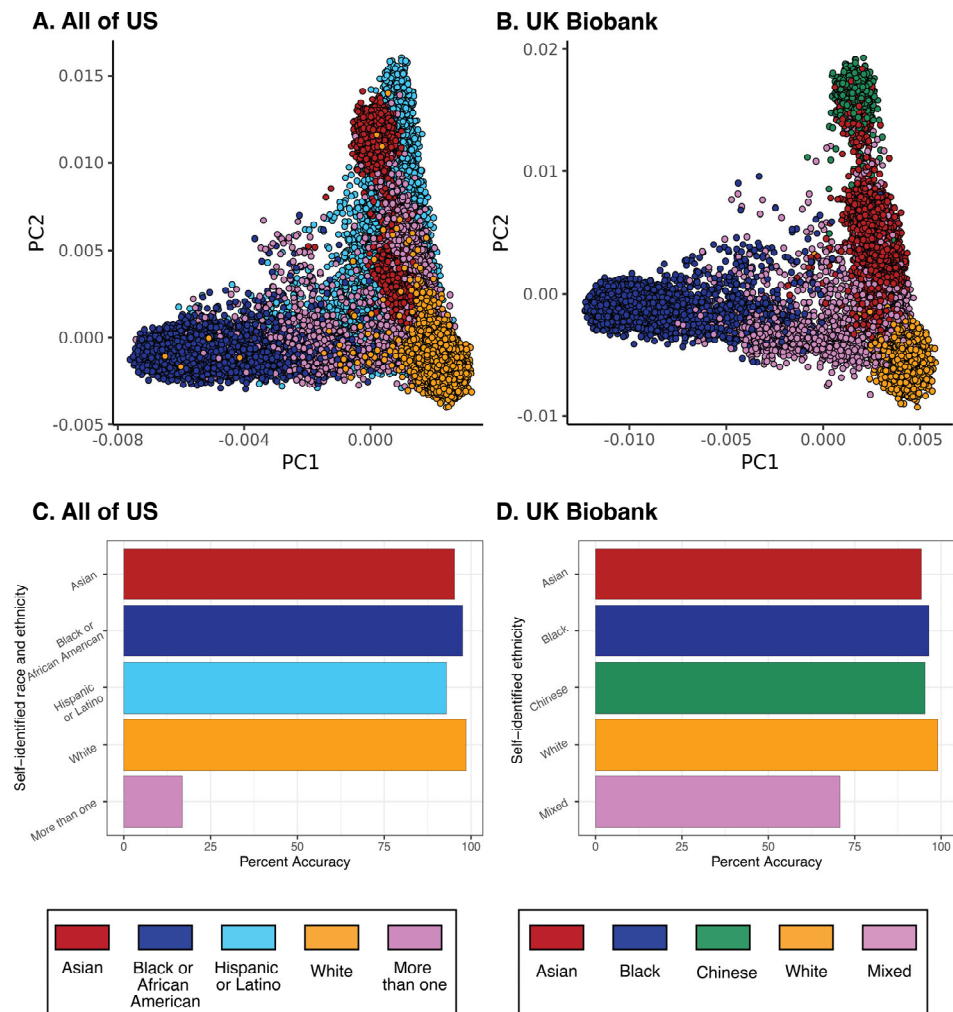


Figure 1. Pharmacogenomic variation corresponds to self-identified race and ethnicity. Principal component analysis for (A) *All of Us* and (B) *UKB* based on pharmacogenomic variants. Individual participants are shown as circles and color-coded by their self-identified race or ethnicity (see key below plot). The percent accuracy with which pharmacogenomic variation predicts race and ethnicity is shown for (C) *All of Us* and (D) *UKB*.

3. Global and Local Views on Race, Ethnicity, and Genetics

Race and Ethnicity Stratify Pharmacogenomic Variation at the Local Level but Do Not Represent Natural Groups That Correspond to Global Patterns of Human Genetic Variation

The racial and ethnic stratification of pharmacogenomic variation described in the previous section does not negate the fact that these groups are social constructs, which do not map well to global patterns of human genetic variation. At the global level, human genetic variation is largely distributed as a continuum, based on reproductive isolation by distance, with discontinuities introduced by major geographic barriers, like oceans, mountain ranges, and deserts [26]. Furthermore, the vast majority of human genetic variation can be found within the African continent, consistent with the fact that modern humans (*Homo sapiens*) evolved in Africa for more than 200,000 years before migrating

out of Africa and populating the rest of the world [27–29]. The deepest human lineages divide the Khoisan groups from all other human populations, with a divergence time of ~200,000 years, followed by the Rainforest Hunter-Gather split at ~150,000 years ago. West Africans and Eurasians who migrated out of Africa share a substantially more recent divergence ~75,000 years ago (note that all dates are approximate and subject to refinement). This means that people of the African diaspora in the US, UK, and Colombia, all of whom are descended from West African or Southwest African groups, are more genetically related to European descendants in these countries than either would be to Khoisan or Rainforest Hunter-Gatherers. This is despite the fact that Khoisan, Rainforest Hunter-Gatherers, and West African individuals would all likely be considered as Black in these countries, based on their phenotype and geographical origins, whereas European descendants would be considered as White.

The seeming contradiction of racially and ethnically stratified pharmacogenomic variation versus distinct global patterns of human genetic variation can be resolved by understanding how race and ethnicity are defined at the local level. As described in the previous section, different countries define race and/or ethnic groups based on the demographic characteristics of their own populations, which are in turn shaped by country-specific patterns of colonization and immigration [4–7]. Definitions of race and ethnicity are also informed by political considerations, including advocacy by groups that seek official recognition, such as the Hispanic/Latino group in the US, which was only fully recognized as distinct ethnicity starting with the 1980 census [30]. The distinct mix of native and immigrant groups that shape the demography of each country can be considered to represent an uneven sampling of (somewhat) continuous global genetic diversity, and it is this uneven sampling that yields meaningful genetic differences between socially constructed groups. This process can be illuminated using a number line analogy introduced by Joseph Graves [31]. Considering human genetic variation as continuously distributed along a number line from one to ten, if a country's population is generated by sampling discontinuously, around the values of one, five, and ten for instance, then this sampling would yield genetically distinct groups from an underlying continuous distribution of genetic variation. While this analogy is an oversimplification, it does reflect the kind of historical processes that generated the populations of modern, cosmopolitan countries and thereby explains how locally defined racial and ethnic groups can differ genetically in ways that are distinct from global patterns of human genetic variation.

4. Pharmacogenomics and Modifiable Risk Factors

The Characterization of Pharmacogenomic Variants Provides a Means to Elucidate Modifiable Risk Factors in Support of Health Equity

Clinical applications of pharmacogenomics are closely tied to the concept of risk stratification. A risk factor is anything that increases the chance of developing a health condition or disease. The characterization and measurement of risk factors can be used to inform disease prevention strategies and treatment programs. Risk factors can be characterized as environmental, behavioral, physiological, demographic, or genetic. Environmental risk factors include chemical exposures and social relationships, behavioral risk factors include smoking, diet, and physical activity, and physiological risk factors include weight, blood pressure, and cholesterol. Demographic risk factors relate primarily to race, ethnicity, and sex, and genetic risk factors are based on individuals' genetic makeup, i.e., their collection of genetic variants.

Effective health interventions require modifiable risk factors. Environmental, behavioral, and physiological risk factors are all (potentially) modifiable, and it makes perfect sense for public health strategies and medical interventions to focus on modifiable risk factors of this kind. A corollary, and seemingly reasonable, critique of demographic and genetic risk factors is that they are not (readily) modifiable. Demographic characteristics are considered to be socially ascribed, meaning that they are assigned at birth or involuntarily later in life. Since socially ascribed characteristics are neither chosen nor earned, they

are typically not considered to be modifiable. Genetic variants are inherited at birth and through the process of development manifest in the ~37 trillion cells of the human body. Thus, genetic risk factors are also non-modifiable. It should be noted that genetic risk factors may become partially modifiable in the not-too-distant future owing to developments in CRISPR or other gene editing approaches, but for the moment these technologies are not widely available in the healthcare setting. The key point with respect to epidemiology and health is that a focus on the discovery of non-modifiable demographic or genetic risk factors may not directly inform disease prevention and treatment.

Pharmacogenomics provides a direct link between non-modifiable demographic or genetic risk factors and modifiable environmental risk factors. This is because pharmacogenomic variants are defined as a special case of gene-by-environment interactions, specifically the interaction between genetic variants and medications used to treat disease. And medications are of course a kind of chemical environmental exposure. The effects of medications—with respect to dosage, efficacy, and toxicity—vary based on the presence or absence of specific pharmacogenomic variants. In this sense, the presence of pharmacogenomic variants in individual patients, or their enrichment in certain demographic groups, can suggest environmental modifications, i.e., changing or modifying drug prescriptions, that lead to improved health outcomes. The way that pharmacogenomic variants are assayed—directly in the case of pharmacogenetic tests or indirectly in the case of variant population frequencies—will determine the level of intervention. Precision medicine approaches require pharmacogenomic information for individual patients, whereas precision public health relies on pharmacogenomic information gleaned at the population level. While demographic data, for race and ethnicity in particular, are not reliable proxies for individual-level pharmacogenetic variation, they can serve as valuable guides for population-level screening and as risk factors for patient group stratification. This crucial distinction is discussed at greater length in the following section on precision medicine versus precision public health.

5. Precision Medicine Versus Precision Public Health

Population Pharmacogenomics Is an Essential Component of Precision Public Health, Where the Focus Is on Population-Level Variation as Opposed to Individual-Level Differences

Precision (genomic) medicine is an emergent medical discipline that involves the use of patients' genomic information to inform their clinical care, with respect to disease diagnosis, prognosis, and treatment [32,33]. Pharmacogenomic data is foundational to the treatment arm of precision medicine, which aims to deliver “the right medicine, to the right patient, at the right time”. In other words, precision medicine is squarely focused on individual-level pharmacogenomic information, i.e., the presence or absence of specific drug-associated variants in any given patient. Precision public health, on the other hand, relies on population-level data to inform public health strategies [34–36]. As it relates to pharmacogenomics, precision public health aims for “the right intervention, to the right population, at the right time.” Population pharmacogenomic profiles—data on the relative frequencies of pharmacogenomic variants within and between populations of interest—can be used to guide precision public health initiatives. For example, as we demonstrated in our Colombian study, population pharmacogenomic profiles can be used to target the provisioning of resources and testing where they are more likely to lead to improved health outcomes [21].

The distinction between precision medicine versus precision public health is key to understanding the utility, or lack thereof, of race and ethnicity as proxies for pharmacogenomic variation. Race and ethnicity are poor proxies for individual-level genetic variation and thus should not be used in place of direct pharmacogenomic testing for precision medicine. However, race and ethnicity do effectively stratify pharmacogenomic variation at the population level and therefore can be used to inform precision public health. This distinction is exemplified by the story of the nationwide rollout of the HIV medication efavirenz in Zimbabwe [37].

In 2015, Zimbabwe followed the World Health Organization's recommendation for HIV public health programs and switched from a three-drug cocktail to the single combination pill efavirenz. This change led to widespread adverse effects, including hallucinations, depression, and suicidal tendencies, with thousands of patients quickly abandoning treatment. It turns out that these adverse effects were associated with a pharmacogenomic variant that is found at anomalously high frequency in the Zimbabwe population. Approximately 20% of the population has two copies of the recessive efavirenz adverse effect associated variant, a fact which had been reported by Collen Masimirembwa, a local scientist working in the capital city Harare seven years earlier [38]. If Masimirembwa's advice on efavirenz dosing tailored to the local population had been heeded, the country could have avoided a public health crisis.

This example underscores the distinction between the individual-level prediction accuracy of population labels compared to their utility for population-level stratification. If Zimbabwe is taken as a population label and used as a surrogate to predict individual patients' adverse reactions to efavirenz, based on population pharmacogenomic data, it would have an extremely low prediction accuracy of 20% in the country. However, if the same pharmacogenomic data were used to inform public health interventions, they could prove to be extremely useful. Precision public health can leverage population pharmacogenomic data of this kind to inform population-wide treatment programs, directing pharmacogenomic testing where it is most needed and avoiding drugs that are predicted to cause adverse effects in a large fraction of the population. Race and ethnicity labels provide sufficient pharmacogenomic risk stratification to support the precision public health approach. We explore our own results on the relationship between race, ethnicity, and adverse drug reactions in the following section.

6. Adverse Drug Reactions

The Mitigation of Adverse Drug Reactions Is an Area Where Population Pharmacogenomics Can Have a Direct and Immediate Impact on Public Health

Toxicity-associated pharmacogenomic variants are linked to adverse reactions to numerous widely prescribed medications. As we have shown previously, differences in the frequencies of toxicity-associated pharmacogenomic variants between racial and ethnic groups have important implications for public health. Our approach to studying the implications of race and ethnicity for adverse drug reactions relies on the excess number of predicted adverse reactions per one thousand patients [24]. Given the population frequency of toxicity-associated variants, along with their documented effect mode as either dominant requiring a single copy of the toxicity-associated allele or recessive requiring two copies of the toxicity-associated allele, the number of predicted adverse effects can be calculated. Population stratification, using race, ethnicity, or any other population grouping approach, can then be used to calculate group differences in the number of predicted adverse reactions for medications with toxicity-associated variants.

Application of this approach to US racial and ethnic groups, using the *All of Us* and *HRS* cohorts, yields striking results [24,25]. The Black group shows up to 726 predicted excess adverse drug reactions per 1000 patients compared to the majority White group for the dominant effect mode and up to 635 excess adverse reactions for the recessive effect mode. The Hispanic group, which is more ancestrally similar to the majority White group, shows up to 351 predicted adverse drug reactions per 1000 patients compared to the White group for the dominant effect mode and up to 286 excess adverse reactions for the recessive effect mode.

By way of example, a pharmacogenomic variant in the *CYP3A4* (Cytochrome P450 Family 3 Subfamily A Member 4) gene has been associated with adverse reactions to methadone, which is used to treat heroin-dependent patients. In the *All of Us* cohort, the toxicity-associated allele of this variant (dbSNP rs4646437) shows high frequency in the Black group (0.73) compared to the majority White group (0.11), and it exerts a dominant effect on adverse reactions to methadone. Given these allele frequency differences

between groups, the dominant effect mode predicts 719 excess adverse drug reactions per 1000 patients in the Black group compared to the majority White group.

A pharmacogenomic variant in the *VKORC1* (Vitamin K Epoxide Reductase Complex Subunit 1) gene has been associated with anticoagulation and excess bleeding in patients treated with the blood thinners warfarin and phenprocoumon. In the *All of Us* cohort, the toxicity-associated allele of this variant (dbSNP rs9923231) shows high frequency in the Asian (0.67) group compared to the majority White group (0.34), and it exerts a dominant effect on adverse reactions to warfarin and phenprocoumon. This allele frequency difference, given the dominant effect mode, predicts 327 excess adverse drug reactions per 1000 patients in the Asian group compared to the majority White group.

A pharmacogenomic variant in the *DPYD* (Dihydropyrimidine Dehydrogenase) gene has been associated with toxic side effects to the chemotherapeutic agents fluorouracil and capecitabine. In the *All of Us* cohort, the toxicity-associated allele of this variant (dbSNP rs1801265) shows high frequency in the Black (0.40) group compared to the majority White group (0.22), and it has been reported to exert both recessive and dominant effects on toxicity. This allele frequency difference predicts 112 excess adverse drug reactions per 1000 patients in the Black group compared to the majority White group under the recessive effect mode and 248 excess adverse reactions under the dominant mode.

A pharmacogenomic variant in the *ERCC1* (Excision Repair 1, Endonuclease Non-Catalytic Subunit) gene has been associated with adverse reactions to nine different drugs, including cisplatin and oxaliplatin chemotherapy drugs. In the *HRS* cohort, the toxicity-associated allele of this variant (dbSNP rs11615) shows high frequency in the Black (0.88) and Hispanic (0.64) groups compared to the majority White group (0.37). These large allele frequency differences, given the recessive effect mode, yield predictions of up to 638 excess adverse drug reactions per 1000 patients in the Black group and up to 273 excess adverse reactions in the Hispanic group compared to the majority White group.

These findings on race, ethnicity, and adverse drug reactions also have implications for the clinical relevance of human genetic variance components. In 1972, Richard Lewontin found that the majority of human genetic variation was found within (85%) rather than between (15%) racial groups [39]. This fundamental result has been replicated many times since, and it is often taken to support the irrelevance of racial classification to human genetic variation [40,41]. This is consistent with the notion that race and ethnicity are purely social constructs with little or no biological significance [42]. Here, it must be reiterated that observed patterns of pharmacogenomic variation, particularly as they relate to adverse drug reactions, clearly support the clinical relevance of race and ethnicity [21,22,24,25]. In fact, we previously showed that when pharmacogenomic variation is partitioned exactly in the way that Lewontin and others have described, 85% within-group variation and 15% between-group variation, there can be up to 700 excess adverse drug reactions per 1000 patients predicted for the recessive effect mode and as many as 300 adverse reactions predicted for the dominant mode [24]. In other words, a high relative excess of within-versus between-group genetic variation, as is almost always seen for human populations, does not preclude the utility of race and ethnicity for pharmacogenomic risk stratification.

7. Conclusions: Embrace Genomic Diversity for Health Equity

A Reckoning with the Implications of Population Pharmacogenomic Diversity Is a Prerequisite for Health Equity

To conclude this perspective, we would like to emphasize once again the relevance of human genomic diversity for research efforts and clinical applications aimed at promoting health equity [43,44]. The extent of actionable pharmacogenomic differences between racial and ethnic groups flies in the face of an emerging orthodoxy, which maintains that race and ethnicity are social constructs with no biological or clinical relevance [3,42]. The ascendance of this strict social constructionist view of race and ethnicity points to a fundamental contradiction at the heart of genomics research efforts. On the one hand, there is an increasing awareness of the need to diversify genomics research cohorts in support of

health equity, driven by the realization that the current bias towards European ancestry cohorts will limit the global reach and impact of genomic medicine [45–48]. This view is exemplified by the National Human Genome Research Institute’s (NHGRI) pangenome initiative, which aims to expand the human reference genome to include representation from scores, and ultimately hundreds, of genome sequences from diverse human populations [49]. On the other hand, more and more genomics researchers are simultaneously advocating for the elimination of race and ethnicity in genomics research and for developing methods that emphasize the genetic sameness among populations rather than their differences [50–54]. This worldview stresses a continuum of human global diversity with population groups representing operationally defined abstractions rather than any real underlying structure in the data. There are even efforts to remove widely used terms like ancestry and admixture from the scientific lexicon, so as to avoid essentializing human genetic variation [55,56]. This body of work is part of a larger trend, which Jerry Coyne and Luana Maroja recently described as “the ideological subversion of biology” [57]. Much as we have done here, Coyne and Maroja soundly reject the ideologically motivated claim that there is no empirical value in studying genetic differences between races, ethnic groups, or populations, and they point out the threat that ideological research prerogatives of this kind pose to open scientific inquiry.

We contend that scientists should be able to accommodate more than one view on the constitution of racial and ethnic groups. Clearly, these groups are socially constructed, given that their boundaries and composition are delineated by humans. Accordingly, they differ with respect to non-genetic factors that are relevant to health outcomes, such as diet, lifestyle, and socioenvironmental factors. And just as clearly, socially defined racial and ethnic groups can and do differ genetically in ways that are relevant to disease treatment decisions. It follows that efforts to eliminate race and ethnicity from genomics research and clinical considerations, however well-intentioned, will ultimately do more harm than good, exacerbating rather than ameliorating existing health disparities. A more nuanced approach to race and ethnicity, one which recognizes both the social and genetic dimensions of these groups, can be effectively leveraged to support health equity following the roadmap laid out by NIMHD, promoting fairness and opportunity in disease treatments and health outcomes.

Author Contributions: Conceptualization, I.K.J. and L.M.-R.; methodology, I.K.J., S.S. and L.M.-R.; formal analysis, I.K.J. and S.S.; visualization, S.S.; writing—original draft preparation, review, and editing, I.K.J., S.S. and L.M.-R.; funding acquisition, I.K.J. and L.M.-R. All authors have read and agreed to the published version of the manuscript.

Funding: The work was supported by the IHRC-Georgia Tech Applied Bioinformatics Laboratory (RF383) to I.K.J., the National Institutes of Health (NIH) Distinguished Scholars Program (DSP) to L.M.-R., and the Division of Intramural Research (DIR) of the National Institute on Minority Health and Health Disparities (NIMHD) at NIH (1ZIAMD000016 and 1ZIAMD000018) to L.M.-R.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. National Institute on Minority Health and Health Disparities. What is Health Equity? Available online: <https://www.nimhd.nih.gov/resources/understanding-health-disparities/health-equity.html> (accessed on 15 July 2023).
2. National Institute on Minority Health and Health Disparities. Minority Health and Health Disparities: Definitions and Parameters. Available online: <https://www.nimhd.nih.gov/about/strategic-plan/nih-strategic-plan-definitions-and-parameters.html> (accessed on 15 July 2023).
3. National Academies of Sciences Engineering and Medicine. *Using Population Descriptors in Genetics and Genomics Research: A New Framework for an Evolving Field*; The National Academies Press: Washington, DC, USA, 2023.
4. Colombian Government. Population and Demography: Ethnic Groups. Available online: <http://www.dane.gov.co/index.php/en/statistics-by-topic-1/population-and-demography/ethnic-groups> (accessed on 15 July 2023).
5. United Kingdom Government. List of Ethnic Groups. Available online: <https://www.ethnicity-facts-figures.service.gov.uk/style-guide/ethnic-groups> (accessed on 15 July 2023).

6. United States Census Bureau. About the Topic of Race. Available online: <https://www.census.gov/topics/population/race/about.html> (accessed on 15 July 2023).
7. United States Census Bureau. About the Hispanic Population and Its Origin. Available online: <https://www.census.gov/topics/population/hispanic-origin/about.html> (accessed on 15 July 2023).
8. Pew Research Center. Race and Multiracial Americans in the US Census. In *Multiracial in America: Proud, Diverse and Growing in Numbers*; Pew Research Center: Washington, DC, USA, 2015; pp. 19–31.
9. Yuan, J.; Hu, Z.; Mahal, B.A.; Zhao, S.D.; Kensler, K.H.; Pi, J.; Hu, X.; Zhang, Y.; Wang, Y.; Jiang, J.; et al. Integrated Analysis of Genetic Ancestry and Genomic Alterations across Cancers. *Cancer Cell* **2018**, *34*, 549–560. [[CrossRef](#)]
10. Witherspoon, D.J.; Wooding, S.; Rogers, A.R.; Marchani, E.E.; Watkins, W.S.; Batzer, M.A.; Jorde, L.B. Genetic similarities within and between human populations. *Genetics* **2007**, *176*, 351–359. [[CrossRef](#)] [[PubMed](#)]
11. Tang, H.; Quertermous, T.; Rodriguez, B.; Kardia, S.L.; Zhu, X.; Brown, A.; Pankow, J.S.; Province, M.A.; Hunt, S.C.; Boerwinkle, E.; et al. Genetic structure, self-identified race/ethnicity, and confounding in case-control association studies. *Am. J. Hum. Genet.* **2005**, *76*, 268–275. [[CrossRef](#)]
12. Paschou, P.; Lewis, J.; Javed, A.; Drineas, P. Ancestry informative markers for fine-scale individual assignment to worldwide populations. *J. Med. Genet.* **2010**, *47*, 835–847. [[CrossRef](#)] [[PubMed](#)]
13. Jin, Y.; Schaffer, A.A.; Feolo, M.; Holmes, J.B.; Kattman, B.L. GRAF-pop: A Fast Distance-Based Method to Infer Subject Ancestry from Multiple Genotype Datasets Without Principal Components Analysis. *G3 Genes Genomes Genet.* **2019**, *9*, 2447–2461. [[CrossRef](#)] [[PubMed](#)]
14. Fang, H.; Hui, Q.; Lynch, J.; Honerlaw, J.; Assimes, T.L.; Huang, J.; Vujkovic, M.; Damrauer, S.M.; Pyarajan, S.; Gaziano, J.M.; et al. Harmonizing Genetic Ancestry and Self-identified Race/Ethnicity in Genome-wide Association Studies. *Am. J. Hum. Genet.* **2019**, *105*, 763–772. [[CrossRef](#)] [[PubMed](#)]
15. Rosenberg, N.A.; Li, L.M.; Ward, R.; Pritchard, J.K. Informativeness of genetic markers for inference of ancestry. *Am. J. Hum. Genet.* **2003**, *73*, 1402–1422. [[CrossRef](#)] [[PubMed](#)]
16. Rosenberg, N.A. A population-genetic perspective on the similarities and differences among worldwide human populations. *Hum. Biol.* **2011**, *83*, 659–684. [[CrossRef](#)]
17. Ding, L.; Wiener, H.; Abebe, T.; Altaye, M.; Go, R.C.; Kercsmar, C.; Grabowski, G.; Martin, L.J.; Khurana Hershey, G.K.; Chakorborty, R.; et al. Comparison of measures of marker informativeness for ancestry and admixture mapping. *BMC Genom.* **2011**, *12*, 622. [[CrossRef](#)]
18. Genomes Project Consortium; Auton, A.; Brooks, L.D.; Durbin, R.M.; Garrison, E.P.; Kang, H.M.; Korbel, J.O.; Marchini, J.L.; McCarthy, S.; McVean, G.A.; et al. A global reference for human genetic variation. *Nature* **2015**, *526*, 68–74. [[CrossRef](#)]
19. Cavalli-Sforza, L.L.; Menozzi, P.; Piazza, A. *The History and Geography of Human Genes*; Princeton University Press: Princeton, NJ, USA, 1994.
20. Bergstrom, A.; McCarthy, S.A.; Hui, R.; Almarri, M.A.; Ayub, Q.; Danecek, P.; Chen, Y.; Felkel, S.; Hallast, P.; Kamm, J.; et al. Insights into human genetic variation and population history from 929 diverse genomes. *Science* **2020**, *367*, eaay5012. [[CrossRef](#)] [[PubMed](#)]
21. Nagar, S.D.; Moreno, A.M.; Norris, E.T.; Rishishwar, L.; Conley, A.B.; O’Neal, K.L.; Velez-Gomez, S.; Montes-Rodriguez, C.; Jaraba-Alvarez, W.V.; Torres, I.; et al. Population Pharmacogenomics for Precision Public Health in Colombia. *Front. Genet.* **2019**, *10*, 241. [[CrossRef](#)] [[PubMed](#)]
22. Nagar, S.D.; Conley, A.B.; Jordan, I.K. Population structure and pharmacogenomic risk stratification in the United States. *BMC Biol.* **2020**, *18*, 140. [[CrossRef](#)] [[PubMed](#)]
23. Jordan, I.K.; Sharma, S.; Nagar, S.D.; Valderrama-Aguirre, A.; Marino-Ramirez, L. Genetic Ancestry Inference for Pharmacogenomics. *Methods Mol. Biol.* **2022**, *2547*, 595–609. [[CrossRef](#)] [[PubMed](#)]
24. Jordan, I.K.; Sharma, S.; Nagar, S.D.; Marino-Ramirez, L. The Apportionment of Pharmacogenomic Variation: Race, Ethnicity, and Adverse Drug Reactions. *Med. Res. Arch.* **2022**, *10*. [[CrossRef](#)]
25. Sharma, S.; Mariño-Ramírez, L.; Jordan, I.K. Race, Ethnicity, and Pharmacogenomic Variation in the United States and the United Kingdom. *Pharmaceutics* **2023**, *15*, 1923. [[CrossRef](#)] [[PubMed](#)]
26. Peter, B.M.; Petkova, D.; Novembre, J. Genetic Landscapes Reveal How Human Genetic Diversity Aligns with Geography. *Mol. Biol. Evol.* **2020**, *37*, 943–951. [[CrossRef](#)]
27. Schlebusch, C.M.; Malmstrom, H.; Gunther, T.; Sjodin, P.; Coutinho, A.; Edlund, H.; Munters, A.R.; Vicente, M.; Steyn, M.; Soodyall, H.; et al. Southern African ancient genomes estimate modern human divergence to 350,000 to 260,000 years ago. *Science* **2017**, *358*, 652–655. [[CrossRef](#)]
28. Ragsdale, A.P.; Weaver, T.D.; Atkinson, E.G.; Hoal, E.G.; Moller, M.; Henn, B.M.; Gravel, S. A weakly structured stem for human origins in Africa. *Nature* **2023**, *617*, 755–763. [[CrossRef](#)]
29. Choudhury, A.; Aron, S.; Botigue, L.R.; Sengupta, D.; Botha, G.; Bensellak, T.; Wells, G.; Kumuthini, J.; Shriner, D.; Fakim, Y.J.; et al. High-depth African genomes inform human migration and health. *Nature* **2020**, *586*, 741–748. [[CrossRef](#)]
30. Rodriguez, C.E. *Changing Race: Latinos, the Census, and the History of Ethnicity in the United States*; NYU Press: New York, NY, USA, 2000.
31. Graves, J.L. Biological Theories of Race beyond the Millennium. In *Reconsidering Race: Social Science Perspectives on Racial Categories in the Age of Genomics*; Suzuki, K., Von Vacano, D.A., Eds.; Oxford University Press: Oxford, UK, 2018; pp. 21–31.

32. Jameson, J.L.; Longo, D.L. Precision medicine—Personalized, problematic, and promising. *N. Engl. J. Med.* **2015**, *372*, 2229–2234. [[CrossRef](#)] [[PubMed](#)]
33. Collins, F.S.; Varmus, H. A new initiative on precision medicine. *N. Engl. J. Med.* **2015**, *372*, 793–795. [[CrossRef](#)] [[PubMed](#)]
34. Weeramanthri, T.S.; Dawkins, H.J.S.; Baynam, G.; Bellgard, M.; Gudes, O.; Semmens, J.B. Editorial: Precision Public Health. *Front. Public Health* **2018**, *6*, 121. [[CrossRef](#)] [[PubMed](#)]
35. Khoury, M.J.; Iademarco, M.F.; Riley, W.T. Precision Public Health for the Era of Precision Medicine. *Am. J. Prev. Med.* **2016**, *50*, 398–401. [[CrossRef](#)] [[PubMed](#)]
36. Khoury, M.J.; Bowen, M.S.; Clyne, M.; Dotson, W.D.; Gwinn, M.L.; Green, R.F.; Kolor, K.; Rodriguez, J.L.; Wulf, A.; Yu, W. From public health genomics to precision public health: A 20-year journey. *Genet. Med.* **2018**, *20*, 574–582. [[CrossRef](#)] [[PubMed](#)]
37. Nordling, L. How the genomics revolution could finally help Africa. *Nature* **2017**, *544*, 20–22. [[CrossRef](#)]
38. Nyakutira, C.; Roshammar, D.; Chigutsa, E.; Chonzi, P.; Ashton, M.; Nhachi, C.; Masimirembwa, C. High prevalence of the CYP2B6 516G→T(*6) variant and effect on the population pharmacokinetics of efavirenz in HIV/AIDS outpatients in Zimbabwe. *Eur. J. Clin. Pharmacol.* **2008**, *64*, 357–365. [[CrossRef](#)]
39. Lewontin, R.C. The apportionment of human diversity. In *Evolutionary Biology*; Springer: Berlin/Heidelberg, Germany, 1972; pp. 381–398.
40. Novembre, J. The background and legacy of Lewontin’s apportionment of human genetic diversity. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2022**, *377*, 20200406. [[CrossRef](#)]
41. Edge, M.D.; Ramachandran, S.; Rosenberg, N.A. Celebrating 50 years since Lewontin’s apportionment of human diversity. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2022**, *377*, 20200405. [[CrossRef](#)]
42. Flanagan, A.; Frey, T.; Christiansen, S.L.; AMA Manual of Style Committee. Updated Guidance on the Reporting of Race and Ethnicity in Medical and Science Journals. *JAMA* **2021**, *326*, 621–627. [[CrossRef](#)]
43. Oni-Orisan, A.; Mavura, Y.; Banda, Y.; Thornton, T.A.; Sebro, R. Embracing Genetic Diversity to Improve Black Health. *N. Engl. J. Med.* **2021**, *384*, 1163–1167. [[CrossRef](#)] [[PubMed](#)]
44. Marino-Ramirez, L.; Perez-Stable, E.J.; Jordan, I.K. Honour genetic diversity to realize health equity. *Nature* **2023**, *613*, 243. [[CrossRef](#)] [[PubMed](#)]
45. Sirugo, G.; Williams, S.M.; Tishkoff, S.A. The Missing Diversity in Human Genetic Studies. *Cell* **2019**, *177*, 1080. [[CrossRef](#)] [[PubMed](#)]
46. Bustamante, C.D.; Burchard, E.G.; De la Vega, F.M. Genomics for the world. *Nature* **2011**, *475*, 163–165. [[CrossRef](#)] [[PubMed](#)]
47. Popejoy, A.B.; Fullerton, S.M. Genomics is failing on diversity. *Nature* **2016**, *538*, 161–164. [[CrossRef](#)] [[PubMed](#)]
48. Petrovski, S.; Goldstein, D.B. Unequal representation of genetic variation across ancestry groups creates healthcare inequality in the application of precision medicine. *Genome Biol.* **2016**, *17*, 157. [[CrossRef](#)] [[PubMed](#)]
49. Liao, W.W.; Asri, M.; Ebler, J.; Doerr, D.; Haukness, M.; Hickey, G.; Lu, S.; Lucas, J.K.; Monlong, J.; Abel, H.J.; et al. A draft human pangenome reference. *Nature* **2023**, *617*, 312–324. [[CrossRef](#)]
50. Lewis, A.C.F.; Molina, S.J.; Appelbaum, P.S.; Dauda, B.; Di Rienzo, A.; Fuentes, A.; Fullerton, S.M.; Garrison, N.A.; Ghosh, N.; Hammonds, E.M.; et al. Getting genetic ancestry right for science and society. *Science* **2022**, *376*, 250–252. [[CrossRef](#)]
51. Carlson, J.; Henn, B.M.; Al-Hindi, D.R.; Ramachandran, S. Counter the weaponization of genetics research by extremists. *Nature* **2022**, *610*, 444–447. [[CrossRef](#)]
52. Biddanda, A.; Rice, D.P.; Novembre, J. A variant-centric perspective on geographic patterns of human allele frequency variation. *Elife* **2020**, *9*, e60107. [[CrossRef](#)]
53. Yudell, M.; Roberts, D.; DeSalle, R.; Tishkoff, S. NIH must confront the use of race in science. *Science* **2020**, *369*, 1313–1314. [[CrossRef](#)]
54. Yudell, M.; Roberts, D.; DeSalle, R.; Tishkoff, S. Taking race out of human genetics. *Science* **2016**, *351*, 564–565. [[CrossRef](#)]
55. Kampourakis, K.; Peterson, E.L. The racist origins, racialist connotations, and purity assumptions of the concept of “admixture” in human evolutionary genetics. *Genetics* **2023**, *223*, iyad002. [[CrossRef](#)]
56. Coop, G. Genetic similarity versus genetic ancestry groups as sample descriptors in human genetics. *arXiv* **2022**, arXiv:2207.11595.
57. Coyne, J.A.; Maroja, L.S. The Ideological Subversion of Biology. *Skept. Inq.* **2023**, *47*.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.