*Trude Bennett, DrPH*

# "Racial" and Ethnic Classification: Two Steps Forward and One Step Back?

Twenty years ago, in the wake of civil rights legislation designed to combat the legacy of racial and ethnic discrimination, the Office of Management and Budget (OMB) established standards "to provide consistent and comparable data on race and ethnicity throughout the Federal government for an array of statistical and administrative programs."[1]

The most far-reaching of the recommendations in OMB's Statistical Policy Directive No. 15 was the designation of four "racial" groups—American Indian or Alaskan Native, Asian or Pacific Islander, black, and white—and one "ethnic" category, Hispanic origin. Directive No. 15 stated that the categories that most closely reflect individuals' recognition in their communities should be used.[1] When practical, self-identification was designated as the most desirable method of collecting data on "race" and ethnicity.

Growing controversy over the appropriateness and adequacy of the OMB standards led to a four-year Federal review process, the outcome of which coincided with President Clinton's recent call for a "national conversation on race." In what follows I give an overview of the final recommendations of the OMB review process, the contradictions revealed in the definitions and uses of "racial" and ethnic categories, and potential problems in the implementation of new guidelines.

## Data Issues

OMB stated explicitly from the outset that its 1977 standards were not based on anthropological or scientific principles. Neither were the standards meant to determine individual eligibility for Federal programs but rather to facilitate consistent monitoring of population demographics and differential treatment in education, employment, housing, lending practices, the legal system, and health and other public services.[2] Yet Federal funding allocations, Congressional districts, and enforcement of equal access provisions have been influenced by the OMB classification system. OMB's "racial" and ethnic categories have also had far-reaching effects on public health research and practice through their impact on the collection, tabulation, and reporting of vital statistics data (frequently used as numerators for health indicators) and Census data (the primary source of population-based denominators).

Since OMB Directive No. 15 was issued in 1977, scientific and political controversies have arisen about the social and statistical validity of "racial" and ethnic data. Debates have ensued about the potential for reinforcing stereotypes, reproducing inequalities, and obscuring true causal relationships through the use of "racial" and ethnic classifications in health research and surveillance.[3,4] Growing recognition of race as a socially constructed, non-biologically based category has cast doubt on the utility of racial analyses.

Critiques of health research that uses "race" and ethnicity as variables stress the need to explore complex interactions between race as defined in this society, ethnic origin, and socioeconomic status (SES) and to search for causal pathways that explain why "race" is often a marker of adverse health risks and outcomes.[5] Among the recommendations that resulted from a 1993 Workshop on the Use of Race and Ethnicity in Public Health Surveillance sponsored by the Centers for Disease Control and Prevention (CDC) and the Agency for Toxic Substances and Disease Registry (ATSDR) were that data on "race" and ethnicity should be collected only when useful for public health purposes; that data on "racial" and ethnic differences should be analyzed in relation to potential confounding variables such as SES; and that the justification and methods for measurement of "race" and ethnicity should be clearly defined and explicitly stated.[6]

Recent efforts to incorporate measures of social class into public health research and surveillance[7] aim to elucidate social inequalities and place "racial" and ethnic differences in a broader context (see article in this issue by Krieger et al., page 481, and associated commentary by Williams, page 492).

## Possible Solutions

The persistence of racial discrimination, however, necessitates improved methods of identifying and enumerating populations for purposes of equal opportunity and legal redress. Hahn and colleagues[8,9] have demonstrated serious inconsistencies and systematic biases in major sources of "racial" and ethnic data. Though self-identification changes over time for many individuals, more significant problems result from assignment of "race" and ethnicity by observers such as hospital admission clerks or funeral home directors. Hahn et al. con-

cluded that public health research and planning may be hindered by subsequent errors such as the underestimation of infant mortality for populations other than blacks and whites. Some steps have been taken to rectify these problems—the linkage of birth and infant death certificates by the National Center for Health Statistics, for example, promotes consistency and self-identification.

Other proposed remedies to deficiencies in "racial" and ethnic data have met with less success. Legal efforts to include items on "race" and ethnicity in the uniform hospital claims form (UB-92) ended in failure. Proponents feared that violations of Title VI of the Civil Rights Act (that is, segregation of patients and discriminatory treatment in hospitals receiving Federal funds) would continue undetected without such information. One indication of the difficulty in tracking "race" and ethnicity in hospital records is the extent to which these data are missing in the National Hospital Discharge Survey—in nearly 20% of all records in 1992.[10]

Another instance of broad-based advocacy was a lawsuit seeking the release of 1990 Census data adjusted for the undercount of "racial" and ethnic populations that was verified by Census Bureau studies. After prolonged litigation, the U.S. Supreme Court supported the Commerce Department's decision not to adjust data from the 1990 Census, which had missed an estimated 4.6% to 5.7% of "blacks," 5.0% of "Hispanics," and 12.2% of "reservation Indians" according to the Census Bureau's 1990 Post-Enumeration Survey and demographic analyses.[11,12]

## Interagency Review

Major demographic changes (increases in immigration and multiracial families) and growing public awareness of the social and political importance of data combined to induce public scrutiny of OMB Directive No. 15. In 1994, OMB formed the Interagency Committee for the Review of the Racial and Ethnic Standards for the Classification of Federal Data on Race and Ethnicity, which initiated a process of research, public testimony, and documentary review. Subjects of the most intensive investigation were inclusion of a multiracial category, addition of "Hispanic" as a "racial" rather than an ethnic category, classification of Native Hawaiians as American Indians rather than Pacific Islanders, changes in terminology ("African American" instead of "Black;" "Native" American instead of "American Indian or Alaskan Native"; "Latino" instead of "Hispanic"), and addition of several new ethnic categories ("Middle Eastern," "Arab," "Cape Verdean," "Native Hawaiian").

The Interagency Committee established principles for the review process and gathered extensive evidence to weigh the scientific, social, and political merits of the many suggestions it received. The controversy receiving the most publicity was the demand for a "multiracial" category allowing people to acknowledge their complete family heritage without, for example, being forced to identify with one parent to the exclusion of the other. Opponents argued for maintaining the comparability of data with existing sources and avoiding the potential diminution of certain "racial" and ethnic groups (which could affect the allocation of Federal resources).

After extended periods of public comment and lengthy consideration of special studies and public sentiment, the Interagency Committee published its unanimous recommendations in the July 9, 1997, *Federal Register* (see box on page 480). *[Ed. note: OMB was expected to announce its final decision regarding these recommendations in mid-October, 1997.]*

Of the major changes considered by the Committee, the only one adopted in its recommendations was to allow for multiple responses on a self-identified "race" question. The Committee reaffirmed self-identification of "race"/ethnicity as the preferred method of data collection, adding to the complexity of self-identification allowed, but rejected the proposal for a "multiracial" category.

The only changes in terminology recommended by the Committee were use of "Black or African American" instead of "Black" and "Alaska Native" instead of "Alaskan Native, Eskimo, or Aleut." In addition, the "American Indian or Alaska Native" category would be expanded to include the original peoples of South and Central America, but no new ethnic categories would be added.

## Future Prospects

The only immediate consequences of revised standards would be changes in data collection for the 2000 Census. Problems associated with data tabulation and reporting and with data collection by other agencies have been deferred. The Interagency Committee recommended that "[t]o assist the agencies, OMB should issue guidelines on data tabulation and reporting, instructions for interviewers, and suggested wording for questions by January 1, 1999.... Federal and state agencies are encouraged to work together, under the auspices of OMB, to develop methods that would produce consistent results for program purposes and for comparisons with historical data."[1] Compliance by Federal agencies other than the Census Bureau would be required by January 1, 2003. Agencies would face major challenges, and it is possible that new methods of collecting "racial" and ethnic data would introduce more problems than solutions.

## Devil in the Details

The impetus for reviewing data standards was the

growing diversity of the U.S. population, but the danger of the OMB recommendation is that we may be taking a step backward from self-identification. Permitting multiple responses to "racial" identification opens the door for greater complexity of self-identification and offers the potential for more detailed examination of social inequalities as determinants of health and well-being. Yet the resulting complexity could also cause the process to backfire.

As Eric Rodriguez of the National Council of La Raza said in response to the decision on multiple check-offs: "Really, the devil is in the details. It really depends on how they collect and use the information. That tells you whether policy will be affected or data will be obscured."[13]

The Interagency Committee advocated detailed reporting on the number and nature of multiple "racial identities," yet this leaves statisticians with stark choices. Separate tallies of all "racial" combinations risk numbers too small for statistical meaning, and lumping of "mixed" identities simply results in a "multiracial" category—the solution that was avoided because it might "add to racial tensions and further fragmentation of our population." By not asking individuals their preference for a single or primary identity, statistical imputations of "race" will be based on population distributions—inevitably flawed for small and dispersed groups. Indeed, assignment of "race" based on any kind of chosen algorithm requires a ranking of identities based on criteria other than—and perhaps contradictory to—self-identification. It would be unfortunate to lose the opportunities provided by the new level of detail and tragic to relapse into biased "racial" hierarchies based on new data.

Like many compromises, the OMB decision may ultimately fail to satisfy any stakeholders. However, if nothing else, the process of reviewing "racial" and ethnic classifications has brought into focus the universality of multiracial heritage and the fluidity of "racial" identity. Sociologist Orlando Patterson's reaction to the OMB decision[14] was a plea to reject "racial" labels in favor of more descriptive ethnic terms such as Japanese American, Chinese American, or Pakistani American, a solution not unlike recent guidelines proposed by the *British Medical Journal*.[15]

While the value of "racial" and ethnic classification in U.S. statistics will continue to be debated, the 1993 CDC/ATSDR recommendations[6] delineate the possibility of progressive uses of "racial"/ethnic data, including documentation of the effects of racism. Attention to data issues engendered by the OMB review and the upcoming decennial Census offers an opportunity to take a broader view of the meaning and significance of "race" and ethnicity.

The public health community's participation in the "national conversation on race" must entail not only a commitment to social justice but also a renewed search for causal explanations of health disparities experienced by "racial" and ethnic groups and frequently (but mistakenly) attributed to "race" or "ethnicity." Whatever terminology and reporting mechanisms are utilized, the fulfillment of public health objectives demands the elimination of inequities in health risk factors, access to health care services, and health status indicators. When such disparities correlate with commonly perceived groupings, we must intensify the search in public health research and practice for an accurate understanding of the physical and social environments, working and living conditions, perceptions, and behaviors that selectively impair the longevity, health, and well-being of the U.S. population.

---

Trude Bennett is an Assistant Professor of Maternal and Child Health and Co-Investigator for the Assessment of Statistical Materials and Methods for Minority Health Research Project at the University of North Carolina at Chapel Hill School of Public Health.

*Address correspondence to Dr. Bennett, Dept. of Maternal and Child Health, School of Public Health, CB# 7400, Rosenau Hall, Univ. of North Carolina at Chapel Hill, Chapel Hill NC 27599-7400; tel. 919-966-5977; fax 919-966-0458; e-mail <trude_bennett@unc.edu>.*

## References

1. Recommendations from the Interagency Committee for the Review of the 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 [notice]. Federal Register 1997;62: 36873–946.
2. Evinger S. How shall we measure our nation's diversity? Chance 1995;8:7–14.
3. Osborne NG, Feit MD. The use of race in medical research. JAMA 1992;267:275–9.
4. Schulman KA, Rubenstein LE, Chesley FD, Eisenberg JM. The roles of race and socioeconomic factors in health services research. Health Serv Res 1995;30:180–95.
5. Kaufman JS, Cooper RS. In search of the hypothesis. Public Health Rep 1995;110:662–6.
6. Use of race and ethnicity in public health surveillance: summary of the CDC-ATSDR Workshop. MMWR Morb Mortal Wkly Rep 1993;42(RR-10):12–16.
7. Krieger N, Williams DR, Moss NE. Measuring social class in U.S. public health research; concepts, methodologies, and guidelines. Annu Rev Public Health 1997;18: 341–78.
8. Hahn RA. The state of Federal health statistics on racial and ethnic groups. JAMA 1992;267:268–71.
9. Hahn RA, Mulinare J, Teutsch SM. Inconsistencies in coding of race and ethnicity between birth and death in U.S. infants: a new look at infant mortality, 1983 through 1985. JAMA 1992;267:259–63.
10. Kozak JL. Underreporting of race in the National Hospital Discharge Survey. Advance Data from Vital and Health Statistics No. 265. Hyattsville (MD): National

Center for Health Statistics; 1995.
11. Hogan H. The 1990 Post-Enumeration Survey: operation and results. J Am Statistical Assoc 1993;88:1047–60.
12. Robinson JG, Ahmed B, Gupta PD, Woodrow KA. Estimation of population coverage in the 1990 United States Census based on demographic analysis. J Am Statistical Assoc 1993;88:1061–71.
13. Barr S, Fletcher MA. U.S. proposes multiple racial identification for 2000 census. The Washington Post 1997 Jul 9;Sect. A:1.
14. Patterson O. The race trap. The New York Times 1997 Jul 11;Sect. A:21.
15. Ethnicity, race, and culture: guidelines for research, audit, and publication. BMJ 1996;312:1094.

## Interagency Committee's Recommended Changes to OMB Standards on "Race and Ethnicity Data"

**The following is excerpted from Office of Management and Budget Directive No. 15, *Recommendations From the Interagency Committee for the Review of the 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* (July 9, 1997). Bold print represents the Interagency Committee's recommended changes to the 1977 standards.**

*6.2.2 Recommended Standards*

The **minimum categories for data on race and ethnicity** for Federal statistics and program administrative reporting are defined as follows:

a. *American Indian or **Alaska** Native.* A person having origins in any of the original peoples of North **and South America (including Central America)**, and who maintains cultural identification through tribal affiliation or community recognition.

b. *Asian or Pacific Islander.* A person having origins in any of the original peoples of the Far East, Southeast Asia, the Indian subcontinent, or the Pacific Islands. This area includes, for example, China, India, Japan, Korea, the Philippine Islands, **Hawaii**, and Samoa.

c. *Black **or African-American**.* A person having origins in any of the black racial groups of Africa.

d. *Hispanic.* A person of Mexican, Puerto Rican, Cuban, Central or South American or other Spanish culture or origin, regardless of race.

e. *White.* A person having origins in any of the original peoples of Europe, North Africa, or the Middle East.

To provide flexibility **and assure data quality**, it is preferable to collect data on race and ethnicity separately. **When race and ethnicity are collected separately, ethnicity should be collected first. Persons of mixed racial origins can, but are not required to, report more than one race.** If **race and ethnicity are collected separately**, the minimum designations are:

a. *Race:*
—American Indian or **Alaska** Native
—Asia or Pacific Islander
—Black **or African-American**
—White

b. *Ethnicity:*
—Hispanic origin
—Not of Hispanic origin

**When the data are reported, a minimum of one additional racial category, designated "More than one race," must be included, if the criteria for data quality and confidentiality are met, in order to report the aggregate number of multiple race responses. Data producers are encouraged to provide greater detail about the distribution of multiple responses. Terms such as "Haitian" or "Negro" can be used in addition to "Black" and "African-American." Terms such as "Latino" or "Spanish origin" can be used in addition to "Hispanic."**

If a combined format **must be** used to collect racial and ethnic data, **both race and ethnicity or multiple races should be collected when appropriate, although the selection of one category will be acceptable. If a combined format is used, the minimum categories are:**

—American Indian or Alaska Native
—Asian or Pacific Islander
—Black or African-American
—Hispanic
—White

**When the data are reported, a minimum of two additional categories, designated "Hispanic and one or more races" and "More than one race," must be included if the criteria for data quality and confidentiality are met and both race and ethnicity and multiple races were collected.**

In no case should the provisions of this Directive be construed to limit the collection of data to the categories described above. **In fact, the collection of subgroup detail is encouraged.** However, any reporting required which uses more detail shall be organized in such a way that the additional categories can be aggregated into these minimum categories for data on race and ethnicity.

*The full recommendations, published in the July 9, 1997,* Federal Register, *are available from the OMB at* **www.whitehouse.gov/wh/eop/omb/html/fedreg.html** *or from the OMB Publications Office, 727 17th St. NW, NEOB, Rm. 2200, Washington DC 20503; tel. 202-395-7332; fax 202-395-6137.*