Race/Ethnicity and the 2000 Census: Implications for Public Health

Edward J. Sondik, PhD, Jacqueline Wilson Lucas, MPH, Jennifer H. Madans, PhD, and Sandra Surber Smith, MPH

ABSTRACT

Objectives. This article addresses the potential impact of the revised standards for race and ethnicity on data from the 2000 census and public health data sources, policies, and programs.

Methods. The authors examine the relationship between race/ethnicity and health in selected measures, identify the factors that influence race/ethnicity identification, consider past experience in race/ethnicity reporting, and explore the challenges in understanding and managing the effects of new racial/ethnic categories in various data sets.

Results. The multiple-race group seems to compose only a small percentage of the US population and may have little impact on data for single-race groups. Actual effects will vary according to a number of factors, including the size, composition, and geographic distribution of the group.

Conclusions. More research is needed to support a thorough understanding of the reporting of multirace data and the development of techniques for analyzing these data. Given the importance of understanding the relationship between race/ethnicity and health, the ability to produce useful, comparable, and meaningful data is essential. (*Am J Public Health.* 2000;90:1709–1713)

The 2000 census has adhered to the 1997 revised US Office of Management and Budget (OMB) standards for race and ethnicity, and other government agencies have followed or will soon follow suit. In this article, we explore what this change will mean to health data and health research and what we must do to understand and manage the impact.

Importance of Race and Ethnicity in Health

We start from the fact that race and ethnicity are critical to many aspects of life in the United States, in some ways that we support and in others we reject. Certainly race and ethnicity are important in public health. Whether representing actual differences or a constellation of factors that affect health and health status, race and ethnicity are important determinants of health patterns in the United States. In studies that control for income, education, and other measures of socioeconomic status, differences by race are often reduced, but other times remain, for many of the key health indicators we track.

For example, the prevalence of HIV infection among Blacks is 8 times that among Whites and 2 times that among Hispanics; the rate of annual Papanicolaou tests is one third higher among Blacks than among Asians and Pacific Islanders; and diabetes-related amputations are 25% higher among Blacks than among Whites.[1] These 3 topics are part of a major Department of Health and Human Services initiative to eliminate health disparities that also focuses on disparities in cardiovascular disease outcomes, infant mortality, and adult and childhood immunizations. To improve health and health care, we must recognize and be able to accurately measure and monitor demographic shifts and population changes. We must understand these dynamics not just as independent phenomena but also in relation to health outcomes.

Importance of Trends

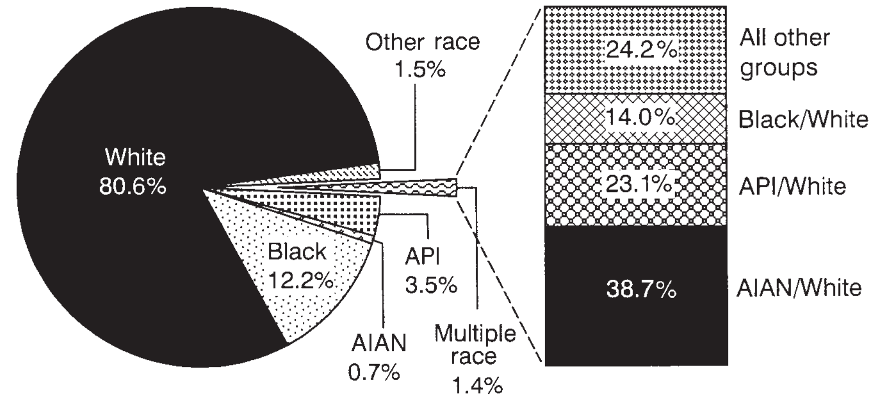
Trend data shed light on the process of change. Perhaps even more important, they signal emerging problems or improvements in health status. And perhaps most important, trend data are essential in evaluating the effects of interventions. While changes in reporting are not new, we always work to minimize the effects of those changes. It is important that we set standards and control processes to allow us to collect information that reflects only actual change rather than the accidental, the inadvertent, or the coincidental.

We have been collecting data that for a given individual might change as follows. Suppose individual *X* reports under the new guidelines that her race is Black and Asian. Under the former (single response) guidelines, she could have reported, we hypothesize, either Black or Asian or perhaps another race, although the last-mentioned possibility seems unlikely. In any case, trends tracked as Asian or Black will change to some degree under the new guidelines. In effect, the guidelines force us to define a new set of independent variables that consist of the new race categories and their combinations. Because race and ethnicity are reported separately, there are actually 2 sets of race categories, one for Hispanics and one for non-Hispanics. The discontinuity in trends may be significant, especially in some sections of the country.

The authors are with the National Center for Health Statistics, Centers for Disease Control and Prevention, Hyattsville, Md.

Requests for reprints should be sent to Edward J. Sondik, PhD, National Center for Health Statistics, Centers for Disease Control and Prevention, 6525 Belcrest Rd, Room 1140, Hyattsville, MD 20782 (e-mail: efs2@cdc.gov).

This article was accepted June 26, 2000.



Source. Centers for Disease Control and Prevention/National Center for Health Statistics
Note. API = Asian/Pacific Islander; AIAN = American Indian/Alaska Native.

FIGURE 1—Distribution of single- and multiple-race groups in the 1997 National Health Interview Survey.

TABLE 1—Primary Racial Identifications Selected by Multiple-Race Groups: National Health Interview Survey, 1997

	Primary Racial Identification, %				
	White	Black	AI/AN	API	Multiple Race
AI/AN/White	81.6	…	16.5	…	1.8
API/White	39.1	…	…	48.7	12.7
Black/White	26.7	50.2	…	…	23.2

Note. Data are age adjusted. AI/AN = American Indian/Alaska Native; API = Asian/Pacific Islander.

Factors That Influence the Reporting of Race

Discontinuity will be no less apparent as we implement changes in race and ethnicity classifications. The 2000 census is not the first time we have changed race and ethnicity classifications, but it is the first time that there will be systemwide reporting of multiple races by individuals. The percentage of the population that will report itself as multiracial has been estimated at less than 2% in the United States. This estimate is based on the reporting of multiple races in the National Center for Health Statistics' National Health Interview Survey (NHIS) (Lucas JW, unpublished NHIS data, 1999).

But there are many unknowns. We do know that interracial births are increasing.[2] Will the number of those who report multiracial identity grow? We do not know enough about what factors affect the selection of a single racial or ethnic identity, let alone any combination of races or ethnicities. We believe that there are 3 principal factors that influence reporting, however. First, the choices offered in the OMB guidelines, and the respondent's understanding of those choices, affect the way race and ethnicity are reported. At least some of the criticism of the race and ethnicity choices on the 2000 census (as reported in the *Washington Post* and the *New York Times*) seems to be related to the difficulties encountered by respondents in understanding the categories both in general and in terms of their own background.

Second, fashion or custom may play a role. Today, race is very much a combination of biologic/genetic, cultural, and societal constructs. Interpretation of race and identification of race increasingly are subject to outside and changing influences. Identification of race and ethnicity may certainly—and legitimately—vary over time as experiences and environments change.

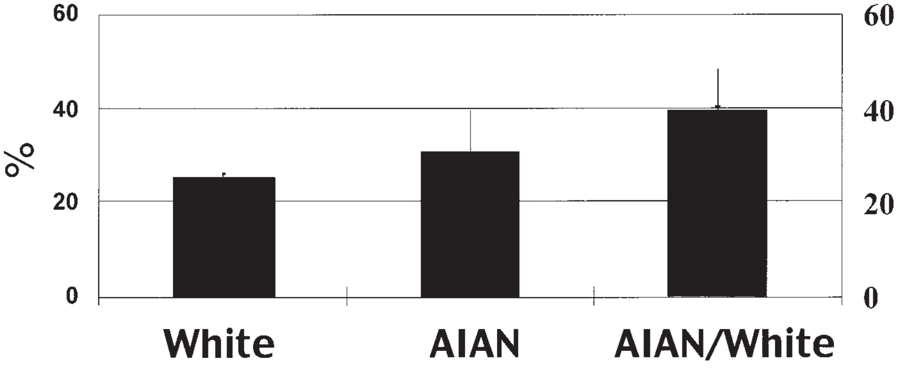
Third, politics must be considered. As long as race plays a role in how individuals are viewed and treated, race may be differentially reported. In addition, the reporting of race and ethnicity is far from just an individual selection. There are groups advocating that multiracial individuals select one race or another to strengthen the political position of that group. Others advocate the identification of multiple races to enhance or strengthen the economic and political clout of the multiracial community, and still others hope that a multiracial identification will lead to a "raceless society."[3] Those of us collecting and analyzing data do not as yet know the impact of all of these forces, but they must be taken into account.

Past Experience: Using Data From the National Health Interview Survey

We have some data on multiple-race reporting in household surveys that may be useful as we discuss this issue. For more than 20 years, NHIS respondents have been asked to select all of the race groups that represent them. As a follow-up question, they have been asked which race best represents them; that is, they have been asked to provide what might be called a primary racial identification. The questions are stated as follows: "What is the number [from a printed list] of the group or groups that represent your race?" and "Which of those groups would you say best describes your race?"

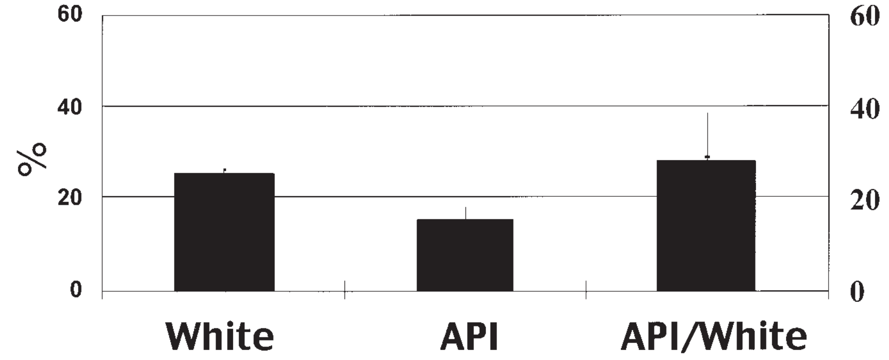
In 1997, about 1.4% of the population selected more than one race or identified themselves as multiracial. Of this group, 39% were American Indian/Alaska Native/White, 23% were Asian or Pacific Islander/White, 14% were Black/White, and 24% were other combinations (see Figure 1). When asked to select a primary racial identification (Table 1), more than 80% of the American Indian/Alaska Native/White group selected White; about half of the Asian or Pacific Islander/White group selected Asian or Pacific Islander; and just over half of the Black/White group selected Black. However, almost a quarter of the Black/White respondents declined to indicate a primary race. These distributions have been relatively stable over the past 20 years (Lucas JW, unpublished NHIS data, 1999).

What do these data tell us about the multiracial population? The multiracial group is composed of individuals of varied racial combinations who identify in varying proportions with a primary race. As a group, they may be the progeny of interracial marriages, children of multiracial parents, or descendants of a multiracial heritage that goes back generations. There may be no common experiences for the multiracial group, except that all live in a predominantly monoracial society. Some say that experience is enough to create a bond, a commonality.[3] If so, then in a survey we might expect different responses from a multiple-race group compared with the responses of each of its component groups. The following 3 examples show the variety of relationships that may exist between a multiple-race group and its component single-race groups.



Source. Centers for Disease Control and Prevention/National Center for Health Statistics
Note. Data are age-adjusted. Vertical lines indicate 95% confidence intervals.

FIGURE 2—Current smoking among respondents 18 years or older for 2 single- and 1 multiple-race group: National Health Interview Survey, 1997.



Source. Centers for Disease Control and Prevention/National Center for Health Statistics
Note. Data are age-adjusted. Vertical lines indicate 95% confidence intervals.

FIGURE 3—Current smoking among respondents 18 years or older for 2 single- and 1 multiple-race group: National Health Interview Survey, 1997.

The 1997 NHIS asked respondents whether they currently were smokers. The responses for Whites, American Indians/Alaska Natives, and Asian/Pacific Islanders are shown in Figures 2 and 3. The rate of current smoking in the American Indian/Alaska Native/White group was higher than the response from the White group, although the confidence intervals show that the response from that group was consistent with the response from the American Indian/Alaska Native group. On the other hand, the response from the multiple-race group Asian or Pacific Islander/White group was higher than the Asian or Pacific Islander response but consistent with the White response.[4]

In still another example, rates of private health insurance coverage, as measured by the NHIS, have been shown to be about 59% for the American Indian/Alaska Native/White group, about 75% for Whites, and about 40% for American Indians/Alaska Natives. Here, the rate for the multiple-race group seems to be consistent with an average of the rates for the single-race groups (Lucas JW, unpublished NHIS data, 1999).

Impact of Multiple-Race Categories

Although the multiple-race group constitutes only a small percentage of the US population, its impact will vary significantly according to a number of factors.[5] For example, in regard to size and composition of biracial groups, smaller single-race groups such as American Indians/Alaska Natives, Asians, and Native Hawaiians/other Pacific Islanders stand a greater chance than larger racial groups (White, Black) to be affected by the use of multiple-race categories. In fact, NHIS data show that the American Indian/Alaska Native/White group is larger than the American Indian/Alaska Native single-race group.

In terms of geographic distribution of multiple–race/ethnicity groups, the concentration of groups in certain states or communities may differentially affect data for those areas. The diversity of some communities may dictate multiracial analyses; for example, the diversity present in California is not present in Iowa.

Techniques for Analyzing Multiple-Race Data

To maintain trends in various health data sets, we will need to develop and use mechanisms for bridging the differences between the current race and ethnicity categories and those under the 1977 standards. As mentioned earlier, the NHIS offers a unique opportunity to study multirace data and develop bridging mechanisms. The impact of bridge methods on estimates for single-race groups will depend on several factors, including (1) the extent to which single- and multiple-race groups differ on characteristics that affect the outcomes of interest, (2) the size of the difference between single- and multiple-race respondents on the outcomes of interest, (3) the relative size of the groups, and (4) how multiple-race respondents answered race questions under the old standard. More research is necessary and, indeed, is under way in the federal statistics agencies to determine the "best" bridge methods for specific applications, data sets, or situations.

Multiple-Race Data and Health Disparities

Healthy People 2010

As noted earlier, data gathered by race and ethnicity are used in public health to set goals, establish programs and policies, and measure progress. Nowhere is this demonstrated more effectively than in the national health promotion and disease prevention objectives set forth in *Healthy People 2010: Understanding and Improving Health*.[6] One of the overall goals of Healthy People 2010 is the elimination of disparities in health. The revised race and ethnicity classifications will have an impact on the Healthy People initiative. Unlike the previous versions of the initiative (Healthy People and Healthy People 2000), Healthy People 2010 has set the same target for all population groups under each objective. Target setting was based on the rates achieved by the racial group exhibiting the healthiest behavior. These targets can be revisited mid-decade, and thus comparability among groups and trend data is critical.

The Healthy People process is a partnership with state and local governments and private-sector institutions, organizations, and groups. To the extent that these organizations feel the impact of the classifications, Healthy People will reflect that as well. In community data generated as part of the Healthy People process, there may be more day-to-day issues revolving around the multiple-race categories than we observe at the national level.

A problem we have yet to address is that targets are set with the baseline data—all on the 1977 standards—but progress must be tracked according to the new standards. Clearly, we need a strong bridge to the past.

REACH Program

The Centers for Disease Control and Prevention's REACH funds programs in 30 communities to eliminate disparities in 6 areas of health: infant mortality, cancer, heart disease, immunizations, diabetes, and AIDS. Data measuring progress in each of these areas may be affected by the new classifications. For example, mortality data are used to monitor 5 of these priorities, and birth data are used to monitor infant mortality rates. Indeed, birth data may be particularly sensitive to the classification changes.

Not all registration areas will implement the 1997 standards at the same time or with complete coverage and compliance at the start. For example, California alone has implemented the revised standards for birth and death records filed in the year 2000; this was done to provide compatibility with the 2000 census. Other states may implement the revised standards for race in 2001. In 2003, it is anticipated that all states will implement the revised standards in conjunction with the implementation of completely revised birth and death records. As a result, for some years, race data will be derived in different states under different concepts, and thus, to create national totals by race, decisions must be made about how to aggregate disparate vital statistics data.

In constructing postcensal population estimates for the nation, states, counties, and cities, the Census Bureau uses race data in annual birth and death data files provided by the National Center for Health Statistics. Production of accurate population estimates by race and other characteristics (age, sex, Hispanic origin) depends on availability of comparable data across the various data sources. Postcensal population estimates will be created by updating the 2000 census with birth, death, and migration information, and comparability of these components with the census is critical if estimates are to be accurate.

Vital statistics data include birth and death rates for various population groups. Typically, the numerator is derived from birth or death record counts, while the denominator comes from postcensal population estimates. Birth certificate data on race are likely to have been self-reported by the mother, and over time these data may become comparable to census data collected under the new standards. Death certificate data, however, are usually provided by an observer, such as the next of kin, and sometimes by a funeral director or coroner. These data, particularly for populations with multiple racial heritages, are likely to be quite different from the information obtained when respondents report about themselves.

Also, people may report differently about themselves over time, depending on age, marriage, children, migration, or perceived benefit of reporting one way over another. Thus, we have the problem of creating valid birth and death rates by race with comparable concepts of race in the numerator and denominator. We have taken some time to outline these specific challenges, because they are real and affect one of the most important public health and population data systems.

Health Insurance Portability and Accountability Act

Certainly, the new standards present different challenges to different data systems. It is possible that some of the most important public health databases will be those generated under the Health Insurance Portability and Accountability Act of 1996. The Department of Health and Human Services has been successful in the first step toward enabling the collection of racial and ethnic data on the act's institutional claim standard, allowing the revised classifications to be used on health care forms. Collecting this type of information in an administrative and financial environment maximizes the opportunities to produce relevant health care statistics for public health purposes, such as addressing disparities in access to and quality of care.

Research Agenda

One of the ongoing research activities at the National Center for Health Statistics focuses on the processing and analysis of multiple-race reporting. However, it is very clear that we will need more research in several different directions if we are to increase knowledge of and manage the impact of the revised standards. We must support the following:

- *Cognitive testing of race and ethnicity questions with multiple-race respondents*. How and why do people select certain racial or ethnic categories? What factors may cause them to vary their responses? How does selection change over time?
- *Bridge methodology*. What are the best ways to bridge the data collected under the new categories with those collected under the old categories? What factors favor one method over another?
- *Modeling of relationships between risk factors and health outcomes*. Are the paths from risk behavior to disease present for single- as well as multiple-race groups? Are the factors affecting health care choice, opportunity, and use the same for both groups?
- *Interaction with other methodological changes*. How do the revised race standards interact with other methodological changes in the nation's data systems, such as implementation of the *International Classification of Diseases, 10th Revision*; new birth and death standard certificates; and the new population standard for age adjustment? Simply in regard to the last-mentioned of these changes, we know that the 2000 population standard—which replaces the 1940 standard and which the National Center for Health Statistics will begin to use with 1999 data—reduces racial disparities. How does the fact that multirace populations are generally younger than single-race groups complicate the analysis?
- *Technologic change*. Finally, what is the impact of technologic change? Electronic birth and death records, computerized medical records, and all the forms of computer-assisted data collection will certainly have some effect that is not yet totally understood in its magnitude or direction.

As we began by saying, data gathered by race and ethnicity are important for understanding and improving the public's health. Our data systems must be up to the task. ☐

Contributors

E. J. Sondik was the lead author and conceived the approach, scope, and direction of the paper. J. W. Lucas analyzed the data and contributed to the conceptual and methodological framework of the paper. J. H. Madans addressed policy issues and programmatic implications. S. S. Smith wrote the manuscript with significant contributions, revision, and review from the other authors.

Acknowledgments

We would like to thank the following individuals for their invaluable advice and information: Jennifer D.

Parker, PhD, senior staff fellow, Office of Analysis, Epidemiology and Health Promotion, National Center for Health Statistics, and James A. Weed, PhD, deputy director, Division of Vital Statistics, National Center for Health Statistics.

References

1. *Healthy People 2000 Review, 1998–99*. Hyattsville, Md: National Center for Health Statistics; 1999.
2. Vital statistics of the United States, 1998, technical appendix. Available at: http://www.cdc.gov/nchs/data/98natali.pdf.
3. Root MPP. *The Multiracial Experience: Racial Border as the New Frontier*. Thousand Oaks, Calif: Sage Publications; 1996.
4. Lucas, JW. OMB's new classification of race and ethnicity and its potential impact on health research. Paper presented at: Scientific Seminar Series on NCHS Data, sponsored by the Office of the Associate Director for Science, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention; Atlanta, Georgia; March 10, 2000.
5. Recommendations from the Interagency Committee for the Review of Racial and Ethnic Standards to the Office of Management and Budget concerning changes to the standards for the classification of federal data on race and ethnicity. *Federal Register*. July 9, 1997: 36873–36946.
6. *Healthy People 2010: Understanding and Improving Health*. Washington, DC: US Dept of Health and Human Services; 2000.