Use and Misuse of "Race" in Biomedical Research

Karama C. Neal, PhD, MA

Adjunct Assistant Professor of Biology at Emory University where she teaches responsible conduct of reseraach. She is the editor of So What Can I Do, The Public Service weblog promoting ethics in action
http://www.karamaneal.com

Abstract

Scientific research provides substantial evidence that there is no genetic or biological basis for our social understanding of race. The confounding social and biological uses of race complicate its use in scientific and biolomedical research; thereby, prompting this discussion of how research informs the concept of race. Recommendations for the use of race in scientific study are provided. Scientists must critically consider when and how to use race in their research so that data are not compromised. Parsing of social, environmental, and biological contributors to health status will contribute to scientific advancement and improved medical care. Clear, consistent, and medically-relevant use of racial concepts in research promotes scientific responsibility, biomedical justice, and an improved social understanding of race. In addition, it is critical to correcting economic, health and other social disparities.

Keywords: race, ethnicity, genetics, research, ancestry

Use and Misuse of "Race" in Biomedical Research

Karama C. Neal, PhD, MA

I. Introduction

In over 30 years of research, scientists have shown numerous times using a variety of molecular and genetic methods that there is no biological basis for our common social understanding of race (Lewontin 1972; Barbujani, Magagni et al. 1997; Serre and Paabo 2004). A common definition of race is **"**One of the great divisions of mankind, having certain physical peculiarities in common." (Oxford English Dictionary, 2008). Though the term is used "imprecisely," (Oxford English Dictionary, 2008) its use suggests that human beings can be divided into discrete categories based on physical or biological properties, sometimes including, but not limited to skin color, hair texture, and eye and nose shape. This broad and vague definition allows it to be used by everyone from Nazi scientists to peace and justice activists to justify their beliefs and behavior.

Before the human genome was cloned, it was difficult to know if or how much genetics contributed to racial classification. That is, if the genetic differences associated with physical 'racial' differences (*e.g*. skin color, hair texture) are correlated with genetic differences that are *not* generally associated with racial divisions (*e.g*. lactose tolerance, height), then the physical differences we observe in the human population mirror the genetic differences across the genome between members. That would suggest that racial categories are biologically real, medically relevant and provide a good summary and proxy for medically relevant biological information. However, if racial differences are *not* similar to other genetic differences, particularly medically relevant ones, then racial categories are of little use in medicine and biomedical research.

Modern genomic science is being used to determine if such correlations exist, and genetic and genomic research supports the latter understanding of race, that is, racial differences *do not* correspond to other genetic differences. In recent work, Serre and Paabo (2004) show that human diversity forms a gradient or continuum rather than distinct clusters. They write, "There is no reason to assume that major genetic discontinuities exist between different continents or 'races' [quotes theirs]." Therefore, although there are differences in eye shape, hair texture, and other physical features, the differences in human biological processes are not very great, when they exist at all. Current research shows that the physical characteristics we use to assign race are distributed differently from other most biological characteristics, including medically relevant ones.

This does not mean that humans cannot be grouped in genetically meaningful ways. People can be classified based on various allele frequencies, but those groups do not correspond to racial categories, and are not necessarily medically relevant. (Cooper, Kaufman et al. 2003) In addition, a greater amount of variation exists within groups (~85%) than between them (~15%). (Barbujani, Magagni et al. 1997) Therefore, there is no genetic basis for the many ways in which race is understood and used in the United States and elsewhere.

Since recent genetic evidence suggests that human populations vary along gradients or clines and do not cluster into discrete races (Barbujani, Magagni et al. 1997; Serre and Paabo 2004), and because is notion of race is prevalent, fluid and arbitrary (as will be discussed below), it is reasonable to ask if and when race should be used in biomedical research, a discipline places high value on precise definitions, data integrity, and scientific reproducibility.

II. Problems with the concept of race.

Human variation is continuous, not discrete. Although a common understanding of race rests on notions of discrete population groups (Oxford English Dictionary, 2008; Dictionary.com 2005), modern scientific data support a continuous model of human genetic and physical variation. Such a model is inconsistent with discrete racial categories. This nondiscrete variation can be seen in physical characteristics, such as hair texture and skin color, and genetic characteristics, such as gene variants and noncoding sequences (Goodman 2000). "Nature has not created four or five distinct nonoverlapping genetic groups of people" (Ossorio and Duster 2005) which some call races.

Racial categories are arbitrary. Careful analysis shows that the physical characteristics used to establish racial categories are arbitrary. For various economic, political, social, and other reasons, racial groups are designated based on characteristics such as skin color, hair texture, and certain facial features. But why are eye and nose shape important and not hanging or attached earlobes or the presence of dimples? Why is skin color relevant and not eye color? Why is hair texture important, but not hair color, quantity or location?

The arbitrary nature of racial categories becomes even more obvious when admixture is considered. Consider the child of a black father and a white mother. According to the pre-1989 racial determination algorithms from the National Center for Health Statistics such a child would be classified as black. A sibling born to the same parents after 1989 would be considered white. (Anderson, Moscou et al. 2001) Note that these designations are not dependent on physical, genetic or other biological features of the children or the parents, but simply reflect "the social nature of racial identity" (Anderson, Moscou et al. 2001). Social and physical characteristics are applied arbitrarily in both biological and hereditary means of determining race. The lack of consistency makes concepts of race difficult to use in a reproducible manner.

Race terms are unclear. A primary goal of scientific writing and communication is clarity and precision. Careful language is required so that others can understand the research, replicate results, and apply the conclusions in practical and research protocols. Because racial terms like white, Spanish, Asian, African-American and Hispanic are often poorly defined or not defined it is almost impossible to ensure that a reader will understand a scientific paper in the way that the authors intended.

This problem is magnified when one considers the wide variety of ethnic communities and relationships in the world. But even if one limits the discussion to the US and to one 'race,' problems exist. Consider the term 'African-American.' This term appears to apply to people of African ancestry who live in or were born in the United States of America. (For brevity, I will ignore the fact that the entire Western hemisphere is America.) I fall into that category, as my ancestors for at least six generations were born in the US. Earlier ancestors were born in Africa. But what about the US-born child

of African-Brazilian immigrants? Or the US-born child of Ghanaian immigrants? Does it matter if the parents are South African whites or Algerians? How many African ancestors must one have to be classified as African-American? Does is matter if those ancestors speak Spanish?

Different people have different answers to these questions, and the answers may vary over space and time. With everyone from Barack Obama (CBS News, 2004) to Tiger Woods (Rosaforte, 1997, pg 181) to Teresa Heinz-Kerry (Washington Post, 2004) claiming the term 'African-American' at one time or another, and with others in disagreement with their use of it, it is difficult to know what exactly 'African-American' means, especially in the absence of clear definition. Since there is no agreement, it is confusing to use such terminology, particularly in scientific work where confusion and imprecision are to be avoided.

This is particularly true now that race is taken into account in drug development and marketing. Some drugs are known to act differently upon people of differing genotypes. However, for various reasons, including bias and lack of resources, tests may not be (made) available to assay genotype in a patient. Consequently, race is often used as a proxy for genotype. For example, 2004 advertisements for the asthma medication Advair warn that "Rare but serious asthma episodes and asthma-related fatalities occurred in a study with SEREVENT®, one of the components of ADVAIR. These risks may be greater in African-Americans."(Advair, 2005; News-medical.net, 2008) Perhaps, a particular group of alleles or cultural characteristics are associated with adverse risks, and these features are more frequent in 'African-Americans.' However, neither the advertisement nor associated prescribing information defines who is included in the group 'African-Americans.' How will physicians know who should be warned about the dangers of this drug? If the relevant alleles or cultural characteristics were studied and assayed in patients, that question could be easily answered, improving patient care.

Racial categories are fluid. Notions of race are rooted in the misunderstanding that people can be divided into relatively distinct groups based on biological characteristics that vary according to ancestral continent of origin. But *Homo sapiens* is one species, the members of which interbreed at will, often without regard to race assignment. In addition to the fact that early and recent work suggests that humans vary along gradients (Lewontin 1972; Barbujani, Magagni et al. 1997), people from different continents have been mating since the second group of hominids migrated from Africa. Racial groups have never been distinct and as migration increases and mating patterns change, our ideas about what people constitute a specific racial group change.

Consider the term "colored." Although it has fallen out of favor in the US, the meaning varies widely depending by time and place. In much of the US, "colored' refers to anyone of African descent, though in some parts of the country (*e.g*. New Orleans) the term was applied only to persons of African descent of a certain skin color. In South Africa, Coloured refers to people of "mixed" racial heritage.

Meyer and colleagues (Meyer, Parkin et al. 2003) studied cataract formation in the South African Coloured, specifically because "We believe that ethnic differences may be excluded as a confounding factor because the Coloured people of South Africa are recognized as a distinct ethnic group." That means that in South Africa the offspring of a black person and white person is of a different ethnic group (Coloured) than either of the parents. This particular group, as defined by the researchers, has not always existed. In

the US (at least until recently) the children of those parents would be considered black, though there is a movement to label people with multiple recent continental ancestries as 'multiracial.'

An examination of the US Census categories shows how governmental understanding of race has changed over time. In 1790, the race categories were "Free White Males, Free White Females, All Other Free Persons; Slaves." These divisions considered not just race, but also gender and bondage status (Nobles and Sciarra 2000). Two hundred years later the Census offered sixteen choices for "race," none of which refer to gender or bondage status (Nobles and Sciarra 2000). However, it doesn't take hundreds of years for official understanding of race to change. The Census Bureau notes the following changes from 1990 to 2000 (United States Census Bureau, 1999):

- In 2000, the term "Latino" is added to the question wording and response options "Spanish/Hispanic/Latino."
- A major change for the 2000 question was adding the instruction "Mark [X] one or more races to indicate what this person considers himself/herself to be." The 1990 question instructed respondents to "Fill ONE circle for the race that the person considers himself/herself to be."
- For 2000, the American Indian and Alaska Native categories were combined; in 1990, these were three separate categories American Indian, Eskimo, and Aleut. The 2000 version allows American Indians and Alaska Natives to write-in their tribal affiliation. In 1990, there was a write-in only for American Indians.
- For 2000, the Asian and Pacific Islander response categories have been split into two groups. Asian categories are listed in alphabetical order. Pacific Islander categories also are listed alphabetically, except that Native Hawaiian is the first category in the Pacific Islander list. The 1990 spanner for Asian or Pacific Islander was deleted in 2000.
- For 2000, the term "Chamorro" is added to the 1990 response option Guamanian, i.e.,"Guamanian or Chamorro." The race question in 2000 has three write-in lines, one for "American Indian or Alaska Native," one for "Other Asian" or "Other Pacific Islander," and one for "Some other race." In 1990, the race question had two write-in lines, one for "Indian (Amer.)" and one for "Other API" or "Other race.

Clearly, the genetic and biological composition of Americans did not change that significantly in 10 years, but the US government interpretation of race did. Temporal and spatial fluidity of perceived racial groups makes it difficult to racially classify people in ways that mean the same thing to different people. Perhaps that low level of precision is acceptable in some disciplines, but it is generally not allowed in scientific discourse.

Race is a proxy for other modes of classification. Scientific and biomedical researchers may use race terminology when they are really studying class, culture, skin color, diet, or some other feature of the population. Because race has biological *and* social connotations, it is difficult to know what is meant by specific race terms. Is a researcher interested in the biology associated with having dark skin? Or is she actually studying the lower economic status that many people with dark skin have, that may correlate with, but not directly cause, a certain health status? The problem is exacerbated

by the fact that many of these characteristics are correlated with each other making it difficult to determine which, if any, are the causative agents. Given this morass, scientists must be diligent about deciding exactly what they want to study and they must be certain that their research protocol addresses that question. Researchers who do not ask the question they want answered risk generating data that is irreproducible, and less meaningful and useful than it could be.

Race terms are frequently misused. In science writing, race is often used in *multiple* poorly defined ways. A single paper may discuss race as self-assigned by study participants, then use race terminology to describe certain genotypes, and finally use the same language to extrapolate their conclusions to massive continental groups (Sankar and Cho 2002). In addition, it is not always clear what is meant when terms associated with race are used in other ways. For example, work by Rijken and colleagues (2004) and Scanlon and Stubbs (2004) discuss the ways in which black skin responds to ultraviolet radiation and pressure ulcers, respectively. It is often unclear whether the writers mean black as a color, the way a physicist might use it, or black as a race or ethnicity, the way a social worker might describe it.

Even if one term is used with a consistent meaning, difficulties often arise when multiple terms are used together. This is because terms signifying race refer to many characteristics, including physical characteristics, ancestral country, ancestral continent, language, and even colonizing country of the ancestral country. It is not unusual to read a paper [*e.g*. (Perry, Rosenblatt et al. 2004)] that compares whites (defined by skin color) to African-Americans (defined by ancestry) to Hispanics (defined by common colonizer). These different ways of delineating populations may not be biologically meaningful, especially when compared to each other.

Defining race with geography and ancestry does little to solve the problem. Sankar and Cho (2002) give the example of comparing Asian-Americans and Mexican-Americans which "implies that people of Asian ancestry now living in the United States represent a level of genetic diversity that is equivalent to that of people of Mexican ancestry now living in the United States." That is not the case, and while that may not be socially or politically relevant, it is certainly biologically relevant and may well be medically relevant.

Race is not a good summary of or substitute for biological information. As discussed previously, the social and political means for racial classification do not adequately represent medically relevant biological information. The genes that determine the physical characteristics used to assign race often vary in dramatically different ways from medically relevant genes, such that racial differences do not necessarily mimic biological or medical differences. Ascertaining race is not a shortcut to understanding biology. If one is interested in biological phenomena, one must measure and describe biology.

Summary. The problems of understanding and using race, particularly as they relate to bioscience and medicine, are complicated and may be insurmountable. Feldman and colleagues (2003) concur and write "race is both too broad and too narrow a definition of ancestry to be biologically useful."

III. Reasons for the continued misuse of racial terminology.

In spite of data to the contrary, race continues to be used in biomedical research because of the mistaken belief that the social or geographic groups mimic biological groups. In light of the many problems associated with the misuse of race, and given the numerous calls for change, why do biomedical scientists continue to misunderstand and misuse the concept of race?

The language of race is familiar. Researchers may continue to use existing race terms for historical reasons. These terms, though confusing, are familiar and widely, if inconsistently, understood. Using them may constitute a bad, but breakable habit.

The language of race is a shorthand. Race terminology can provide a shorter, though less accurate way of describing populations, genotypes and other things. It is not unusual to read papers describing, for example, "Caucasian chromosomes" (e.g. Stephens, Reich et al. 1998). Though there are alleles and markers that occur more frequently in Caucasian people, "Caucasian chromosomes" suggests that there are chromosomes that are the exclusive property of European-derived Caucasian people. Given the state of racial admixture, the lack of a clear definition of "Caucasian people", and the racial gradient in which Caucasian people (however defined) exist, that is not the case.

Although I characterize myself and am usually classified by others as black and African-American, I have at least two great-great-great grandfathers who were white and so may well carry my own "Caucasian chromosomes." It is not unusual for self-described black people to have alleles that are more often found in self-described white or self-described Native American people (African Ancestry, 2008) and *vice versa*. In addition, Sankar and Cho (2002) note that these terms "make an inappropriate link between a rapidly changing social term and a fixed biological entity."

The language of race supports other beliefs. Sankar and Cho (2002) note that "conceptualizing race as a social construct has helped to undermine racism by eliminating its alleged natural basis." However, if writers and researchers do not want to undermine racism (for conscious or subconscious reasons), they may be less willing to change the way they conceive of and write about human population groups, even when scientific data supports doing so and when scientific responsibility and integrity mandate it.

For example, since the genetic determinants of most chronic diseases are unknown, one might hypothesize that genetic predispositions are found equally among all populations. This, however, is not what has happened. Blacks in the United States are "assumed to be genetically predisposed to virtually all chronic diseases" (Cooper, Kaufman et al. 2003). Cooper and colleagues (2003) list additional research problems that are unlikely to be coincidental, given the social status of the groups involved:

- Genetic causes or correlations to disease are suggested even in the absence of genetic data.
- Social and biological contributors are confounded and conflated.
- Correlations are used to make causal arguments, often based on insignificant or minimal data.

They suggest that "the correlation between the use of unsupported genetic inferences and the social standing of a group is glaring evidence of bias and demonstrates how race is used both to categorize and to rank order subpopulations."

In 1972, Lewontin (1972) wrote, " it has always been obvious that organisms vary, even to those pre-Darwinian idealists who saw most individual variation as

distorted shadows of an ideal. [Emphasis mine]" The idea that races form distinct groups can be used to support the belief that one of those groups is ideal or standard, while the rest are "distorted shadows." However, since human beings show a gradation of variation, it is more difficult to use biology to justify that one 'racial' group is significantly different (or better) than another. Scientists may, knowingly or unknowingly, seek to promote the mistaken idea that people are dramatically genetically different from each other, thereby leaving room for the notion that some people are "ideal" and others substandard.

Another common problem occurs when Europeans and those of European descent are used as that standard for comparison or as a control when studying non-European groups (*e.g*. Exner, Dries et al. 2001) and Chen, Rathore et al. 2001). The fact that the juxtaposition of black and white provides maximum contrast on paper does not mean that black people and white people are biological or medical opposites. In addition, white people are quite varied, consisting of British, Greek, Spanish, Australian, Canadian, and in some places, Moroccan, Iranian, and Turkish people, and may not always be the best control in health research (Bhopal and Donaldson 1998). Using these populations in this way may not only be scientifically irresponsible, but it also contributes to the incorrect notion that white people constitute a uniform 'standard' by which all other populations can or should be measured. It is a small step from such incorrect notions to conscious or subconscious beliefs in white supremacy.

IV. Alternatives to the misuse of race.

Several alternatives have been suggested, each of which offers advantages and disadvantages when used in biomedical research.

Replace race with ethnicity. 'Race' and 'ethnicity' are frequently used interchangeably. However, while 'race' has a presumed biological basis, 'ethnicity' is more broadly defined and does not imply biological meaning. Definitions of 'ethnic' include "racial, national, religious, linguistic, or cultural heritage" (Dictionary.com, 2008). When diseases like Tay-Sachs are shown to cluster in Ashkenazi Jewish populations, the reference is to ethnicity, not race (Cooper, Kaufman et al. 2003).

Depending on the context, 'race' and 'ethnicity' may have complementary usage, or mean exactly the same thing. Sankar and Cho (2002) suggest that ethnicity, when used to define a population socially or culturally, could appropriately replace race "when a researcher seeks a variable that corresponds to the behavioral aspects implied by the terms, such as diet, occupation, social status or health beliefs." In order for data to be universally understood and reproducible, the behavioral features of interest must be stated and measured, and not assumed.

Replace race with ancestry. Several groups (*e.g*. Rosenberg, Pritchard et al. 2002; Feldman, Lewontin et al. 2003) have suggested that continental or geographic ancestry, for which race is often used as a proxy, is a more useful concept for human classification in biomedical research.

For example, sickle cell anemia is often assumed to affect only people of African descent. If a person of African descent and a person of European descent present with symptoms of sickle cell, many physicians would rely on the different race of the patients to make preliminary diagnoses about their condition.

However, sickle cell also occurs in Mediterranean and South Asian populations (and is generally absent in Southern and Northern Africa), so it is inappropriate to limit ones scope to just African populations in regard to the disease. If a patient presents with symptoms consistent with sickle cell, a better approach is to ask if the person (regardless of their skin color or hair texture) has West or East African, Mediterranean or South Asian ancestry. Knowing race is insufficient; ancestry provides the necessary information to direct history taking and diagnosis.

Although ancestry may be of value in many instances, care must be taken to account for admixture and migration. Feldman and colleagues (Feldman, Lewontin et al. 2003) note "Many people thus have ancestry from more than one major geographic regions, meaning that the association of phenotype and geography breaks down." This is particularly important given the rise in commercial genetic testing designed to give users information about their ancestry (Bolnick et al. 2007).

Allow self-classification. Burchard and colleagues (2003) suggest that self-defined ancestry is a good proxy for genetically defined clusters. However, physical characteristics, like skin color, do not always correspond to ancestral continent of origin (Parra, Kittles et al. 2004). Self-reporting with open-ended categories will allow for more complex and accurate data. This complexity has always been necessary because of the continuous variation and group admixture in the human species. Still, biomedical researchers must still exercise care because even self-assigned racial groups may be discordant with most medical and biological characteristics. In addition, subjects may not be aware of or willing to reveal all relevant ancestral contributors.

Define race. Several attempts have been made to define racial groups for biomedical purposes (*e.g.* Anderson, Moscou et al. 2001; Bhopal 2004), but the definitions are not always consistent with each other and are not commonly accepted or used. One could argue that this does not matter as long as a particular research group uses clearly defined categories that allow replication of scientific results. However, such definitions must be carefully constructed so that they are precise, useful, and do not reinforce misunderstanding of the relationship between race and biology.

Cooper *et al*. (2003) note that "although everyone, from geneticists to laypersons, tends to use 'race' as if it were a scientific category, . . . no one offers a quantifiable definition of what race is in genetic terms." Indeed, genetic and other scientific research strongly suggests that such a definition is not possible. Nevertheless, biomedical researchers can use specific terms that impart more meaning. For example, "First generation US-born persons of Punjab ancestry who live in the southern US" is more informative than "Asian." More specific terms will allow for a clearer understanding of the scientific results. This allows data to be parsed, reproduced and applied in clinically- and scientifically-relevant and practical ways.

Replace race with measurable markers. In biomedical research and applications, a particular genotype is often at issue. If race is used as a proxy for the more relevant genotype, then the gene, and not the race of the person with the gene, is what should be studied and discussed. However, because we do not yet understand enough about the genetic basis of human health to identify and test for markers for each medical condition, we may not always be able to replace race with a measurable genetic entity. Sankar and Cho (2002) write, "genetic markers [for ancestry] might effectively replace race, but only if the markers chosen happen to be distributed among the selected populations in the

same way as the variable of interest in that particular study." Factors such as diet, economic class, access to health care, diet, religion, metabolic functions, health beliefs, and others that may affect medical outcomes may also need to be measured.

Identification of genetic markers for health status poses special ethical concerns. For such data to be medically useful, genetic information must be gathered from populations. In response to these concerns, the Quebec Network of Applied Genetic Medicine developed a "Statement of Principles on the Ethical Conduct of Human Genetic Research Involving Populations" (Cardinal, 2003). They identify ten "fundamental principles" and use them to make nine recommendations on the various stages on population genetic research, including recruitment, consent (community and individual), confidentiality, and commercialization.

The International HapMap Consortium, a multicenter collaboration that is developing tools that will facilitate work that relates human genetic variation to health and disease, has made a similar response (2004). In 2004, they wrote, "The Project raises many ethical issues because it will allow researchers to compare patterns of variation among both individuals and populations." They review ethical considerations involved in population identification, resource allocation, individual and community respect, and data release.

An analysis of these principles and recommendations is beyond the scope of this paper, but for genetic data to be used in beneficial ways, ethical procedures must be followed. This is important not only for the population being studied, but also for humanity at large. The HapMap Consortium (2004) writes, "Although there are differences among populations . . . it is important that the findings of the HapMap Project not be over-simplified to perpetuate social and historical stereotypes."

Recommendations: In biomedical research, race, ethnicity, ancestry, and continent or country of origin are often used to categorize the subject, and place her in a particular social or biological population group. However, the complex nature of ethnicity, culture and biologically definable population groups complicates this classification and leaves researchers with at least two difficult questions: First, how can we define population groups without resorting to race? Researchers should focus on markers that are relevant to the research questions. One should not assume, *a priori*, that ethnicity, or what was formerly misunderstood, as 'race' is necessarily a factor.

It is critical then to define groups by measurable features: genotype, electrolyte levels, melanin content, diet, etc. When such definitions are not possible, researchers should describe and define the group of interest as specifically a possible, using terms that refer to ethnicity, religion, continent of origin, social class, and the like. Although such descriptions will take more time and be more complex, they will more accurately reflect the complex nature of biological and cultural variation within the human species and allow data and results to be universally understood, reproduced and applied.

Second, when is it justified to use 'race' in biomedical research? The biological connotations of race are so deeply rooted and consciously and subconsciously assumed, that even social understandings of race is compromised by the lingering notion of biological determinism. That lingering notion, whether or not we are conscious of it, influences the ways we study and use 'race'. For that reason, I suggest that the only appropriate time to use the term 'race' is when discussing the human species (or some other species) as a whole, *i.e*. 'the human race'.

Although ethnicity is often used as a synonym for race, it has the advantage of not carrying biological meaning. Ethnicity is not commonly misunderstood as having genetic underpinnings. In addition, 'ethnicity' may include cultural, religious, environmental features and other measurable characteristics. These uses make it somewhat preferable to 'race'.

Without even a social usage of 'race', what are we left with? The NIH (National Institutes of Health and White House, 1997), US Census Bureau and others are beginning to acknowledge the relevance of factors other than race. Consider that the terms 'Hispanic' and 'Latino' used to be 'race' designations. Now, however, these are often considered ethnicities. Presumably, the rationale is the people who identify as 'Hispanic' or 'Latino' may have any variety of ancestral continents of origin. A similar approach could be taken with other population groups. Subjects could be asked to provide information such as grandparental continents or countries of origin, region of birth and residence, religion, and language. Admittedly, this does not address all aspects of culture, ethnicity, environment, and other medically- and socially-relevant factors but it would provide more of the information needed to ensure scientific integrity, reproducibility and clarity. In addition, it helps move away from the notion of 'race' as a biological determinant.

V. Reasons to rethink our use of race

Improved care and better science. Parsing biological, social and environmental contributors to disease will improve medical care and enhance scientific advancement. Though 'race' cannot be described in biologically meaningful ways, the concept still has social meaning. People who do not have a large amount of European ancestry often suffer discrimination. Countries populated by darker skinned peoples are generally less economically wealthy than countries populated by predominantly lighter skinned people. There are large health disparities between and among various ethnic groups in the US and abroad. People of predominantly non-European ancestry carry more of the burden of HIV and AIDS in the US and abroad. This is due, in part, to the fact that as many as 14% of those with predominant European ancestry carry an allele of the CCR5 gene that confers resistance to HIV infection (Stephens, Reich et al. 1998). However, most of the disparity exists because of confounding economic, social and political situations that negatively impact people whose ancestors primarily come from places other than Europe. These types of problems are unlikely to be solved by simply ignoring 'race' or ethnicity. Rather, they will worsen. Clear, consistent, nonbiased, appropriate, and medically-relevant use of biologically and socially relevant and definable population groups in health research is critical to correcting economic, health and other social disparities.

Scientific responsibility and biomedical justice. One understanding of justice is that, in whatever ways are relevant, like people should be treated alike (Beauchamp and Childress 2001). Although there is physical and genetic variation in the human species, it is often not so significant as to deem people of different 'races' as 'unlike' each other. Physicians, bioscience and health researchers, bioethicists, health care practitioners, and others have a responsibility to apply this principle of justice in their research and practice.

Cooper et al. (2003) write, "There is a tendency for scientists to ignore the messy social implications of what they do. At the extreme, the argument is made that 'we just tell the truth about nature,' and its negative consequences are political problems that do

not concern us. Whether or not such a position is defensible from an ethical point of view the debate over race cannot be sidestepped so easily." As a geneticist and a bioethicist, I argue that such views are *not ethically defensible*. Scientists "tell the truth about nature", but by determining which questions are worth asking and answering, we also decide which truth to tell. Inasmuch as scientists choose the things we study and how we study them, we share responsibility for the ways in which our research is used. Perhaps if we asked different scientific questions, and pursued the answers in different ways, our work would not have as many "negative consequences" or cause "political problems."

Clearer understanding of 'race'. Biomedical justice is attained when everyone, regardless of gender, class, mobility or disability status, sexual orientation, color, language, ancestral continent(s) of origin, ethnicity, culture, or 'race' equally benefit from and share the burden of biomedical research and application. A more precise, scientifically-grounded and socially relevant understanding and use of the concept of race is necessary (though not sufficient) for that occur.

Since the advent of genomics, there is renewed interest in assaying the genetic basis of 'race'. Genomic research has shown that differences within the same population account for 84.4% of the total genetic variance in the human species (Barbujani, Magagni et al. 1997), meaning that any two human beings are much more the same than they are different. However, these conclusions are not new. In 1972, Lewontin studied 35 blood groups and concluded, "Human racial classification is of no social value and is positively destructive of social and human relations. Since such racial classification is now seen to be of virtually no genetic or taxonomic significance either, no justification can be offered for its continuance" (Lewontin 1972). As producers of scientific knowledge and as creators of the often misused and misunderstood concept of race, biomedical scientists have a responsibility to reconsider the use of human racial classification to improve scientific integrity and promote biomedical justice.

Acknowledgments

I am grateful to Mark Kuczewski, Kwadjo Boaitey, Nana Gletsu, and Mawunyo Gletsu for helpful comments on early versions of this manuscript.

References

International HapMap Project. 2004. "Integrating ethics and science in the International HapMap Project." *Nature Review Genetics* 5(6): 467-75.

Advair. 2005 from http://www.advair.com. Accessed February 2, 2006.

African Ancestry. From http://www.africanancestry.com/patriclan.html. Accessed March 11, 2008.

Anderson, Matthew R., Susan Moscou, et al. 2001. "The role of race in the clinical presentation." *Family Medicine* 33(6): 430-4.

Barbujani, Guido, Arianna Magagni, et al. 1997. "An apportionment of human DNA diversity." *Proceedings of the National Academy of Sciences U S A* 94(9): 4516-9.

Beauchamp, Thomas L. and James F. Childress. 2001. *Principles of Biomedical Ethics.* Oxford, Oxford University Press.

Bhopal, Raj 2004. "Glossary of terms relating to ethnicity and race: for reflection and debate." *Journal if Epidemiology and Community Health* 58(6): 441-5.

Bhopal, Raj and Liam Donaldson. 1998. "White, European, Western, Caucasian, or what? Inappropriate labeling in research on race, ethnicity, and health." *American Journal of Public Health* 88(9): 1303-7.

Bolnice, Deborah A., Duana Fullwiley, Troy Duster, et al. 2007. "The science and business of Genetic Ancestry Testing." *Science* 318: 399-400.

Burchard, Estaban G., Elad Ziv, et al. 2003. "The importance of race and ethnic background in biomedical research and clinical practice." *New England Journal of Medicine* 348(12): 1170-5.

Cardinal, Genevieve, Mylene Deschenes, et al. 2003. "Statement of Principles on the Ethical Conduct of Human Genetic Research Involving Populations." from http://www.rmga.qc.ca/en/index.htm

CBS news. from http://www.cbsnews.com/stories/2004/11/23/earlyshow/main657289.shtml Accessed March 11, 2008.

Chen, Jersey, Saif S. Rathore, et al. 2001. "Racial differences in the use of cardiac catheterization after acute myocardial infarction." *New England Journal of Medicine* 344(19): 1443-9.

Cooper, Richard S., Jay S. Kaufman, et al. 2003. "Race and genomics." *New England Journal of Medicine* 348(12): 1166-70.

Dictionary.com. from http://dictionary.reference.com/search?q=ethnicity Accessed February 2, 2006.

Dictionary.com. (2005). From http://dictionary.reference.com/search?q=race Accessed February 2, 2006.

Exner, Derek V., Daniel L. Dries, et al. 2001. "Lesser response to angiotensin-converting-enzyme inhibitor therapy in black as compared with white patients with left ventricular dysfunction." *New England Journal of Medicine* 344(18): 1351-7.

Feldman, Marcus W., Richard C. Lewontin, et al. 2003. "Race: a genetic melting-pot." *Nature* 424(6947): 374.

Goodman, Alan H. 2000. "Why genes don't count (for racial differences in health)." *American Journal of Public Health* 90(11): 1699-702.

Lewontin, Richard C. 1972. "The apportionment of human diversity." *Evolutionary Biology* 6: 381-398.

Meyer, David, Donald P. Parkin, et al. 2003. "NAT2 slow acetylator function as a risk indicator for age-related cataract formation." *Pharmacogenetics* 13(5): 285-9.

National Institutes of Health and White House. 1997. From http://www.whitehouse.gov/omb/fedreg/ombdir15.html. Accessed March 11, 2008.

News-Medical.net. 2008. From http://www.news-medical.net/?id=741 Accessed March 11, 2008

Nobles, Anna Y. and Daniel T. Sciarra. 2000. "Cultural determinants in the treatment of Arab Americans: a primer for mainstream therapists." *American Journal of Orthopsychiatry* 70(2): 182-91.

Ossorio, Pilar and Troy Duster. 2005. "Race and genetics: controversies in biomedical, behavioral, and forensic sciences." *American Psychologist* 60(1): 115-28.

Oxford English Dictionary. race, n.2. http://dictionary.oed.com/cgi/entry/50195829?query_type=word&queryword=race&first=1&max_to_show=10&sort_type=alpha&search_id=7U4v-Gr64Kt-7715&result_place=1 Accessed March 11, 2008

Parra, Estaban J., Richard A. Kittles, et al. 2004. "Implications of correlations between skin color and genetic ancestry for biomedical research." *Nature Genetics* 36(11 Suppl): S54-60.

Perry, Arlette C., Evelyn B. Rosenblatt, et al. 2004. "Physical, behavioral, and body image characteristics in a tri-racial group of adolescent girls." *Obesity Research* 12(10): 1670-9.

Rijken, F., P. L. Bruijnzeel, et al. 2004. "Responses of black and white skin to solar-simulating radiation: differences in DNA photodamage, infiltrating neutrophils, proteolytic enzymes induced, keratinocyte activation, and IL-10 expression." *Journal of Investigative Dermatology*122(6): 1448-55.

Rosaforte, Tim. 1997. "Tiger Woods: The Makings of a Champion" St. Martin's Paperbacks, New York. Pg 181.

Rosenberg, Noah A., Jonathan K. Pritchard, et al. 2002. "Genetic structure of human populations." *Science* 298(5602): 2381-5.

Sankar, Pamela and Mildred K. Cho. 2002. "Genetics. Toward a new vocabulary of human genetic variation." *Science* 298(5597): 1337-8.

Scanlon, E. and N. Stubbs. 2004. "Pressure ulcer risk assessment in patients with darkly pigmented skin." *Professional Nurse* 19(6): 339-41.

Serre, David and Svante Paabo. 2004. "Evidence for gradients of human genetic diversity within and among continents." *Genome Research* 14(9): 1679-85.

Stephens, J. Claiborne, David E. Reich, et al. 1998. "Dating the origin of the CCR5-Delta32 AIDS-resistance allele by the coalescence of haplotypes." *American Journal of Human Genetics* 62(6): 1507-15.

United States Census Bureau. 1999. From http://www.census.gov/population/www/cen2000/90vs00.html Accessed March 11, 2008.

Washington Post. from http://www.washingtonpost.com/wp-dyn/articles/A58006-2004Aug11.html. Accessed March 11, 2008.