Group Categories in Pharmacogenetics Research

Lisa Gannett†‡

Current controversy over whether the Office of Management and Budget (OMB) system of racial and ethnic classification should be used in pharmacogenetics research as suggested by the U.S. Federal Drug Administration (FDA) has been couched in terms of realist-social constructionist debates on race. The assumptions both parties to these debates share instead need to be relinquished—specifically, dichotomies between the social and scientific and what is descriptive and evaluative/normative. This paper defends a pragmatic approach to the question of the appropriateness of the OMB group categories in pharmacogenetics research, an approach that is local and context-specific rather than global, incorporates practical and ethical as well as theoretical dimensions, and recognizes intersections of the social and the biological in the constitution of group categories.

1. FDA Guidelines and Realist-Social Constructionist Debates on Race. In January 2003, the U.S. Federal Drug Administration (FDA) published a draft "not for implementation" Guidance for Industry recommending that researchers use the Office of Management and Budget (OMB) system of racial and ethnic classification for the collection of racial and ethnic data in clinical trials. U.S. residents are familiar with this mode of classification because it is used by federal, state, and local government agencies in census-taking and other forms of data collection. The FDA recommendation is that the OMB system be used, not only for studies conducted

†To contact the author, please write to: Department of Philosophy, Saint Mary's University, Halifax, Nova Scotia, B3H 3C3, Canada; e-mail: lisa.gannett@smu.ca.

‡I would like to thank Wendy Hui Kyong Chun for the conference invitation which initiated this paper. Besides PSA 2004, versions of the paper were presented in Fall 2003 at Harvard University's Color Lines Conference and Canadian and Pacific SWIP conferences, and during 2004 to philosophy departments at Cape Breton University, Saint Mary's University, University of Tennessee-Knoxville, and University at Albany–SUNY, and the history and philosophy of science department at Indiana University. My thanks to conference organizers and all of these audiences for their helpful comments and generous hospitality, and to Melinda Fagan, Malia Fullerton, Jim Griesemer, and Kathleen Okruhlik whose contributions made this a better paper.

Philosophy of Science, 72 (December 2005) pp. 1232–1247. 0031-8248/2005/7205-0048$10.00

within the U.S. but elsewhere in the world (with the only modification for international use the replacement of the designation 'African-American' with 'African').[1] The proposed directive is of particular relevance to multinational pharmaceutical companies and the nascent field of pharmacogenetics. Although developments in this field were originally touted for their ability to provide drug therapies personally tailored to individual genomes, a number of researchers have come to believe that the investigation of group rather than individual patterns of DNA variability in drug responses might be more promising, at least at the outset.[2] The pharmaceutical company NitroMed, based in Lexington, Massachusetts, recently submitted research data in support of its application for FDA approval for a drug to treat heart failure in African-Americans. With this approval, NitroMed's product, called BiDil, would become the first 'ethnic drug' in what has been described as "the emerging field of race-based pharmacogenomics" (Henig 2004).

The FDA proposal has not gone without criticism. In an article published in *Science*, Susanne B. Haga and J. Craig Venter (2003) voice opposition to the FDA plan to carry out pharmaceutical testing using the OMB classification system, arguing that these are "social" not "scientific" categories of race and ethnicity. This is consistent with the view expressed by Venter at the June 26, 2000 White House press conference announcing the "completion" of the Human Genome Project that "the concept of race has no genetic or scientific basis" ("White House Remarks" 2000).[3] Haga and Venter believe the promise of pharmacogenetics lies in its implementation as individualized medicine given the likelihood that variation in drug responses will vary more within racial and ethnic groups than among them. This position finds support in longstanding scientific evidence that genetic differences among individuals irrespective of racial or ethnic group membership are greater than differences that exist among individuals taken as representatives of such racial or ethnic groups (Lewontin 1972; Nei and Roychoudhury 1974).

Other scientists have responded differently, not dismissing the validity

1. This reflects the FDA's aim that it be possible to assess the relevance of data collected on non-U.S. populations for U.S. populations.

2. This reflects more widespread attempts to overcome difficulties encountered in efforts to map complex (non-Mendelian) traits to the human genome. The International HapMap Project, for example, seeks to identify common descent, or haplotype, groups; because of this shared ancestry, such groups are believed likely to share allelic variants involved in differences in drug responses as well as disease propensity and behavior.

3. Venter founded Celera Genomics, the biotech start-up company that launched the sequencing race for the human genome against the publicly funded project. He is now with The Institute for Genome Research (TIGR).

of these group categories outright, but instead conducting empirical research to attempt to discern whether the OMB and other social systems of racial and ethnic classification are adequate to serve as 'proxies' for underlying biological groups such that the self-reported race or ethnicity of subjects might prove useful for pharmacogenetic studies. The technique of multi-locus genotyping is used to derive clusters on the basis of patterns of similarity and dissimilarity at variable sites across the genome (Pritchard, Stephens, and Donnelly 2000, 946).[4] Clustering is possible because, even if at each locus of the genome within-group differences far exceed between-group differences, across many loci, these small statistical differences accumulate. Individuals are assigned to either a single cluster or multiple clusters, in this latter case, belonging to a given cluster as a matter of degree. Genetic similarity is assumed to measure shared ancestry, and the correspondence of genetic clusters to the geographical origins/populations of sampled individuals is taken to provide empirical evidence for this assumption. It is the ability to infer in the opposite direction—that self-reported ancestry, or racial and ethnic affiliation, is a suitable proxy for genetic similarity—which is relevant to researchers with interests in pharmacogenetics. On the basis of the most extensive multi-locus genotyping study carried out to date, the genotyping of 377 autosomal microsatellite loci in 1,056 individuals from 52 populations, Noah A. Rosenberg et al. (2002) conclude that for most purposes "self-reported population ancestry likely provides a suitable proxy for genetic ancestry" (2384).[5]

Reporting on these results in the *New York Times*, in an article headlined "Gene Study Identifies 5 Main Human Populations," journalist Nicholas Wade (2002) surmises that the Rosenberg et al. findings provide scientific validation for popular conceptions of race, given the correspondence of genetic clusters to major geographical regions of the world. The apparent success of BiDil in treating cardiac failure in self-identified African-Amer-

4. Early inroads in this direction were made in the 1980s with the sampling and comparison of mitochondrial genomes in individuals from a number of populations to address questions concerning human origins and evolutionary history (including the famous 'mitochondrial Eve' hypothesis by Cann, Stoneking, and Wilson in 1987). Multi-locus genotyping enlists distance- or model-based computational methods for the derivation of genetic clusters.

5. In a similar study, Michael J. Bamshad et al. (2003) genotyped 100 *Alu* insertion polymorphisms in 302 individuals from 23 ethnic groups from sub-Saharan Africa, East Asia, and Europe, and an additional 60 microsatellites in 206 of the same individuals, and successfully assigned all of these samples to their continent of origin. Less success was obtained when 263 individuals from various Indian caste groups were also included. The authors concluded that "a proxy is sometimes, but not always, an accurate guide to population structure" (587).

ican research subjects has been interpreted by Sally Satel (2002), resident scholar at the American Enterprise Institute, and author of *PC, M.D.: How Political Correctness Is Corrupting Medicine*, as support for an albeit "rough" biological classification of race to which the "politically correct" are blind. Similarly, a recent *New York Times* article written by Wade (2004) poses the questions: "Is there a biological basis for race? If there is not, as many social scientists and others argue, how can a drug like BiDil work so well in one race?" In this way, the controversy over the use of the OMB or some other system of racial and ethnic classification in pharmacogenetics has come to be framed in terms of realist-social constructionist debates about race. The realist claims that our racial classifying practices identify things in nature, and that race is therefore a scientific and objective category, while the social constructionist asserts that race is not a genuine natural category but an invention of racialist/racist societies, hence subjective.

2. Getting Past Realist-Social Constructionist Debates on Race. It is unfortunate that the controversy that has arisen concerning the FDA proposal has taken on the tenor of pre-existing realist-social constructionist debates about race. Those who ask the question, regardless of the side of the issue on which they fall, share the assumption that a wedge can be inserted and maintained between the social and scientific, the cultural and natural, and what is subjective and objective. Along these lines, Haga and Venter dichotomize the social and scientific in their belief that because OMB designations are based on social group categories they are unsuitable for genetic research. Alan H. Goodman, who opposes the use of racial and ethnic categories in biomedical research (see Goodman 2000), is cited by Wade as reacting to the Rosenberg et al. study by saying that the genetic clustering results do not reflect race but geography. In doing so, Goodman opposes what is subjective and social (race) to what is objective and scientific (genetic clusters; geography), while at the same time ignoring that ancestry is reflected in the clusters and that geographical origin and ancestral relations contribute to prevalent social constructions of race. I contend that genetic clusters and geography are not objective and scientific in the way suggested either: multi-locus genotyping does not succeed in delineating an authoritative natural taxonomy of groups that exist independently of our social classifying practices.

It may seem at the outset that the question whether the OMB or some other system of social group classification serves as an adequate proxy for the underlying biological groups of interest to researchers in pharmacogenetics and other areas is straightforwardly empirical. It is an empirical question whether multi-locus genotyping will produce genetic clusters that are discrete enough to support a typology of groups for use in

biological research. It is an empirical question the extent to which multilocus genotyping will sort individuals who self-identify as members of the same social group or groups into the same genetic cluster or clusters. Rosenberg et al. recognize that these questions remain open since multilocus genotyping studies to date have sampled individuals who are well rooted within ancestral groups, and the result may have been a collection of groups more clearly demarcated from one another than would have been the case if individuals had been sampled indiscriminately.[6] However, my concern is not whether there is adequate discreteness to support the delineation of *some* typology of groups; it is whether there is a single authoritative typology of groups that would provide theoretical foundations for diverse research agendas, from the history of human evolution to the marketing of pharmaceuticals.

The Rosenberg et al. article, as suggested by its title "Genetic Structure of Human Populations," implies that there is a privileged representation of the genetic structure of *Homo sapiens*. Yet, it is not desirable to treat all of the genome's variable sites it is possible to genotype as equally relevant or informative—this depends on the research context. For the purpose of phylogenetic reconstruction, markers that are selectively neutral and/or functionally inert are preferable; it is well known that using adaptive rather than neutral traits for classification is misleading in the construction of phylogenies because of convergent evolution. The mutability of different kinds of markers makes a difference concerning the evolutionary period of interest: since microsatellites are more mutable than SNPs (single nucleotide polymorphisms), they are more useful for studying relatively recent evolutionary events and relationships. Biomedical researchers, in contrast, are interested in functional regions of the genome; for pharmacological studies, it is the genetic basis of physiological variability associated with drug responses that matters. Hence, the research context guides the scientist's choice of which genetic markers to use for clustering. And marker choice affects clustering results. In a given research context, it is not enough that clusters are genetically similar; what is important is that they are genetically similar in the relevant ways.

Population geneticists might argue, however, that evolutionary theory provides justification for their confidence that multi-locus genotyping dis-

6. David Serre and Svante Pääbo (2004) argue that study design did indeed bias the Rosenberg et al. results. They equalized population sizes using a subset of individuals from the Rosenberg et al. study, and substituted an assumption of uncorrelated for correlated allele frequencies, to find that the clusters generated do not correspond to the continents. Serre and Pääbo also compared clustering results for data sets obtained via different sampling strategies, and found that the continental groups produced when populations are sampled disappear when sampling is carried out geographically according to population density.

covers *the* genetic structure of *Homo sapiens* and yields an authoritative natural taxonomy of groups that exist independently of our social classifying practices. Recall the Rosenberg et al. assumption that the genetic clusters generated computationally on the basis of patterns of similarity and dissimilarity at multiple microsatellite loci across the genome reflect ancestral or genealogical relations: the more two individuals are genetically alike, the closer their genealogical ties. When computational analysis is carried out at multiple levels of resolution, it becomes possible to interpret the array of hierarchically-ordered sets of clusters produced phylogenetically. Previous multi-locus clustering studies have constructed phylogenetic trees based on genetic distances between individuals in order to explore evolutionary relationships among the human populations the genetic clusters are conceived as representing (Bowcock et al. 1994; Mountain and Cavalli-Sforza 1997).[7] Consequently, the biological reality of the clusters derived by multi-locus genotyping could be assumed to rest not in their genetic similarity but in their identification as phylogenetic groups. This justifies marker choice: though functional markers might seem more appropriate for multi-locus genotyping for pharmacological purposes, a phylogenetic group's unique evolutionary history means that it may exhibit characteristic allelic frequencies for any variable segment of the genome—functional or non-functional, adaptive or neutral—even if concordance among these is lacking. However, it is doubtful that the contingencies of our evolutionary past as a species fulfill the conditions required for phylogenetic reconstruction using multi-locus genotyping to yield a definitive typology of groups: an entirely branching human evolutionary history, complete subdivision into populations reproductively isolated from one another, and spatiotemporal continuity among the presumed biological populations of today and yesteryear.[8]

Nevertheless, all of us have ancestors, to whom we have genetic as well as cultural ties, and each of our ancestors lived somewhere on this planet. Genetic clustering tells us *something* about the ways in which individual genomes have been shaped by past, even long past, reproductive choices and environments. But what is it to say, further, that these genetic clusters themselves represent biological groups whose component individuals share

7. This represents a departure from traditional studies of human evolution where clustering methods are applied not to individuals but to groups in an attempt to discover the genealogical relationships between these groups by calculating genetic distances on the basis of similarities in allelic frequencies found at a number of loci (Cavalli-Sforza, Menozzi, and Piazza 1994).

8. Admittedly, this is an empirical question, and one which remains open. The ability to use genetic distances to construct phylogenetic trees provides no resolution, however, because the computational models used themselves assume branching and reproductive isolation (Templeton 1998). For further discussion, see Gannett 2004.

"genetic ancestry" or a common "ancestral geographical origin"? Which ancestors (matrilineal, patrilineal, recent, remote)? Which geographical locations (country, city, village, riverbank, valley, continent, island)? Research context determines the geographical scale or historical period of interest, and this will determine the desired level of resolution of the clusters, that is, the range in the number of clusters the computer is asked to generate.[9] The recent studies by Rosenberg et al. and others that have sought to ascertain patterns of distribution of human genome diversity are not 'pure' or 'basic' research devoid of practical aims: in some research context other than pharmacogenetics, and some society other than the U.S., the focus on (presumed) ancestry and on a handful of clusters that could correspond to OMB categories might have been directed elsewhere, to a different mode of clustering or level of resolution. A priori classification schemes, suited to the priorities of specific research contexts and embedded in social classifying practices, in this way, become incorporated in any typology of genetic clusters derivable by multi-locus genotyping.[10] Conceiving of the OMB classes as entirely social proxies for the underlying wholly biological entities of actual scientific interest is the wrong approach: there are no such entities. It would be better to recognize that the group categories most appropriate for pharmacogenetics will incorporate both biological and social relations. This may be just as well since social group categories are not easily expended in population-based human genetics research: in the U.S., biological tissues for genotyping are racially/ethnically coded by the OMB class assigned to the patient on hospital admission; social designations like "African American" or "American Indian" may be needed to recruit subjects and secure funding; there may be interest in studying gene-environment interactions; and so forth.

It need not imply that, to the extent the group categories are social, they are not scientific. In his *Pandora's Hope* essay "Circulating Reference: Sampling the Soil in the Amazon Forest," Bruno Latour (1999) describes the circulation of reference that occurs in the construction of objects of scientific knowledge. Latour asks: How is it that a published figure in a scientific journal manages to represent properties of soil at a location in the Amazon of interest to researchers who wonder if the forest is en-

9. Rosenberg et al. (2002) tested from two to six clusters; Bamshad et al. (2003) tested from one to six clusters.

10. Serre and Pääbo (2004) explicitly characterize the aim of the Rosenberg group as investigating "whether individuals can be assigned to culturally predefined populations on the basis of their genotypes" (1683). They argue that this led to a study design that biased results (finding cladistic rather than clinal patterns of genetic diversity) by sampling discontinuously at the extremes of continents instead of continuously across continents and focusing on those parts of the world that have contributed most to the genetic makeup of the U.S.

croaching on the savannah or the savannah on the forest? With the help of maps and compasses, a grid system is set up in the forest; soil samples are bagged and marked by depth and grid locations; portions of these soil samples are sorted by depth into a location-coded two-dimensional correlate system embedded in a wood frame; a standardized chart is used to assign a numerical number for soil color by depth and location; and ultimately a graphical representation is produced for publication. In this way, according to Latour, scientists connect world to word via a series of small steps he calls a "movement of abstraction." Each step in this series involves the extraction of matter to be used subsequently as form in the representation of a new phenomenon or referent and for the purposes of the next step's extraction. No "gap in representation" that must be bridged by "correspondence" arises because, along each of these steps, in each successive phenomenon, there is a "trace" of the former. "Traceability" means that the chain of reference is reversible, with truth-value circulating in both directions, word to world as well as world to word. Latour emphasizes the social network—the institutions, technologies, expertise, infrastructure, etc.—required to establish and to maintain the circulation of reference. If this network breaks down, reference disappears. On Latour's view, then, it makes no sense to ask whether race really exists apart from our social classifying practices because there is no such reality available as the object of scientific study.

3. Context Matters: A Pragmatic Epistemological Framework. Besides dichotomizing the scientific and social and what is objective and subjective, parties to realist-social constructionist debates about race share the belief that the question whether race exists or is socially constructed is wholly theoretical and descriptive, containing a propositional claim that, simply, is either true or false. This divorces the question from the variety of practical contexts in which classification by race and ethnicity might occur. The scientific study of 'group' DNA differences in pharmacogenetics requires categories of classification that adequately capture biological and social relations alike, and that remain salient, stable, and effective—at least sufficiently so—as they are transported across all sites comprising the research context: sites of the genome and social sites such as the university or research lab, the regulatory agency, the clinic, the advertising agency, the medicine cabinet. The question whether races *really* exist, with its scope global rather than local and theoretical rather than practical, is futile. My suggestion is that a pragmatic epistemological framework instead be adopted. Such a framework not only rejects dichotomies between the social and scientific and what is subjective and objective, but supports an approach that is local and context-specific, and attuned to practice in

its incorporation of evaluative and normative as well as descriptive dimensions.

In a recent paper titled "The ABO Blood Groups: Mapping the History and Geography of Genes in *Homo sapiens*," James R. Griesemer and I extend Latour's account by emphasizing the judgments that occur at each step along the chain of reference as it is established (Gannett and Griesemer 2004). The range of choices available is determined by the research context, and the judgments that are made reflect the privileging of some theoretical interests, practical aims, and value preferences over others. These choices cannot simply be ruled 'subjective' in a way that suggests that they detract from, rather than contribute to, the ability to make knowledge claims. Rather, such judgments make objectivity possible because, without them, the links of the chain of reference could not be established and the circulation of reference would be impossible. At each link in the chain of reference, there is an extraction of matter from its context and a subtraction in thought of some of its properties: selection of some individuals as subjects not others, sampling of blood or some other tissue; extraction of mRNA or nuclear or mitochondrial DNA; genotyping certain segments of the genome and not others; and so forth. Each accompanying judgment is the result of a contingent but reasoned choice among alternatives. It is possible to ask what choices are available and which choices are made and why.

This adds an evaluative dimension to Latour's descriptive account and supports a local pragmatic epistemology that is sensitive to the broader scientific, technological, and societal contexts. The relevant question about the adequacy of group categories of DNA difference is not simply descriptive but evaluative, that is, whether and how particular group categories work and not whether they are true (in the correspondence sense). It may well be that in certain biomedical contexts the continent upon which some minimum number of a person's ancestors lived during a particular historical period, say 500 years ago, provides a reasonable place to draw a boundary. But "nature" provides no authoritative response to the question whether this is the best way mode of representing species genome diversity. This will depend on the specific purposes for employing group categories of DNA difference.

Questions about group categories in pharmacogenetics need first to be localized on the genome. What we know about the statistical distribution of genome diversity in *Homo sapiens* suggests that the majority of genetically based traits that vary among us are found in similar proportions across space and time. Consequently, self-reported ancestry or phenotypic markings of race (skin color, hair texture, etc.), or membership in a genealogically defined cluster, will fail for the most part to predict anything of significance whether for individuals or the group as a whole. For

genetically-based traits that do occur at varying frequencies across space and time, patterns of distribution exhibit continuity far more often than discontinuity. Consequently, because of the substantial overlap between groups, reliance on self-reported ancestry or phenotypic markings of race, or membership in a genealogically defined genetic cluster, is quite limited in what can be predicted for individuals even if accurate predictions about group tendencies or average group differences are possible. More accurate predictions about individuals on the basis of race, ethnicity, or genetic clustering are possible for the relatively few genetically-based traits that do exhibit discontinuity across space and time such that overlap between groups is not substantial. But even if a rare allele is localized to a particular racial or ethnic group, if it is found in only a minority of group members, the ability to make predictions about individuals remains limited.

These patterns of distribution of human genome diversity suggest that, to the extent the OMB or some other system of racial and ethnic classification is useful in pharmacogenetics, it will be most so in those contexts in which attention is focused on groups not individuals. University and industry researchers in pharmacogenetics, and those concerned with public health, focus on populations and population-level regularities. They have reason to welcome any proxy assistance that racial or ethnic classification provides in delimiting groups which differ statistically in their genetic composition. Researchers who are interested in comparing drug efficacy and side effects in treatment and control groups worry about false positives arising in cases where alleles affecting drug responses differ in frequency between treatment and nontreatment groups because of population substructure. Although it would be preferable for researchers to rely on clusters derived from directly genotyping the loci for specific alleles involved in drug metabolism once these have been identified (whether OMB classes are adequate proxies for these specialized clusters remains to be seen),[11] there are additional features of the research context that encourage the use of racial and ethnic classification. The FDA recommendation that researchers use the OMB system of racial and ethnic classification in clinical trials satisfies its aim to achieve consistency with other governmental agencies (Schultz 2003). The pharmaceutical industry could make use of statistical patterns of allelic distribution or variation in nongenetic factors of physiological significance (from diet to stress to access to medical care)—or their interaction—to turn governmental reg-

11. James F. Wilson et al. (2001) use multi-locus genotyping (39 microsatellites on two chromosomes in 354 individuals from eight populations) to infer four genetic clusters found "broadly" to correspond to geographical regions. When the frequencies of drug-metabolizing enzyme alleles (11 variants at six loci) were compared across the clusters, cluster membership was found to be a better predictor for these than ethnic labels.

ulatory requirements to its advantage by patenting and marketing a drug that failed to meet the standard of efficacy or safety in the general population to a subpopulation where it may have nosed above the standard (see Duster 2003). The advertising aims of pharmaceutical companies are also implicated in assessing the utility of group categories in pharmacogenetics, especially because the U.S. is one of a small number of countries where the direct marketing of prescription drugs to patients is permitted by law. Potential consumers will not respond favorably to a television advertisement for a product that is FDA-approved for some racial or ethnic groups and not others, or is designed to treat or test for a condition that is more prevalent in some racial or ethnic groups than others, unless they self-identify with the group in question. This makes social categories of group identification indispensable to the pharmaceutical industry; the OMB classification is useful because Americans and others residing in the U.S., at least, become used to ticking off these particular boxes for themselves and applying the categories to others.

Where does this leave the physician and, most importantly, the patient? Clinicians treat individuals, not populations. Should FDA approval be forthcoming for a particular pharmaceutical product for use only in a certain racial or ethnic group because the necessary threshold of statistical significance regarding the presence of treatment effects and absence of side effects was met in industry tests only for that group, the designation has little or no relevance for patients in the clinical setting because such results are consistent with substantial genetic overlap among groups. The patenting and licensing of 'ethnic drugs' is less likely to protect patients and further their health than to aid their capture as a niche market of consumers. Nor are clinicians and patients likely to be served well by the advertising priorities of pharmaceutical companies, with clinicians left to explain to patients why the television commercial for the 'special' drug for African-Americans with hypertension or American Indians with diabetes is misleading or perhaps seduced themselves by glossy ads in a professional journal or a professional jaunt to a sunny locale into prescribing the drug to hypertensive or diabetic patients who self-identify as, or appear to be, African-American or American Indian. In the clinic, a patient's self-reported ancestry or phenotypic markings of race matter only if there are sequences of DNA involved in the metabolism of the particular drug the clinician is contemplating prescribing that exhibit minimal overlap in their distribution between OMB-classified racial or ethnic groups; since the frequency of any such alleles is likely to be low, the potential benefits or risks would have to be great for racial or ethnic identification to be decisive. More reasonably, racial or ethnic identification might be used to identify at-risk individuals, for whom direct genetic testing would follow.

In such situations, the transportability of group categories in pharmacogenetics between university and industry lab and clinic depends on not only the site of the genome and spatio-temporal distribution of relevant allelic variants but where the clinic is located and the clientele it serves, as well as the particularities of prevailing social constructions of race and ethnicity. Take carrier screening in the U.S. as an example. If a particular allelic variant arose or became common in a particular region of the world during a particular historical period, someone with ancestors who lived there and then has an elevated probability (however small) of carrying the allele compared to someone who has none, and this will increase with the number of such ancestors. To the extent that social constructions of race or ethnicity incorporate ancestral relations, a person's racial or ethnic identity may be at least somewhat informative, and this is reflected in the racialization of diseases like sickle cell anemia (as black) and cystic fibrosis (as white). How instructive these labels are depends on the particular ways in which race and ethnicity are socially constructed. The tendency to essentialize race and ethnicity by assigning each and every person to one and only one racial or ethnic group encourages clinicians to ignore multiple ancestral lines of descent, with the result that people may be screened for either sickle cell or cystic fibrosis but not both. The 'one drop of blood' rule potentially promotes the identification of a higher proportion of the sickle cell than cystic fibrosis carriers (since people with both African and northern European ancestry are more likely to be screened for a 'black' than a 'white' disease). How informative these labels are also depends on the clinic's clientele: in diverse communities, where people's ancestors lived in many places other than northern Europe and western Africa, sickle cell is not a 'black' disease at all, having arisen as an adaptive response to malarial environments whether outside or inside Africa (in the Mediterranean, for example), and only in some areas of Africa (not southern Africa, for example).

The emphasis the adoption of this alternate local pragmatic epistemological framework places on the utility of scientific categories of difference brings into focus the ends of research and also therefore the normative context in which it makes sense to ask whether the ends supported by the group categories are fair, good, and just. An explicitly normative approach is important for ensuring that the adoption of group categories is justified relative to all sets of interests, aims, and preferences, not just those of the pharmaceutical industry. Questions need to be asked about whether the FDA's adoption of OMB categories and its licensing of drugs for use in only some OMB-defined racial or ethnic groups will optimize patient care. Should the maximization of pharmaceutical profits or patient wellbeing win out if these fail to coincide? Because race was socially constructed as a natural category and racist oppression justified histori-

cally by appealing to natural or biological differences, intersections of the biological and the social in racialized group categories of DNA difference are unlikely to be innocuous. Harm may be done to people associated with communities oppressed and marginalized by racialism and racism if the very categories designed to track and rectify the effects of these wrongs become biologized through their incorporation in scientific research, clinical practice, and the marketing of pharmaceuticals. And yet, attention to health-related group differences need not perpetuate—and could even compensate for—the racist history that has seen some communities shoulder a disproportionate share of the burdens associated with biomedical research while reaping fewer of the benefits. At the same time, any such cost-benefit analysis in pharmacogenetics must also consider whether these new 'ethnic drugs' will be accessible to those who are not affluent and lack health benefits, and already disproportionately bear the risks of biologization. Potential harms arising through the globalization of social group categories specific to the U.S. and its particular history of racialism and racism need also to be assessed.[12]

4. Beyond ELSI. These ethical questions are no less a part of the research context than are the theoretical questions of the scientist or the practical aims of the pharmaceutical company. Here, I echo Camille Limoges' call for an "upstream" bioethics. As Limoges points out, "issues of considerable social and ethical relevance" are decided "far upstream from where most ethicists intervene" (1994, 124). Because ethical choices are implicated in the conceptualization of scientific research, and not only in its applications and the distribution of technologies that are developed as a result, Limoges suggests that a "downstream ethics" will be of "limited effectiveness" (1994, 124). Unfortunately, social and ethical issues surrounding categories of group DNA difference have been conceptualized in this downstream way by the NHGRI Ethical, Legal and Social Implications (ELSI) Research Program. The last of three five-year plans (1998–2003) makes no mention of scientific categories. Concerns are with

12. Approached from a pragmatic epistemological framework that incorporates evaluative and normative as well as descriptive dimensions, BiDil serves as more a cautionary than triumphal tale. BiDil is a combination of two generic products, the antioxidant and vasodilator hydralazine and the nitric acid donor isosorbide dinitrate, a dose of which costs about 44 cents (Maugh 2004). Why develop an expensive Cadillac-combo—tens of millions of dollars in venture capital have been invested in BiDil—when the same treatment benefits are already available at so little cost? As Jonathan Kahn writes: "The primary forces driving the re-invention of BiDil as an ethnic drug . . . were legal and commercial, rather than biomedical" (2004, 4). Investors' hopes are buoyed by patent protection rights: while the patent for BiDil for the general population expires in 2007, the patent for BiDil for African-Americans extends to 2020.

how "new genetic information" will affect racial/ethnic self-identification among the nonscientific public, whether members of certain racial/ethnic groups are likely to be less welcoming of genetic technologies or to interpret genetic information differently because of "past misuses" of science, and whether some racial/ethnic groups more than others are likely to be susceptible to genetic discrimination or to have less access to genetic services (Collins et al. 1998; see http://www.genome.gov/10001793). Although the NHGRI's most recent policy paper, "A Vision for the Future of Genomics Research" (Collins et al. 2003), continues to represent ethical concerns as downstream, as "research into the social consequences of increased availability of new genetic technologies and information" (2), the authors recognize that a better understanding of the relationships between genomics, race, and ethnicity requires a "critical examination of how the scientific community understands and uses these concepts in designing research and presenting findings" (10).

Impressive efforts have already been made in this direction: National Human Genome Center researchers at Howard University organized a May 2003 workshop, "Human Genome Variation and 'Race'—the State of the Science," and papers from the workshop were subsequently published in a November 2004 *Nature Genetics Supplement* titled "Genetics for the Human Race." While authors' recommendations on the use of social categories of race and ethnicity in biomedical research are varied, ranging from outright dismissal, to proxy use, to employment of skin pigmentation as scientific variable, among these are conclusions consistent with my arguments here. Write *Nature Genetics* editors: "Thinking of our genetic identities flexibly, gene by gene, level by level, region by region, with criteria for grouping examined and made explicit, fits well with the contingent way in which people describe their own ethnicity, ancestry, nationality, phenotype, family ties and genetic legacy" ("The Unexamined Population" 2004, S3). Bioethicists Mildred K. Cho and Pamela Sankar urge the upstream integration of ELSI considerations into science, arguing that "the awareness and involvement of scientists in thinking about downstream uses is needed at the earliest stages of research" given that "the way in which science is designed and carried out fundamentally affects how it can be used" (2004, S8). Less helpful is looking to "genetic reality" and "the competed Human Genome Project [to] define a concept of race that is scientifically credible" (Patrinos 2004, S1), or characterizing the scientific outcomes from which ethical, legal, and social implications follow as "inevitable" (Royal and Dunston 2004, S6).

I have argued that scientific outcomes are not inevitable but the product of contingent choices, and that these choices are shaped by ethical, legal, and social factors embedded within specific research contexts. There is no truth about race to be found in genetic data alone. In realist-social

constructionist debates about race, realists presume that the correct answer to the question 'are races real?' is 'yes' and look to biologists to furnish an adequate definition of a biological race concept and an authoritative natural taxonomy that might realign linguistic practices accordingly and provide theoretical foundations for social practices. Social constructionists presume that the correct answer to the question 'are (biological) races real?' is 'no' and suspect that realists about race are perpetrating covert racialist or racist agendas under the guise of science. Whether races exist or not is a futile question, if 'nature' is regarded as the source of a response that would be accepted as definitive. While I have focused on the significance of this question as it arises in pharmacogenetics, it is a question that those concerned with social policy often ask. I hope that this paper reinforces the view that social policies do not find justification in scientific representations of a primordial asocial nature. Instead, scientific representations of 'nature' should be more responsive to social policies.

REFERENCES

Bamshad, Michael J., Stephen Wooding, W. Scott Watkins, Christopher T. Ostler, Mark A. Batzer, and Lynn B. Jorde (2003), "Human Population Structure and Inference of Group Membership", *American Journal of Human Genetics* 72: 578–589.

Bowcock, A. M., A. Ruiz-Linares, J. Tomfohrde, E. Minch, J. R. Kidd, and L. L. Cavalli-Sforza (1994), "High Resolution of Human Evolutionary Trees with Polymorphic Satellites", *Nature* 368: 455–457.

Cann, Rebecca L., Mark Stoneking, and Allan C. Wilson (1987), "Mitochondrial DNA and Human Evolution", *Nature* 325: 31–36.

Cavalli-Sforza, L. Luca, Paolo Menozzi, and Alberto Piazza (1994), *The History and Geography of Human Genes*. Abridged. Princeton, NJ: Princeton University Press.

Cho, Mildred K., and Pamela Sankar (2004), "Forensic Genetics and Ethical, Legal and Social Implications beyond the Clinic", *Nature Genetics Supplement* 36: S8–S12.

Collins, Francis S., Eric D. Green, Alan E. Guttmacher, and Mark S. Guyer (2003), "A Vision for the Future of Genomics Research: A Blueprint for the Genomic Era", *Nature* 422: 1–13.

Collins, Francis S., Ari Patrinos, Elke Jordan, Aravinda Chakravarti, Raymond Gesteland, and LeRoy Walters (1998), "New Goals for the U.S. Human Genome Project: 1998–2003", *Science* 282: 682–689.

Duster, Troy (2003), "Unlikely Mix: Race, Biology, and Drugs", *San Francisco Chronicle* (March 17).

Gannett, Lisa (2004), "The Biological Reification of Race", *British Journal for the Philosophy of Science* 55: 323–345.

Gannett, Lisa, and James R. Griesemer (2004), "The ABO Blood Groups: Mapping the History and Geography of Genes in *Homo sapiens*", in Hans-Jörg Rheinberger and Jean-Paul Gaudillière (eds.), *Mapping Cultures of Twentieth Century Genetics*. New York: Routledge, 119–172.

Goodman, Alan H. (2000), "Why Genes Don't Count (for Racial Differences in Health)", *American Journal of Public Health* 90: 1699–1702.

Haga, Susanne B., and J. Craig Venter (2003), "FDA Races in Wrong Direction", *Science* 301: 466.

Henig, Robin Morantz (2004), "The Genome in Black and White (and Gray)", *New York Times* (October 10).

Kahn, Jonathan (2004), "How a Drug Becomes 'Ethnic': Law, Commerce, and the Production of Racial Categories in Medicine", *Yale Journal of Health Policy, Law, and Ethics* 4: 1–46.

Latour, Bruno (1999), *Pandora's Hope: Essays on the Reality of Science Studies*. Cambridge, MA: Harvard University Press.

Lewontin, Richard C. (1972), "The Apportionment of Human Diversity", *Evolutionary Biology* 6: 381–398.

Limoges, Camille (1994), "Errare Humanum Est: Do Genetic Errors Have a Future?", in Carl F. Cranor (ed.), *Are Genes Us? The Social Consequences of the New Genetics*. New Brunswick, NJ: Rutgers University Press, 113–124.

Maugh, Thomas H., II (2004), "Drugs for Only Blacks Stirs Hope, Concern", *Los Angeles Times* (November 9).

Mountain, Joanna L., and L. Luca Cavalli-Sforza (1997), "Multilocus Genotypes, a Tree of Individuals, and Human Evolutionary History", *American Journal of Human Genetics* 61: 705–718.

Nei, Matoshi, and A. K. Roychoudhury (1974), "Genic Variation Within and Between the Three Major Races of Man, Caucasoids, Negroids, and Mongoloids", *American Journal of Human Genetics* 26: 421–443.

Patrinos, Ari (2004), "'Race' and the Human Genome", *Nature Genetics Supplement* 36: S1–S2.

Pritchard, Jonathan K., Matthew Stephens, and Peter Donnelly (2000), "Inference of Population Structure Using Multilocus Genotype Data", *Genetics* 155: 945–959.

Rosenberg, Noah A., Jonathan K. Pritchard, James L. Weber, Howard M. Cann, Kenneth K. Kidd, Lev A. Zhivotovsky, and Marcus W. Feldman (2002), "Genetic Structure of Human Populations", *Science* 298: 2381–2385.

Royal, Charmaine D. M., and Georgia M. Dunston (2004), "Changing the Paradigm from 'Race' to Human Genome Variation", *Nature Genetics Supplement* 36: S5–S7.

Satel, Sally (2002), "I Am a Racially Profiling Doctor", *New York Times* (May 5).

Schultz, James (2003), "FDA Guidelines on Race and Ethnicity: Obstacle or Remedy?", *Journal of the National Cancer Institute* 95: 425–426.

Serre, David, and Svante Pääbo (2004), "Evidence for Gradients of Human Genetic Diversity Within and Among Continents", *Genome Research* 14: 1679–1685.

Templeton, Alan R. (1998), "Human Races: A Genetic and Evolutionary Perspective", *American Anthropologist* 100: 632–650.

"The Unexamined Population" (2004) (editorial), *Nature Genetics Supplement* 36: S3.

U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, Center for Devices and Radiological Health, and Office of the Commissioner (FDA) (2003), *Guidance for Industry: Collection of Race and Ethnicity Data in Clinical Trials*. http://www.fda.gov/cder/guidance/5054dft.pdf

Wade, Nicholas (2002), "Gene Study Identifies 5 Main Human Populations", *New York Times* (December 20).

——— (2004), "Race-Based Medicine Continued . . .", *New York Times* (November 14).

"White House Remarks on Decoding of Genome" (2000) (transcript), *New York Times* (June 27).

Wilson, James F., Michael E. Weale, Alice C. Smith, Fiona Gratix, Benjamin Fletcher, Mark G. Thomas, Neil Bradman, and David B. Goldstein (2001), "Population Genetic Structure of Variable Drug Response", *Nature Genetics* 29: 265–269.