

POLICY BRIEF

RACE AND ETHNICITY DATA COLLECTION: BEYOND STANDARDIZATION

BROOKE CUNNINGHAM, M.D., PH.D.

RACE AND ETHNICITY DATA COLLECTION: BEYOND STANDARDIZATION

BROOKE CUNNINGHAM, M.D., Ph.D.

JUNE 2012

**HEALTH POLICY INSTITUTE
JOINT CENTER FOR POLITICAL AND ECONOMIC STUDIES**



Opinions expressed in Joint Center publications are those of the authors and do not necessarily reflect the views of the staff, officers or governors of the Joint Center for Political and Economic Studies or of the organizations that support the Joint Center and its research.

Joint Center for Political and Economic Studies, Washington, DC 20005

www.jointcenter.org

© 2012

All rights reserved. Published 2012

Printed in the United States.

FOREWORD

Racial and ethnic disparities in health status persist across the lifecycle, with many people of color suffering from poorer health from birth to death relative to national averages. A large and growing body of research illuminates the many complex causes of these disparities, which are primarily related to social and economic inequities that are rooted in historic and contemporary racism and discrimination. Many policymakers, healthcare professionals, and researchers, however, share an often-unstated assumption that racial and ethnic health inequities are the result of fundamental biologic or genetic differences. Not only are such assumptions incorrect, they point policy makers who seek to eliminate health inequities in the wrong direction.

The U.S. Department of Health and Human Services recently released new standards for race and ethnicity data collection, as mandated by the Patient Protection and Affordable Care Act of 2010. Defining the categories and requiring data collection have been important steps in furthering efforts to reduce health and health care disparities. But because policy makers are charged with responding to racial data, and particularly health-related race data, it is important to clarify how and why terms like “race” and “ethnicity” are relevant in medicine and health research.

This brief serves as a starting point for those who wish to better understand race, ethnicity, and the use of these categories in health research. This brief, prepared by Brooke Cunningham, M.D., Ph.D., a fellow in General Internal Medicine and a Greenwall Fellow at the Berman Institute of Bioethics at the Johns Hopkins University, reviews how the terms race and ethnicity evolved and provides an overview of genetics in order to explain why race is in fact not genetic. This distinction is important because the results of biomedical research are frequently reported by race without information on how “race” is defined. Readers, and at times authors themselves, often inappropriately interpret racial differences as differences at the genetic level.

This brief also reviews the history of race data collection by federal agencies. This history demonstrates the social construction of race by chronicling how race categories have changed over time. In addition, it reminds us that government agencies are crucial sites for determining what race is and how it will be used. Finally, the brief reviews major challenges to accurate data collection and recommends several actions that the federal government could take to move the field of health and health care disparities research to its next stage.

It is our hope that this brief will be helpful to policy makers, researchers, health professionals, and members of the general public who seek to better understand the complexity of race and ethnicity in America and to advance health equity for all.

Ralph B. Everett, Esq.

President and CEO

The Joint Center for Political and Economic Studies

EXECUTIVE SUMMARY

In October 2011, the Department of Health and Human Services (HHS) released new standards for race and ethnicity data collection. Defining the categories and requiring data collection have been important steps in furthering efforts to reduce health and health care disparities. Standardization should build our knowledge base by making it easier to compare and aggregate data from multiple sources. In addition, the standards are expected to improve data collection by state and local governments and the private sector, including health insurance corporations and hospitals.

Standardization, however, does not address the difficulties that researchers and policy makers have interpreting race and ethnicity data. This brief gives an overview of race and ethnicity that will hopefully create the foundation necessary for a more accurate understanding of the mechanisms through which race and ethnicity impact health. Our review of these concepts reiterates the social nature of race, clarifies the difference between race and ethnicity, and explains why race is not biological or genetic.

Although federal agencies have not clearly defined the terms “race” and “ethnicity,” the federal government has played a central role in how we think about race. Historically the Census Bureau and the Office of Management and Budget (OMB) have been most influential. The Census Bureau has been assigning individuals to “color” or “race” categories since 1790. The OMB first directed federal agencies with regard to practices in race and ethnicity data collection in 1977, and it subsequently revised its standards in 1997. HHS has been actively addressing race and ethnicity data collection by its constituent agencies since 1992. Responding to new federal laws passed in the 1990s, the National Institutes of Health and the Food and Drug Administration have required biomedical researchers to collect race and ethnicity data.

The latest HHS standards comply with those issued by the OMB and mirror data collection by the Census. While standardization is essential, the shortfalls of the OMB and Census categories are reproduced in the subsequent HHS standards. Researchers would benefit from further direction from HHS on formally defining the terms race and ethnicity, on collecting more granular ethnicity data, on analyzing multiple race responses, and on using indirect means for ascertaining race and ethnicity data. This brief makes several recommendations to further improve the accuracy and utility of race and ethnicity data. We hope that policy makers will find the brief helpful in planning next steps toward addressing health and health care disparities.

I. INTRODUCTION

Many nonwhite populations in the United States have poorer health relative to whites across a wide array of health measures (Agency for Healthcare Research and Quality 2011). Many also face greater barriers to accessing high-quality health care than whites, barriers that a growing number of public and private sector actors seeks to eliminate. These efforts, however, have been impaired by an inadequate understanding of race and ethnicity in the U.S. context. What defines a “racial” or “ethnic” group? Why does race and ethnicity matter when it comes to health research and health care? Absent clear answers to these questions, researchers, policy makers, and the general public frequently resort to antiquated assumptions about race, ethnicity, and human variability that obscure a more accurate understanding of the role that these factors play in contributing to health inequities.

The federal government and increasingly state and local health departments track health and health care disparities. In 1999, Congress passed the Healthcare Research and Quality Act, which mandated that the Agency for Healthcare Research and Quality (AHRQ) produce annual health care disparity reports. These annual reports have been produced by AHRQ since 2003 and summarize disparities in health care quality and delivery for racial and ethnic minorities and other groups (Agency for Healthcare Research and Quality 2011). Other initiatives such as the Institute of Medicine’s high-impact report, *Unequal Treatment* (Institute of Medicine 2003), and the Department of Health and Human Services’ Healthy People 2010 and 2020 programs also call for elimination of health disparities. Last October, in response to a mandate of the 2010 Affordable Care Act, HHS set standards for data collection for race, ethnicity, sex, language, and disability in order to better understand and to reduce health and health care disparities (Dorsey & Graham 2011; Office of Minority Health 2011). The standards are intended for population surveys and likely will be extended to health care delivery settings.

Because policy makers are charged with responding to “racial data,” and particularly health-related race data, this brief serves as a starting point for those who wish to better understand race, ethnicity, and the use of these categories in health. In the first half of the brief, the difference between race and ethnicity is reviewed and an overview of genetics is presented in order to explain why race is in fact not genetic. This is important because results of biomedical research are frequently reported by race.

Readers, and at times authors themselves, often inappropriately interpret racial differences as differences at the genetic level.

The second half of this brief reviews the history of race data collection by federal agencies. This history demonstrates the social construction of race by chronicling how race categories have changed over time. In addition, it reminds us that government agencies are crucial sites for determining what race is and how it will be used. Finally, in the last section of the brief, major challenges to accurate data collection are reviewed and several actions are recommended that the federal government could take to move the field of health and health care disparities research to its next stage.

II. UNDERSTANDING RACE AND ETHNICITY

WHAT IS RACE?

Common sense understandings of race as biological are strong and pervasive in both public and academic circles, despite genetic and social science evidence to the contrary. The scientific consensus is that race is not biological. Dupre (2008) makes this point when he states “...no serious scientist thinks these [racial] categories...have any biological grounding that could justify any claim to the status of natural kind” (p. 51). Rather, race is a social construct. This means that race is “a folk idea, a culturally invented conception about human differences” (Smedley & Smedley 2005, p. 22).

Foundational to the idea of race is that human variation is best understood by classifying people into groups according to easily discerned physical traits. Members of different racial groups are thought to be inherently different from one another in crucial ways. Group membership supposedly confers certain biological or cultural features upon us; these characteristics are considered unique to the group and invariably transmitted, often genetically, from one generation to the next.

Historically, group membership has been determined by such physical markers as skin color, hair texture, and facial features. We have been using these markers to divide ourselves into groups for so long that it seems natural to do so. However, the physical markers that have come to obviously mark one race off from another were not always so self-evident. Rather, human beings collectively have decided, and continue to decide, which

characteristics were (are) salient at a particular time and in a particular place (Cornell & Hartmann 2007).

Notably color has not always been the main marker of difference. The historical record shows that when groups encountered each other, they remarked upon color differences, but blackness was not universally disparaged and did not automatically mark inferiority (Drake 1987, 1990; Smedley 1998). However, the idea of race developed into a worldview when proponents of European expansionism and New World slavery needed such an idea to maintain hierarchical relationships and power in these new social systems.

In colonial America, the laboring classes initially consisted of indigenous Native Americans, white indentured and wage workers, and African laborers. Though there were negative stereotypes of Africans, “throughout much of the seventeenth century they received treatment only marginally different from that afforded other members of the ‘lower ranks’” (Kolchin 1993, p. 15). As the supply of indentured servants began to decrease, the planter class institutionalized chattel slavery in order to secure a stable labor force. Virginia, which served as a model for other colonies, began passing slave laws that differentiated black slaves from white servants. Those laws formalized slavery as a condition for life (Kolchin 1993). In 1676, a group of laborers—black and white united—unsuccessfully rose up against the planter class in Bacon’s rebellion. Subsequently, stronger laws were passed in Virginia that further reduced the social standing of blacks and served to form a “‘white race’ out of a heterogeneous and motley collection of Europeans *who had never before perceived that they had anything in common*” (Smedley & Smedley 2011, p. 115).

In addition to the slave laws, Enlightenment intellectuals cast Africans as inferior to Europeans. In the 18th century, scientists such as Carl Linnaeus and Johann Blumenbach were preoccupied with developing classification systems for plants and animals, and in the process rank ordered organisms in a hierarchical fashion (Krieger 2005; Zuberi 2001); philosophers such as David Hume and Immanuel Kant theorized that there were fundamentally different human types (Eze 1997). For these scientists, outward physical appearance was taken to be highly informative, correlating well with the inheritance of other physical and even cultural traits. As was done with other organisms, humans were then rank ordered according to their race, and Europeans were placed at the top.

Although slavery had existed since ancient times, these ideas supported a new form of slavery. In the New World, enslavement meant permanent inferior status, and slavery became racialized for the first time (Smedley & Smedley 2011). To achieve this, it was useful to link the idea of race tightly to biology, and, since then, race has significantly and differentially shaped the experiences and opportunities afforded individuals. Unfortunately, we now find it difficult to disentangle biology from race despite scientific and historical evidence that supports severing this link. In addition, with the increase in data stratified by race, some who have historically maintained a strong endorsement of race as a social construct have retreated from this position, and are now equivocal about race as a social versus biological reality (Fullwiley 2007b; Soo-Jin Lee 2008).

In summary, race is socially constructed because human beings developed the idea of race and have subsequently attributed variable meanings and significance to race. In addition, the markers that signify membership in one racial group versus another change over time and space, a fact that further illustrates that race is socially constructed. For example, a person may find that his or her racial classification changes as the context changes (Penner & Saperstein 2008). However flexibly race may (or may not) be employed by human beings, at the societal level, race tends to work in the same way across contexts. Race is used to restrict the rights and access of some (in the U.S., historically non-white persons) to resources, services, self-determination, and other opportunities, thereby affording privileged access to these same social goods to others (whites) (Bonilla-Silva 1997). Because race has been used to delineate individuals’ place in the world, it has been important for individuals and for the state to classify people in order to know how to treat them. Historically racial groups have been mutually exclusive for this reason.

WHAT IS ETHNICITY?

Because most of us do not know the history of race, the terms ethnicity and race are often used interchangeably (Afshari & Bhopal 2010; Oppenheimer 2001) and incorrectly. There are a few major differences between ethnicity and race, particularly as these ideas play out in the United States (Bonilla-Silva 1999; Cornell & Hartmann 2007; Ford & Harawa 2010; Omi & Winant 1994; Wimmer 2008). First, as noted above, the idea of race plays on ideas of biological difference. Although there is

indeed a sense of common descent, ethnicity, in contrast to race, speaks to cultural difference.

Ethnicity refers to clusters of people who have common cultural traits that they distinguish from those of other people. People who share a common language, geographic locale or place of origin, religion, sense of history, traditions, values, beliefs, food habits, and so forth, are perceived, and view themselves as constituting an ethnic group. (Smedley & Smedley 2005, p. 17)

Secondly, though individuals may self-report both race and ethnicity, and strongly hold to both types of identities, historically others (that is, society-at-large) confer racial assignments upon us. In contrast, ethnic identity originates from within the group itself (Waters 1990).

Race typically has its origins in the classifications that a dominant group imposes upon a less powerful collection of others. . . . [R]ace has been first and foremost a way of describing “others,” of making clear “they” are not “us.”...An ethnic group emerges only when that [ethnic] identity becomes part of the group’s own self-concept....[E]thnicity...often has its origins in assertion, in the claims groups make about themselves, instead of the claims others make about them. (Cornell & Hartmann 2007, pp. 28, 31)

Thirdly, because race has been a central way of organizing ourselves politically in the United States, one of the first things we do when we meet a stranger is attempt to place him or her into a racial category (Omi & Winant 1994). These racial assignments structure our interactions with one another, shaping both whether and how interpersonal interactions take place. In the United States, ethnicity does not shape interpersonal interaction, nor access to social opportunities, to the same extent. This is because “power is almost invariably an aspect of race; it may or may not be an aspect of ethnicity” (Cornell & Hartmann 2007, p. 31). Though today the relationships between Hispanic/Latino Americans and other groups are almost certainly structured by ethnicity, race historically has served as the primary means to parse out opportunity (Omi & Winant 1994; Waters 1999). Race historically has served as a marker of worth; such connotations may or may not be present with ethnicity.

Finally, in contrast to race, we historically have allowed, even encouraged, ethnicity to change because ethnicity primarily refers to cultural background. Indeed, many expect immigrants to assimilate, that is, to cast off previous customs for “the American way of life,” and to be absorbed into the American melting pot (Gordon 1964). The degree to which this occurs may be debated, but the point is that our *idea of ethnicity* allows it to happen. The *idea of ethnicity* as culture allows for the possibility of a change in culture. Through acculturation a member of an immigrant group (or his or her child) can come to be seen as an American, no hyphen necessary. In contrast, the *idea of race* as a biological variable does not allow for a black person (or his child) to come to be identified as white.

RACE IS NOT GENETIC

Despite awareness that race functions as an important social mechanism, biomedicine may be acting to reify race as biological (Azoulay 2006; Frank 2007; Lee 2009; Roberts 2011). Until 2004, the medical subject heading (MeSH) for “racial stocks” defined race as “major living subspecies of man differentiated by genetic and physical characteristics” (Sankar 2003). Many recent biomedical articles seem to suggest that the most important genetic differences are to be found between individuals from different racial categories. In this section we will briefly review why race is not genetic.

Nucleotides are the basic building blocks of DNA. A nucleotide is formed from a sugar, a phosphate group, and a base. The four bases are adenine (A), cytosine (C), guanine (G), and thymine (T). In forming the double helix, adenine always pairs with thymine; guanine always pairs with cytosine (Alberts 2008). SNPs (“snips”) are single-point nucleotide differences along the genome. For example, people may have an “A” at a site where others have a “C.” There are at least 10 million SNPs; thus SNPs are the most common form of human genetic variation (National Human Genome Research Institute 2012).

Most SNPs are found in all populations, though in different amounts or frequencies. This is because human genetic variation is clinal, which means that it diverges along spatial gradients in a continuous and overlapping way (Kittles & Weiss 2003). It follows then that populations geographically more distant from each other have more variation between them than populations that are geographically closer to one another. Nevertheless,

85-95% of all genetic variance occurs within populations. The variation between populations has been estimated to be 5-15%, and is increasingly thought to be at the lower end of this range (Cavalli-Sforza 2007).

That said, “there are no gene variants that are present in all individuals of one population group and in no individuals of another. . . . However frequencies of genetic variants and haplotypes differ across the world” (Bonham, Warshauer-Baker, & Collins 2005, p. 12). In instances where a SNP is more frequent in one population versus another, it is due to (a) recent emergence of the variant (also known as mutation) such that there has not been enough time for it to spread; (b) positive selection of the variant in a particular environment (i.e., having the variant gives you a survival advantage in a particular environment); or (c) to the fact that humans migrated out of Africa, taking only a subset of the overall genetic diversity found in Africa with them (Rotimi & Jorde 2010). Again, variants are not carried by all group members, and may only be present in a minority of group members.

Thus, there is no genetic evidence for nonoverlapping or homogenous racial groups. This refutes ideas that racial groups are discrete types or subspecies. While most scientists recognize that “race” does not make biological sense in terms of categorical fixed groups, there are some who would seek to define races as groups that differ in the frequencies of at least some genes. Those who adhere to this view argue that gene frequency differences can be used to divide individuals into clusters that resemble our current racial typology (Burchard et al. 2003; Risch, Burchard, Ziv, & Tang 2002). Critics counter that several aspects of this process of “clustering” are problematic, including the reference samples from which the SNPs are identified, the analytic software used, and the assumptions that researchers make in order to proceed (Bolnick 2008; Fujimura & Rajagopalan 2011; Roberts 2011).

In addition, self-reported race or ethnicity does not reliably predict which variants will be present in a person’s genome, nor does a person’s genome predict to which race he or she belongs (Feldman & Lewontin 2008; Rotimi & Jorde 2010). SNPs called ancestral informative markers (AIMs) do provide limited information about a person’s recent geographic (often described as continental) ancestry (Royal et al. 2010). Often the analyses of AIMs correlate well with self-identified race. At other times, results are incongruent. Neither finding should be

surprising. We have created racial labels that are based in part on continental geography, but because race is not genetic, everyone classified into the same racial group will not have the same AIMs. “Self-identified race is a [crude] surrogate for ancestral geographic origin, which is a surrogate for variation across the genome, which is a surrogate for variation in disease-relevant alleles, which is a surrogate for individual disease risk” (Bonham et al. 2005, p. 13). By nature, surrogate markers are imprecise.

Unfortunately, the facts that racial categories are socially constructed, that human beings share at least 99.5% of their genomes (National Human Genome Research Institute 2012), and that the cause and severity of the vast majority of traits and diseases cannot be attributed to any one gene acting alone often gets lost in how results are reported in studies (Dupre 2008). Race continues to be used as a proxy for genetic or biological variation, often without acknowledging the limitations and pitfalls of doing so (Fullwiley 2007a; Hunt & Megyesi 2008). As genomic sequencing becomes easier and cheaper, there should be less of a need to use race in this way. However, as has been the case thus far, it is very likely that many medical studies will continue to employ race as if there is a biological basis to it. This is problematic for understanding and ultimately improving the health of individuals and populations. Though the idea of race is grounded in an illogic of biology (Goodman 2000), race is a social category and remains most informative when used as a lens to understand social processes that impact health.

III. RACE AND ETHNICITY DATA COLLECTION BY FEDERAL AGENCIES

THE U.S. CENSUS BUREAU

We can see the social construction of race in action by reviewing the history of racial data collection by the U.S. Census Bureau. Race has been collected since the very first Census. However, initially racial data was collected as “color.”

In the first Census of 1790, the government counted free whites, Native Americans who were taxed, and enslaved black men and women (who were counted as 3/5 of a person). In 1820, “free colored” persons and “foreigners not naturalized” were also counted. In 1850, the government began placing those who had both black and white parentage in their own category, mulatto, and by 1890 went so far as to specifically estimate their

degree of black ancestry (e.g., quadroon, octoroon). Counting the multiracial population ceased in 1920. At that point the “one-drop rule” came into effect, and any perceptible African ancestry led to one being classified as black (Humes & Hogan 2009; Nobles 2000; U.S. Census Bureau 2002).

An opposite policy existed for Native Americans. For Native Americans, anyone with one drop Native American ancestry would *not* be considered “Indian” (Robbin 1999). Rather, the federal government, which had financial obligations to anyone designated as Indian, sought to limit how many people were categorized as such by “setting a lower bound on ancestry below which people would not be recognized as Native American” (Snipp 2003, p. 568).

The government began enumerating the Asian population in 1860 when it started counting Chinese immigrants in California. In 1870, the government extended the enumeration of Chinese individuals beyond California and also began counting the Japanese in California. It was not until 1890, following the 1882 Chinese Exclusion Act, that the counting of all Asian persons was extended beyond California (Humes & Hogan 2009; Snipp 2003).

In 1900, the heading for data collection changed from “color” to “color or race”; “color” was subsequently dropped from the category heading in 1950 (U.S. Census Bureau 2002). In its audit of the 1950 Census, the Census Bureau discovered that nonwhites had been undercounted by 12-13% by the observer report of enumerators (Snipp 2003). Concerned about accuracy, for the 1960 Census self-reported race was obtained by sending households Census forms in advance in the mail. The 1960 count for American Indian populations increased by 48% over the 1950 counts (Snipp 2003); since then self-report has been considered the best method to collect race data by the Census Bureau.

THE OFFICE OF MANAGEMENT AND BUDGET

Like the Census Bureau, the Office of Management and Budget impacts how we, as a country, think about race. By setting standards for the federal government, the OMB influences the way race data are collected and reported by other entities. Often other groups, including state and local governments, private sector researchers, and academicians, use government data or compare their data to that generated by the federal government.

In 1977, the OMB released *Statistical Policy Directive No. 15*, which instructed federal agencies on how to collect race and ethnicity data. In the 1990s, the OMB decided to revise the directive in order to address the increasing diversity of the United States (McKenney & Bennett 1994; Wallman, Evinger, & Schechter 2000). In 1997, the OMB released *Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity*. These standards were used in the 2000 Census; adopted by all other federal agencies by January 1, 2003; and remain current today.

A major difference between the current OMB standards and the previous directive is that individuals now can select more than one racial group with which to self-identify. In addition, in the 1977 directive there was one racial group for “Asian and Pacific Islander,” and now those two categories have been broken apart. Finally, the choices for ethnicity are now “Hispanic or Latino,” rather than simply “Hispanic,” to acknowledge broader use of the term Latino (Office of Management and Budget 1997).

OFFICE OF MANAGEMENT AND BUDGET STANDARDS FOR RACE AND ETHNICITY (1997)

ETHNICITY

- Hispanic or Latino**
- Not Hispanic or Latino**

RACE (Mark one or more)

- American Indian or Alaska Native**
- Asian**
- Black or African American**
- Native Hawaiian or Other Pacific Islander**
- White**

The OMB standards are considered the minimal requirements for data collection on race and ethnicity by federal agencies. The OMB permits the collection of more detailed information on population groups “provided that any additional categories can be aggregated into the minimum standard set of categories” (Office of Management and Budget 1997). For example the Census Bureau collects other categories of race data and rolls the data up to the OMB categories. Thus, instead of there being one “Asian” box to mark (as there is with “White,” “Black or African-American,” and “American Indian or Alaska Native”),

the race question in both the 2000 and 2010 Censuses lists the following groups which roll-up to Asian—Asian Indian, Chinese, Filipino, Japanese, Vietnamese, and Other Asian. Similarly there are four checkboxes listed for, and that roll-up to, Native Hawaiian or Other Pacific Islander—Native Hawaiian, Guamanian or Chamorro, Samoan, and Other Pacific Islander. In choosing to format the question in this way, the Census conflates ethnic groups with races, but likely does so to make the question meaningful for individuals (e.g., new immigrants) who may not see themselves as part of a larger “Asian” race (Humes & Hogan 2009). The Census also includes a “Some Other Race” category.

The OMB guidelines “underscore that self-identification is the preferred means of obtaining information about an individual’s race and ethnicity, except in instances where observer identification is more practical (e.g., completing a death certificate)” (Office of Management and Budget 1997). When the data are not collected by self-report, the OMB does allow for the collection of race and ethnicity in the same question. If data are collected through the preferred means of self-report, the OMB requires that separate questions for race and ethnicity be asked (Office of Management and Budget 1997).

THE NATIONAL INSTITUTES OF HEALTH AND THE FOOD AND DRUG ADMINISTRATION

Both the National Institutes of Health (NIH) and the Food and Drug Administration (FDA) require researchers to report study participants by race. In 1993, Congress passed the NIH Revitalization Act, which sought to address the low rates of inclusion for minorities and women in clinical research, in part because it had become clear that there could be real benefits to participation in research, and in part because there was/is a belief that people from different races respond differently to interventions (Epstein 2007). In 1997, in the Food and Drug Modernization Act, Congress similarly mandated that the FDA also ensure the inclusion of women and minorities in clinical trials.

The NIH and FDA subsequently released guidelines for researchers that emphasize the need for the analysis of subgroups in clinical studies (National Institutes of Health 2001; Food and Drug Administration 2005). For example, the FDA states that “differences [in response to medical

products] may be attributable to intrinsic factors (e.g., genetics, metabolism, elimination), extrinsic factors (e.g., diet, environmental exposure, sociocultural issues), or interactions between these factors” (Food and Drug Administration 2005). Although the FDA acknowledges that “extrinsic factors” can affect drug response, the basis for subgroup analysis becomes apparent upon examining the other groups for which subgroup analysis is recommended. With these other groups (e.g., children, the elderly, patients with renal failure), biological concerns, such as the potential for decreased clearance of medications from the body or for adverse effects of medications on development, require attending to dosing and safety.

A similar “biological reality” does not exist for race. As discussed above, human genetic variation is continuous, not absolute. That is, the genetic variants that have been linked to disease tend to be present in all groups, and in no race do all group members possess the same variants. Nevertheless race-specific drug labeling has already begun. Drugs are now indicated or contraindicated for whole groups of people because race is being used as a proxy for underlying biology (Bibbins-Domingo & Fernandez 2007; Ellison, Kaufman, Head, Martin, & Kahn 2008; Kahn 2008; Temple & Stockbridge 2007).

If a genetic variant is linked to drug response, then genotype, not race, should be used in prescribing drugs. Even so, a narrow focus on genes still neglects important mechanisms that lead to poor health, including environmental factors, behaviors, and gene-environment interactions (De Melo-Martín 2008). These factors likely have larger effects on morbidity, particularly for complex diseases such as hypertension, diabetes, and cancer, which are major contributors to health disparities. This overemphasis on biology casts aside what much social science and epidemiological research demonstrates about the relationship between race and health.

Race is a sociopolitical construct, but it nevertheless has biological consequences. Race is a “fundamental cause” of disease; it systematically shapes an individual’s exposure to risky environments and his or her ability to marshal the resources to effectively respond to illness once it occurs (Phelan, Link, & Tehranifar 2010). Thus, social scientists and epidemiologists have argued for the continued use of race in research studies in order to better understand how variables such as income, wealth, education, employment, insurance status, neighborhood resources, stress, and access to appropriate and timely care shape

health outcomes. Unfortunately, researchers typically do not control for these factors or their interactions well. They are then left with a “race effect” that may be erroneously attributed to biological differences (Kaufman 2008).

IV. THE NEW RACE AND ETHNICITY STANDARDS FROM THE DEPARTMENT OF HEALTH AND HUMAN SERVICES

HHS has been actively addressing data issues since at least 1992, when the Public Health Service released the report, *Improving Minority Health Statistics: Report of the PHS Task Force on Minority Health Data*. This was followed by a 1999 report authored by HHS workgroup members, *Improving the Collection and Use of Racial and Ethnic Data in HHS*. The department subsequently commissioned the National Research Council to review race and ethnicity data; that review resulted in the publication of a 2004 report, *Eliminating Health Disparities: Measurement and Data Needs* (Ver Ploeg & Perrin 2004). In addition, the National Committee on Vital and Health Statistics, a public advisory body to the secretary of HHS, also issued its report, *Recommendations on the Nation’s Data for Measuring and Eliminating Health Disparities Associated With Race, Ethnicity, and Socioeconomic Position* in 2004. As noted in the Introduction, the 2010 Affordable Care Act called for HHS to once again revisit race data collection, and the new standards were released in October 2011.

RECONCILING THE NEW HHS STANDARDS WITH THE INSTITUTE OF MEDICINE REPORT

Because the categories of race and ethnicity data collection adhere to the OMB categories, and mirror the Census, the limitations of the OMB categories and shortfalls of the Census are reproduced within the new HHS standards. For example, the new HHS standards do not incorporate fully the recommendations from the 2009 IOM report, *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement* (Ulmer, McFadden, & Nerenz 2009).

The strongest recommendation in the IOM report was to call for more granular, or detailed, ethnicity data. In doing so, the IOM committee acknowledged that Hispanic or Latino

ethnicities are not the only salient ethnicities in American society. Thus, the IOM committee recommended separate questions for race, Hispanic ethnicity, and granular ethnicity. The IOM committee made this recommendation after citing multiple examples from the literature that demonstrate different health outcomes for members of the same race but of different ethnic backgrounds.

HHS STANDARDS FOR RACE AND ETHNICITY (2011)

ETHNICITY DATA STANDARD

**Are you Hispanic, Latino/a, or Spanish origin?
(One or more categories may be selected)**

- No, not of Hispanic, Latino/a, or Spanish origin**
- Yes, Mexican, Mexican American, Chicano/a**
- Yes, Puerto Rican**
- Yes, Cuban**
- Yes, Another Hispanic, Latino, or Spanish origin**

RACE DATA STANDARD

**What is your race?
(One or more categories may be selected)**

- White**
- Black or African American**
- American Indian or Alaska Native**
- Asian Indian**
- Chinese**
- Filipino**
- Japanese**
- Korean**
- Vietnamese**
- Other Asian**
- Native Hawaiian**
- Guamanian or Chamorro**
- Samoan**
- Other Pacific Islander**

With the new HHS standards, more granular ethnicity data will be collected in federal health surveys for Asians and Native Hawaiians/Pacific Islanders. Data are collected on seven different Asian ethnicity categories and four Native Hawaiian/Pacific Islander categories. Unfortunately, like the Census, these ethnic identities are captured under a question that asks about race. The other groups—White, Black, and American Indian/Alaska Native—are left as monolithic groups.

Again, these are minimal standards. In theory organizations can collect more granular ethnicity data as they see fit. Ideally this data would be collected by expanding the ethnicity question, rather than by collecting ethnicity data from a question that asks about race. As is, the question conflates race with ethnicity. As explained above, although these are similar constructs, race and ethnicity do not measure the same thing.

The IOM recommends that researchers retain the flexibility to determine the granular ethnicity categories appropriate to their setting. Unfortunately, there are currently few incentives for organizations to become more informed about the communities in which they practice and do research. In addition, there are no mandates for regular review of these categories even though we know population demographics change over time. Although a consensus is emerging that more granular ethnicity data are better, HHS has not yet released guidance to help organizations and individuals determine when it is useful to collect more granular data, nor how to do so.

MARK ONE OR MORE

Of more consequence may be the lack of guidance regarding analyzing multiple race responses. Again in line with the OMB, the new HHS standards allow respondents to mark more than one race. As a result, there are 63 possible response combinations of the race category. Add in Hispanic/Latino versus not Hispanic/Latino and the number increases to 126 possible combinations (Perlmann & Waters 2002). Although multiple race selections make data analysis more difficult (Campbell 2007; Woo, Austin, Williams, & Bennett 2011), HHS rightly decided not to include a multiracial category, which would be a hodgepodge of individuals from very different backgrounds and would be minimally informative, if at all.

To monitor improvements in health and health care disparities, we will need to compare new data to historic data, in which respondents chose only one race category. In order to track progress on health disparities, health researchers will likely find that they need to allocate multiracial individuals back to single race categories (Mays, Ponce, Washington, & Cochran 2003). Unfortunately, the allocation back to a single race category is done without guidance from the survey respondent as to which, of all of their racial identifications, would be the most appropriate designation for them.

The methods for bridging data in order to evaluate health will need to be different from those used by the OMB to monitor and enforce compliance with civil rights legislation. In addition, depending on the method of bridging that is chosen for health outcomes, the results may vary tremendously, especially for small groups such as Native Hawaiians or Pacific Islanders and American Indians or Alaska Natives (Parker & Makuc 2002). Individuals who identify themselves as members of both minority and white racial groups may have better socioeconomic and health outcomes than those who identify only with the minority group. If organizations choose to allocate multiracial individuals to the minority group, it may appear as though disparities have lessened, when in fact no change has actually occurred.

To solve this problem, other surveys such as the National Health Interview Survey ask respondents reporting more than one race to select the one race with which they most closely identify (Tucker, Miller, & Parker 2002). Others have begun asking about socially assigned race. The question, “How do other people usually classify you in this country?” is a measure of socially assigned race developed by the Centers for Disease Control (Jones et al. 2008). This question has been used in the CDC’s Behavioral Risk Factor Surveillance System and was strongly considered for inclusion in the 2009 and 2010 National Health and Nutrition Examination Surveys. Jones et al. (2008) found that members of minority groups who self-identified as the minority group but were usually classified by others as white reported better health than those who self-identified as minority and were socially assigned to the minority group category. Although self-identification is the gold standard for race data collection, and may correlate highly with socially assigned race, race is arguably more about how others classify and treat you than about how you self-identify. Thus, as the Jones et al. (2008) study shows, questions that solely measure self-identity may not fully get at processes, like racism, that contribute to health and health care disparities.

STRENGTH OF STANDARDS

Although the health sector would benefit from HHS thoroughly resolving the above identified limitations, the Department made a few very important, and perhaps highly contested, decisions in its data collection standards. First, it did not include an “unknown” category. The risk of an unknown category is that individuals—for example, respondents wary of

how the data will be used or researchers uncomfortable about asking for race or ethnicity—will simply fill in unknown.

For similar reasons, we support the decision not to include a category “other” under race. The category “other” presents challenges (Burhansstipanov & Satter 2000). In the 2010 Census, the vast majority of those who selected other race identified their ethnicity as Hispanic/Latino (Humes, Jones, & Ramirez 2011). For those who do so, in part, this may be because Hispanic/Latino may be operating as a racial category for them. They may choose “other” because people more frequently classify them as Hispanic/Latino than any of the other OMB categories, and this classification as Hispanic/Latino likely shapes their life chances significantly (Amaro & Zambrana 2000; Hitlin, Brown, & Elder 2007; Roth 2010).

HHS agencies may request permission from OMB to add an option to write-in another race to interviewer-administered surveys. At best, the varied data entered into an “other” field requires guidelines regarding the allocation of these individuals to broad OMB categories, and the manpower to do so. At worst, analysts may report out data for a problematic “other group” or decide not to analyze these write-in responses at all. Appropriately, HHS has instructed the agency to recode the response as one of the HHS categories before publicly reporting the results. As the standards are extended into other sectors, HHS should continue to advise organizations and researchers not to report out data as “other.”

VI. CONCLUSION

The new HHS standards are a good and necessary start to standardize race and ethnicity data collection in the health sector. Standardizing data collection will improve our ability to monitor health and health care disparities and to compare data from multiple studies. HHS will likely extend these standards to health care delivery settings that receive federal funding. Nongovernmental organizations will likely follow suit and adopt the standards as well.

Formalizing the data categories has been an important first step. However, problems with the collection and interpretation of race and ethnicity data will persist as long as individuals are free to employ whichever definition of race and ethnicity that they wish. Medical editors instruct authors to define their use of the

terms race and ethnicity, but most authors still avoid doing so (Lee 2009).

Unfortunately, HHS has not commented on what race and ethnicity are. In the 1997 revisions, the OMB hedged, stating “the categories represent a social-political construct,” “are not anthropologically or scientifically based,” and “should not be interpreted as being *primarily* [emphasis added] biological or genetic in reference. Race and ethnicity may be thought of in terms of social and cultural characteristics as well as ancestry.” Though the OMB emphasizes the social nature of race in this statement, clearly it has also left the door open for race to be interpreted biologically, even if only secondarily.

However, by not strongly reiterating the scientific consensus that racial data are informative as a measure of social, not biological, realities, we may be stifling progress, particularly in addressing health and health care disparities. Physicians and biomedical researchers would benefit from more guidance about what race represents, means, and can measure. Several authors have made useful recommendations about race and ethnicity data collection in health surveys and biomedical science (Bilheimer & Klein 2010; Bhopal 2006; Caulfield et al. 2009; Hahn & Stroup 1994; Kahn 2006; Mays et al. 2003; Shields et al. 2005; Williams 1999). The following recommendations focus on next steps that the federal government can take to further improve the accuracy and utility of race and ethnicity data.

RECOMMENDATIONS

1. With input from social scientists, geneticists, and other experts, OMB, HHS, and NIH should draft a consensus statement defining race, ethnicity, and ancestry. This statement should make the differences between the concepts clear and outline the type of knowledge that can be gained from studies employing these variables.
2. HHS should disseminate best practices for asking respondents for race and ethnicity data, including guidance on how to address respondents’ concerns about the uses of the data. Additionally, it would be helpful for HHS to encourage organizations to provide formal training to individuals who collect these data, including researchers, funeral directors, and clinical staff who register patients.

3. HHS may consider issuing guidance to researchers and organizations about common resources and methods to determine appropriate granular ethnicity categories for their settings. Alternatively, HHS may consider disseminating a standard list of granular ethnicity categories.
4. HHS should provide guidance on how multiracial data should be tabulated and analyzed.
5. A question for “socially assigned race” should be further developed and tested.
6. The Center for Medicare and Medicaid Services should verify the accuracy of current Medicare enrollees’ race and ethnicity data, which may have been imported from the Social Security Administration prior to the implementation of improved standards for data collection.
7. HHS should develop guidance indicating appropriate circumstances under which indirect means, such as surname and geocoding, can be used for ascertaining race and ethnicity of populations when directly collected data are not available.
8. HHS should require that electronic health technology software packages include fields for race, Hispanic/Latino origin, and granular ethnicity to obtain certification.
9. As these standards are extended into health care delivery, HHS should consider the risks and benefits of collecting and sharing race and ethnicity data, as race and ethnicity data are not covered by the Health Insurance Portability and Accountability Act (HIPAA).
10. As these data standards are extended into health delivery settings, HHS should require the analysis of health care quality metrics by race and ethnicity, and consider creating pay for performance incentives aimed at reducing racial and ethnic disparities.

REFERENCES

- Afshari, R., & Bhopal, R. S. (2010). Ethnicity has overtaken race in medical science: MEDLINE-based comparison of trends in the USA and the rest of the world, 1965-2005. *International Journal of Epidemiology*, 39(6), 1682-1683. doi:10.1093/ije/dyp382
- Agency for Healthcare Research and Quality. (2011). *2010 National Healthcare Quality & Disparities Reports*. Retrieved April 3, 2012, from <http://www.ahrq.gov/qual/qrd10.htm>
- Alberts, B. (2008). *Molecular Biology of the Cell* (5th ed.). New York: Garland Science.
- Amaro, H., & Zambrana, R. E. (2000). Criollo, mestizo, mulato, LatiNegro, indigena, white, or black? The U.S. Hispanic/Latino population and multiple responses in the 2000 Census. *American Journal of Public Health*, 90(11), 1724-1727.
- Azoulay, K. G. (2006). Reflections on race and the biologization of difference. *Patterns of Prejudice*, 40(4-5), 353-380. doi:10.1080/00313220601020098
- Bhopal, R. (2006). Race and ethnicity: Responsible use from epidemiological and public health perspectives. *The Journal of Law, Medicine & Ethics: A Journal of the American Society of Law, Medicine & Ethics*, 34(3), 500-507.
- Bibbins-Domingo, K., & Fernandez, A. (2007). BiDil for heart failure in black patients: Implications of the U.S. Food and Drug Administration approval. *Annals of Internal Medicine*, 146(1), 52-56.
- Bilheimer, L. T., & Klein, R. J. (2010). Data and measurement issues in the analysis of health disparities. *Health Services Research*, 45(5 Pt 2), 1489-1507. doi:10.1111/j.1475-6773.2010.01143.x; 10.1111/j.1475-6773.2010.01143.x
- Bolnick, D. A. (2008). Individual ancestry inference and the reification of race as a biological phenomenon. In B. A. Koenig, S. Soo-Jin Lee, & S. Richardson (Eds.), *Revisiting Race in a Genomic Age* (pp. 70-85). New Brunswick, N.J.: Rutgers University Press.
- Bonham, V. L., Warshauer-Baker, E., & Collins, F. S. (2005). Race and ethnicity in the genome era: The complexity of the constructs. *The American Psychologist*, 60(1), 9-15. doi:10.1037/0003-066X.60.1.9
- Bonilla-Silva, E. (1997). Rethinking racism: Toward a structural interpretation. *American Sociological Review*, 62(3), 465-480.
- Bonilla-Silva, E. (1999). The essential social fact of race. *American Sociological Review*, 64(6), 899-906.
- Burchard, E. G., Ziv, E., Coyle, N., et al. (2003). The importance of race and ethnic background in biomedical research and clinical practice. *New England Journal of Medicine*, 348(12), 1170-1175. doi:10.1056/NEJMs025007
- Burhansstipanov, L., & Satter, D. E. (2000). Office of Management and Budget racial categories and implications for American Indians and Alaska Natives. *American Journal of Public Health*, 90(11), 1720-1723.
- Campbell, M. E. (2007). Thinking outside the (black) box: Measuring black and multiracial identification on surveys. *Social Science Research*, 36(3), 921-944. doi:10.1016/j.ssresearch.2006.07.001
- Caulfield, T., Fullerton, S. M., Ali-Khan, S. E., et al. (2009). Race and ancestry in biomedical research: Exploring the challenges. *Genome Medicine*, 1(1), 8. doi:10.1186/gm8
- Cavalli-Sforza, L. L. (2007). Human evolution and its relevance for genetic epidemiology. *Annual Review of Genomics and Human Genetics*, 8, 1-15. doi:10.1146/annurev.genom.8.080706.092403
- Cornell, S. E., & Hartmann, D. (2007). *Ethnicity and Race: Making Identities in a Changing World* (2nd ed.). Thousand Oaks, Calif.: Pine Forge Press.
- De Melo-Martín, I. (2008). Genetic research and reduction of health disparities. *New Genetics & Society*, 27(1), 57-68. doi:10.1080/14636770701843667
- Dorsey, R., & Graham, G. (2011). New HHS data standards for race, ethnicity, sex, primary language, and disability status. *JAMA: The Journal of the American Medical Association*, 306(21), 2378-2379. doi:10.1001/jama.2011.1789
- Drake, S. C. (1987; 1990). *Black Folk Here and There: An Essay in History and Anthropology*. Los Angeles: Center for Afro-American Studies, University of California.

- Dupre, J. (2008). What genes are and why there are no genes for race. In B. A. Koenig, S. Soo-Jin Lee, & S. Richardson (Eds.), *Revisiting Race in a Genomic Age* (pp. 39-55). New Brunswick, N.J.: Rutgers University Press.
- Ellison, G. T., Kaufman, J. S., Head, R. F., Martin, P. A., & Kahn, J. D. (2008). Flaws in the U.S. Food and Drug Administration's rationale for supporting the development and approval of BiDil as a treatment for heart failure only in black patients. *The Journal of Law, Medicine & Ethics: A Journal of the American Society of Law, Medicine & Ethics*, 36(3), 449-457. doi:10.1111/j.1748-720X.2008.290.x
- Epstein, S. (2007). *Inclusion: The Politics of Difference in Medical Research*. Chicago: University of Chicago Press.
- Eze, E. C. (1997). *Race and the Enlightenment: A Reader*. Cambridge, Mass.: Blackwell.
- Feldman, M. W., & Lewontin, R. C. (2008). Race, ancestry, and medicine. In B. A. Koenig, S. Soo-Jin Lee, & S. Richardson (Eds.), *Revisiting Race in a Genomic Age* (pp. 89-101). New Brunswick, N.J.: Rutgers University Press.
- Food and Drug Administration. (2005). *Guidance for Industry—Collection of Race and Ethnicity Data in Clinical Trials*. Retrieved April 4, 2012 from [http://www.fda.gov/RegulatoryInformation/Guidances/ucm126340.htm?utm_source=fdaSearch&utm_medium=website&utm_term=guidelines race&utm_content=1](http://www.fda.gov/RegulatoryInformation/Guidances/ucm126340.htm?utm_source=fdaSearch&utm_medium=website&utm_term=guidelines%20race&utm_content=1)
- Ford, C. L., & Harawa, N. T. (2010). A new conceptualization of ethnicity for social epidemiologic, and health equity research. *Social Science & Medicine*, 71(2), 251-258. doi:10.1016/j.socscimed.2010.04.008
- Frank, R. (2007). What to make of it? The (re)emergence of a biological conceptualization of race in health disparities research. *Social Science & Medicine*, 64(10), 1977-1983. doi:10.1016/j.socscimed.2007.01.010
- Fujimura, J. H., & Rajagopalan, R. (2011). Different differences: The use of 'genetic ancestry' versus race in biomedical human genetic research. *Social Studies of Science*, 41(1), 5-30.
- Fullwiley, D. (2007a). The molecularization of race: Institutionalizing human difference in pharmacogenetics practice. *Science as Culture*, 16(1), 1-30. doi:10.1080/09505430601180847
- Fullwiley, D. (2007b). Race and genetics: Attempts to define the relationship. *BioSocieties*, 2(2), 221-237. doi:10.1017/S1745855207005625
- Goodman, A. H. (2000). Why genes don't count (for racial differences in health). *American Journal of Public Health*, 90(11), 1699-1702.
- Gordon, M. M. (1964). *Assimilation in American Life: The Role of Race, Religion, and National Origins*. New York: Oxford University Press.
- Hahn, R. A., & Stroup, D. F. (1994). Race and ethnicity in public health surveillance: Criteria for the scientific use of social categories. *Public Health Reports*, 109(1), 7-15.
- Hitlin, S., Brown, J. S., & Elder, G. H., Jr. (2007). Measuring Latinos: Racial vs. ethnic classification and self-understandings. *Social Forces*, 86(2), 587-611.
- Humes, K. R., Jones, N. A., & Ramirez, R. R. (2011). *Overview of Race and Hispanic Origin: 2010*. U.S. Census Bureau. Retrieved April 3, 2012, from <http://www.census.gov/prod/cen2010/briefs/c2010br-02.pdf>;
- Humes, K., & Hogan, H. (2009). Measurement of race and ethnicity in a changing, multicultural America. *Race and Social Problems*, 1(3), 111-131. doi:10.1007/s12552-009-9011-5
- Hunt, L. M., & Megyesi, M. S. (2008). The ambiguous meanings of the racial/ethnic categories routinely used in human genetics research. *Social Science & Medicine*, 66(2), 349-361. doi:10.1016/j.socscimed.2007.08.034
- Institute of Medicine, Committee on Understanding and Eliminating Racial and Ethnic Disparities in Health Care. (2003). *Unequal treatment: Confronting racial and ethnic disparities in health care*. Smedley, B. D., Stith, A.Y., & Nelson, A.R. (Eds.). Washington, DC: National Academies Press.
- Jones, C. P., Truman, B. I., Elam-Evans, L. D., et al. (2008). Using "socially assigned race" to probe white advantages in health status. *Ethnicity & Disease*, 18(4), 496-504.
- Kahn, J. (2008). Exploiting race in drug development: BiDil's interim model of pharmacogenomics. *Social Studies of Science*, 38(5), 737-758.

- Kahn, J. (2006). Genes, race, and population: Avoiding a collision of categories. *American Journal of Public Health*, 96(11), 1965-1970. doi:10.2105/AJPH.2005.067926
- Kaufman, J. S. (2008). Epidemiologic analysis of racial/ethnic disparities: Some fundamental issues and a cautionary example. *Social Science & Medicine*, 66(8), 1659-1669. doi:10.1016/j.socscimed.2007.11.046
- Kittles, R. A., & Weiss, K. M. (2003). Race, ancestry, and genes: Implications for defining disease risk. *Annual Review of Genomics and Human Genetics*, 4, 33-67. doi:10.1146/annurev.genom.4.070802.110356
- Koenig, B. A., Lee, S. S., & Richardson, S. S. (Eds.). (2008). *Revisiting Race in a Genomic Age*. New Brunswick, N.J.: Rutgers University Press.
- Kolchin, P. (1993). *American Slavery, 1619-1877*. New York: Hill and Wang.
- Krieger, N. (2005). Stormy weather: Race, gene expression, and the science of health disparities. *American Journal of Public Health*, 95(12), 2155-2160. doi:10.2105/AJPH.2005.067108
- Lee, C. (2009). "Race" and "ethnicity" in biomedical research: How do scientists construct and explain differences in health? *Social Science & Medicine*, 68(6), 1183-1190. doi:10.1016/j.socscimed.2008.12.036
- Mays, V. M., Ponce, N. A., Washington, D. L., & Cochran, S. D. (2003). Classification of race and ethnicity: Implications for public health. *Annual Review of Public Health*, 24, 83-110. doi:10.1146/annurev.publhealth.24.100901.140927
- McKenney, N. R., & Bennett, C. E. (1994). Issues regarding data on race and ethnicity: The Census Bureau experience. *Public Health Reports*, 109(1), 16-25.
- National Human Genome Research Institute. (2012). *International HapMap Project*. Retrieved April 3, 2012, from <http://www.genome.gov/10001688>
- National Institutes of Health. (2001). *NIH guidelines on the inclusion of women and minorities as subjects in clinical research - Amended, October, 2001*. Retrieved April 3, 2012, from http://grants.nih.gov/grants/funding/women_min/guidelines_amended_10_2001.htm
- Nobles, M. (2000). History counts: A comparative analysis of racial/color categorization in U.S. and Brazilian censuses. *American Journal of Public Health*, 90(11), 1738-1745.
- Office of Management and Budget. (1977). *OMB Directive 15: Race and Ethnic Standards for Federal Statistics and Administrative Reporting*. Retrieved April 3, 2012, from <http://wonder.cdc.gov/wonder/help/populations/bridged-race/Directive15.html>
- Office of Management and Budget. (1997). *Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity*. Retrieved April 3, 2012, from http://www.whitehouse.gov/omb/fedreg_1997standards
- Office of Minority Health, U.S. Department of Health and Human Services. (2011). *Explanation of Data Standards for Race, Ethnicity, Sex, Primary Language, and Disability*. Retrieved April 3, 2012, from <http://minorityhealth.hhs.gov/templates/content.aspx?ID=9228&lvl=2&lvlID=208>
- Omi, M., & Winant, H. (1994). *Racial Formation in the United States: From the 1960s to the 1990s*. New York: Routledge.
- Oppenheimer, G. M. (2001). Paradigm lost: Race, ethnicity, and the search for a new population taxonomy. *American Journal of Public Health*, 91(7), 1049-1055.
- Parker, J. D., & Makuc, D. M. (2002). Methodologic implications of allocating multiple-race data to single-race categories. *Health Services Research*, 37(1), 203-215.
- Penner, A. M., & Saperstein, A. (2008). How social status shapes race. *Proceedings of the National Academy of Sciences of the United States of America*, 105(50), 19628-19630. doi:10.1073/pnas.0805762105
- Perlmann, J., & Waters, M. C. (2002). *The New Race Question: How the Census Counts Multiracial Individuals*. New York: Russell Sage Foundation.
- Phelan, J. C., Link, B. G., & Tehranifar, P. (2010). Social conditions as fundamental causes of health inequalities: Theory, evidence, and policy implications. *Journal of Health and Social Behavior*, 51 Suppl, S28-40. doi:10.1177/0022146510383498
- Risch, N., Burchard, E., Ziv, E., & Tang, H. (2002). Categorization of humans in biomedical research: Genes, race and disease. *Genome Biology*, 3(7), comment 2007.

- Robbin, A. (1999). The problematic status of U.S. statistics on race and ethnicity: An "imperfect representation of reality." *Journal of Government Information*, 26(5), 467-483.
- Roberts, D. (2011). *Fatal Invention: How Science, Politics, and Big Business Re-Created Race in the Twenty-First Century*. New York: New Press.
- Roth, W. D. (2010). Racial mismatch: The divergence between form and function in data for monitoring racial discrimination of Hispanics. *Social Science Quarterly*, 91(5), 1288-1311. doi:10.1111/j.1540-6237.2010.00732.x
- Rotimi, C. N., & Jorde, L. B. (2010). Ancestry and disease in the age of genomic medicine. *New England Journal of Medicine*, 363(16), 1551-1558. doi:10.1056/NEJMra0911564
- Royal, C. D., Novembre, J., Fullerton, S. M., Goldstein, D. B., Long, J. C., Bamshad, M. J., & Clark, A. G. (2010). Inferring genetic ancestry: Opportunities, challenges, and implications. *American Journal of Human Genetics*, 86(5), 661-673. doi:10.1016/j.ajhg.2010.03.011
- Sankar, P. (2003). MEDLINE definitions of race and ethnicity and their application to genetic research. *Nature Genetics*, 34(2), 119; discussion 120. doi:10.1038/ng0603-119
- Shields, A. E., Fortun, M., Hammonds, E. M., King, P. A., Lerman, C., Rapp, R., & Sullivan, P. F. (2005). The use of race variables in genetic studies of complex traits and the goal of reducing health disparities: A transdisciplinary perspective. *American Psychologist*, 60(1), 77-103. doi:10.1037/0003-066X.60.1.77
- Smedley, A. (1998). "Race" and the construction of human identity. *American Anthropologist*, 100(3), 690-702.
- Smedley, A., & Smedley, B. D. (2005). Race as biology is fiction, racism as a social problem is real: Anthropological and historical perspectives on the social construction of race. *American Psychologist*, 60(1), 16-26. doi:10.1037/0003-066X.60.1.16
- Smedley, A., & Smedley, B. (2011). *Race in North America*. New York: Westview Press.
- Snipp, C. M. (2003). Racial measurement in the American Census: Past practices and implications for the future. *Annual Review of Sociology*, 29, 563-588. doi:10.1146/annurev.soc.29.010202.100006
- Soo-Jin Lee, S. (2008). Racial realism and the discourse of responsibility for health disparities in a genomic age. In B. A. Koenig, S. Soo-Jin Lee, & S. Richardson (Eds.), *Revisiting Race in a Genomic Age* (pp. 342-358). New Brunswick, N.J.: Rutgers University Press.
- Temple, R., & Stockbridge, N. L. (2007). BiDil for heart failure in black patients: The U.S. Food and Drug Administration perspective. *Annals of Internal Medicine*, 146(1), 57-62.
- Tucker, C., Miller, S., & Parker, J. (2002). Comparing Census race data under the old and new standards. In J. Perlmann, & M. C. Waters (Eds.), *The New Race Question: How the Census Counts Multiracial Individuals* (pp. 365-390). New York: Russell Sage Foundation.
- U.S. Census Bureau. (2002). *Measuring America: The Decennial Censuses from 1790 to 2000*. Retrieved April 3, 2012, from <http://www.census.gov/prod/www/abs/ma.html>
- U.S. Congress. (1993). *NIH Revitalization Act of 1993, Public Law 103-43*. Retrieved April 3, 2012, from <http://orwh.od.nih.gov/inclusion/revitalization.pdf>
- U.S. Congress. (1997). *Food and Drug Administration Modernization Act of 1997, Public Law 105-115*. Retrieved 4/4/2012, 2012, from <http://www.fda.gov/RegulatoryInformation/Legislation/FederalFoodDrugandCosmeticActFDCAAct/SignificantAmendmentstotheFDCAAct/FDAMA/FullTextofFDAMAlaw/default.htm>
- U.S. Congress. (1999). *Healthcare Research and Quality Act of 1999*. Retrieved April 12, 2012, from <http://www.ahrq.gov/hrqa99a.htm>
- U.S. Congress. (2010). *Patient Protection and Affordable Care Act. Understanding Health Disparities: Data Collection and Analysis. Public Law 111-148, Sec. 4302*. Retrieved April 3, 2012, from <http://www.gpo.gov/fdsys/pkg/PLAW-111publ148/pdf/PLAW-111publ148.pdf>
- U.S. Department of Health and Human Services. (1999). *Improving the Collection and Use of Racial and Ethnic Data in HHS*. Retrieved April 4, 2012, from <http://aspe.hhs.gov/datacncl/racerpt/index.htm>
- U.S. Public Health Service. (1992). *Improving Minority Health Statistics: Report of the PHS Task Force on Minority Health*.

Ulmer, C., McFadden, B., & Nerenz, D. R. (Eds.). Subcommittee on Standardized Collection of Race/Ethnicity Data for Healthcare Quality Improvement. (2009). *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement*. Washington, D.C.: National Academies Press.

Ver Ploeg, M., & Perrin, E. (Eds.). Panel on DHHS Collection of Race and Ethnic Data. (2004). *Eliminating Health Disparities: Measurement and Data Needs*. Washington, D.C.: National Academies Press.

Wallman, K. K., Evinger, S., & Schechter, S. (2000). Measuring our nation's diversity: Developing a common language for data on race/ethnicity. *American Journal of Public Health*, 90(11), 1704-1708.

Waters, M. C. (1990). *Ethnic Options: Choosing Identities in America*. Berkeley: University of California Press.

Waters, M. C. (1999). *Black Identities: West Indian Immigrant Dreams and American Realities*. New York: Russell Sage Foundation.

Williams, D. R. (1999). The monitoring of racial/ethnic status in the USA: Data quality issues. *Ethnicity & Health*, 4(3), 121-137. doi:10.1080/13557859998092

Wimmer, A. (2008). The making and unmaking of ethnic boundaries: A multilevel process theory. *American Journal of Sociology*, 113(4), 970-1022.

Woo, M., Austin, S. B., Williams, D. R., & Bennett, G. G. (2011). Reconceptualizing the measurement of multiracial status for health research in the United States. *Du Bois Review: Social Science Research on Race*, 8(1), 25. doi:10.1017/S1742058X11000038

Zuberi, T. (2001). *Thicker Than Blood: How Racial Statistics Lie*. Minneapolis: University of Minnesota Press.

ABOUT THE AUTHOR

Brooke Cunningham, M.D., Ph.D., is a fellow in General Internal Medicine and a Greenwall Fellow at the Berman Institute of Bioethics at the Johns Hopkins University. Dr. Cunningham majored in both history and African American and African Studies as an undergraduate at the University of Virginia. She earned her medical degree and her doctorate in sociology at the University of Pennsylvania, prior to entering residency in internal medicine at Duke University Medical Center. She began exploring how physicians think about difference in her dissertation, which examined the controversy surrounding international trials to prevent the transmission of HIV from mother to child. Her current research examines physicians' use of race in medical decision-making and seeks to describe factors that shape how health care organizations address health care disparities. As part of the Johns Hopkins Medicine Health Equity Alliance, she currently is working to improve the collection of race and ethnicity data across the Hopkins entities. It was through her work with the Joint Center as a Greenwall Fellow that she first began examining in detail the challenges associated with race and ethnicity data collection.

ABOUT THE JOINT CENTER FOR POLITICAL AND ECONOMIC STUDIES

The Joint Center for Political and Economic Studies is one of the nation's leading research and public policy institutions and the only one whose work focuses primarily on issues of particular concern to African Americans and other people of color. For more than four decades, our research and information programs have informed and influenced public opinion and national policy to benefit not only African Americans but every American.



BOARD OF GOVERNORS
June 2012

Cynthia G. Marshall, *Chair*
President
AT&T North Carolina

Earl W. Stafford, *Secretary*
Chief Executive Officer
The Wentworth Group, LLC

Dwight L. Bush, *Vice Chair*
Managing Director
D.L. Bush & Associates

Robert R. Hagans, Jr., *Treasurer*
Executive Vice President
and Chief Financial Officer
AARP

Barbara L. Johnson, Esq., *At Large*
Partner
Paul Hastings LLP

Ralph B. Everett, Esq.
President and Chief Executive Officer

Lawrence D. Bobo, Ph.D.
W.E.B. DuBois Professor
of the Social Sciences
Department of Sociology
Harvard University

Marva Smalls
Executive Vice President of Global
Inclusion Strategy, MTV Networks
& Executive Vice President
of Public Affairs, and Chief of Staff
Nickelodeon/MTVN Kids & Family Group

Donna Brazile
Founder and Managing Director
Brazile & Associates L.L.C.

Reed V. Tuckson, M.D., FACP
Executive Vice President and Chief
of Medical Affairs
UnitedHealth Group

The Honorable Carolyn C. Kilpatrick
Managing Partner
J & A Global Partners, L.L.C.

The Honorable J.C. Watts, Jr.
Chairman
J.C. Watts Companies

Freeda Kapor Klein, Ph.D.
Trustee
Mitchell Kapor Foundation

Robert L. Wright, O.D.
Chairman and CEO
FE Holdings, Inc.

Robert L. Mallett, Esq.
The Peter P. Mullen Visiting
Professor of Law
Georgetown University Law Center

Cynthia M. Bodrick
Assistant Secretary of the Corporation

MEMBERS EMERITI

Dr. William B. Boyd
President Emeritus
The Johnson Foundation

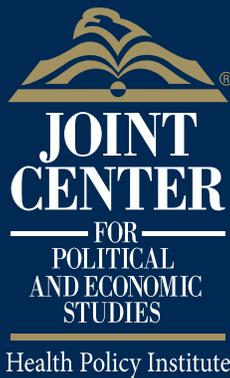
James D. Wolfensohn
President & CEO
Wolfensohn and Company

Eddie N. Williams
President
Eddie Williams and
and Associates, LLC

FOUNDERS

Dr. Kenneth B. Clark
Served from 1970 to 2005

Louis E. Martin
Served from 1970 to 1997



Joint Center for Political and Economic Studies
1090 Vermont Avenue, NW, Suite 1100
Washington, DC 20005
www.jointcenter.org



This product is printed on paper that is 50% recycled, 30% post consumer, Elemental Chlorine Free (ECF), and acid free.

