SAFER • HEALTHIER • PEOPLE™

CDC™

Monitoring the Nation's Health

Vital and Health Statistics

Series 4, Number 31

September 2001

Methodological Issues for Vital Rates and Population Estimates: 1997 OMB Standards for Data on Race and Etnicity

DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention • National Center for Health Statistics

Copyright information

All material appearing in this report is in the public domain and may be reproduced or copied without permission; citation as to source, however, is appreciated.

Suggested citation

Durch JS, Madans JH. Methodological issues for vital rates and population estimates: The 1997 OMB standards for data on race and ethnicity. National Center for Health Statistics. Vital Health Stat 4(31). 2001.

Library of Congress Cataloging-in-Publication Data

Madans JH
Methodological Issues for Vital Rates and Population Estimates: 1997 OMB Standards for Data on Race and Ethnicity
p.cm.—(Vital and health statistics. Series 4,
ISBN 0-8406-0570-6
1. Ambulatory medical care—Utilization—United States—Statistics. 2. Women's health services—Utilization—United States—Statistics. 3. Physician services utilization—United States—Statistics. 4. Women—Diseases—United States—Statistics. 5. Medical care surveys—United States. I. Madans JH II. Series.
RA407.3 .A349 no. 31
362.1'0973'02 s—dc21
[362.1'082'0973] 2001030839

For sale by the U.S. Government Printing Office
Superintendent of Documents
Mail Stop: SSOP
Washington, DC 20402-9328
Printed on acid-free paper.

Vital and Health Statistics

Series 4, Number 31

Methodological Issues for Vital Rates and Population Estimates: The 1997 OMB Standards for Data on Race and Ethnicity

Documents and Committee Reports

U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES
Centers for Disease Control and Prevention
National Center for Health Statistics

Hyattsville, Maryland
September 2001
DHHS Publication No. (PHS) 2001-1468

National Center for Health Statistics

Edward J. Sondik, Ph.D., *Director*

Jack R. Anderson, *Deputy Director*

Jack R. Anderson, *Acting Associate Director for International Statistics*

Jennifer H. Madans, Ph.D., *Associate Director for Science*

Lawrence H. Cox, Ph.D., *Associate Director for Research and Methodology*

Jennifer H. Madans, Ph.D., *Acting Associate Director for Analysis, Epidemiology, and Health Promotion*

P. Douglas Williams, *Acting Associate Director for Data Standards, Program Development, and Extramural Programs*

Edward L. Hunter, *Associate Director for Planning, Budget, and Legislation*

Jennifer H. Madans, Ph.D., *Acting Associate Director for Vital and Health Statistics Systems*

Douglas L. Zinn, *Acting Associate Director for Management*

Charles J. Rothwell, *Associate Director for Information Technology and Services*

Contents

Abstract 1
Introduction 1
Background 2
Population Estimates and Vital Statistics 2
Primary Concerns 3
Research Strategies 4
Matching Studies 4
Vital Record Studies 6
Studies of Error 7
Studies on Racial and Ethnic Identity 7
Other Information Needs 8
References 8
Appendix I. OMB Federal Register Notice, October 30, 1997: Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity 9
Appendix II. Workshop Participants 27
Appendix III. Questions on Race and Hispanic Origin 28

Appendix Figures

Figure I. Questions on race and hispanic origin from Census 2000 28
Figure II. Questions on Hispanic origin and race on the current U.S. standard certificate of live birth 28
Figure III. Questions on Hispanic origin and race on the draft revised U.S. standard certificate of live birth 29
Figure IV. Questions on Hispanic origin and race on the current U.S. standard certificate of death 30
Figure V. Questions on Hispanic origin and race on the draft revised U.S. standard certificate of death 30

Abstract

A workshop in July 2000 explored research needed to address methodological challenges for population estimates and vital rates arising from the revision of the standards for Federal data on race and ethnicity; the use of different standards for Census 2000 and vital records during the implementation of the revised standards; and underlying differences in the collection of race and ethnicity data in censuses and surveys and in vital records. *Matching studies* were proposed to compare race and ethnicity reports in vital records and in Census 2000 or survey responses. Work on *vital records* might include exploring enhanced collection of race and ethnicity data and documenting State vital statistics reporting mandates and practices. Key work on *sources of error* includes identifying, quantifying, and reporting on bias and random errors related to race and ethnicity in population estimates and vital rates. Also needed are comparisons of Census Bureau and State population estimates and more frequent tests of the accuracy of population estimates and projections. Studies on *racial and ethnic identity* were proposed to examine changes in reporting over time and to explore origins of racial and ethnic identities. More information on these issues should be developed for technical and nontechnical audiences.

Keywords: *population estimates • birth rates • death rates • race and ethnicity • data collection*

Methodological Issues for Vital Rates and Population Estimates: 1997 OMB Standards for Data on Race and Ethnicity

Jane S. Durch, M.A., George Washington University, and Jennifer H. Madans, Ph.D., National Center for Health Statistics

Introduction

In October 1997, the Office of Management and Budget (OMB) issued revisions to Statistical Policy Directive No. 15, Race and Ethnicity Standards for Federal Statistics and Administrative Reporting (1, and see Appendix I). The revised standards introduced new categories for collecting and reporting data on race and ethnicity and also require that Federal data collection programs allow respondents to select more than one race. Use of these revised standards has major implications for population and health statistics.

The revised standards were implemented for the collection of population data in Census 2000, but they will not be reflected in the standard birth and death certificates until 2003 at the earliest. Moreover, although the States collaborate with the National Center for Health Statistics (NCHS) in developing those standard certificates, authority for determining the specific data elements to be collected officially rests with the States. At least one State implemented the revised race and ethnicity data standards for birth and death certificates in 2000 while most others will not do so until later in the decade.

The use of two different standards for collection of race and ethnicity data during the transition to full implementation of the revised OMB standards is a significant but short-term concern. A more fundamental and longstanding concern is that data on race and ethnicity from censuses and vital records are never totally comparable because of differences in the way the data are collected. Census data are primarily based on self-identification while data from death records depend on identification of a subject's race and ethnicity by other informants or observers. Birth records primarily rely on self-identification, but the data on race and ethnicity are for the parents and do not refer directly to the newborn.

The comparability of data from decennial censuses and vital records is of special concern to the OMB, Census Bureau, and NCHS because both sorts of data are used to construct population estimates, and the population estimates are used in turn to calculate vital rates.[1] At the request of these agencies, the Epidemiology and Biostatistics Department of the George Washington

[1]Additional information related to the topics addressed at the workshop is available from the agencies at their Web sites.

Census Bureau: www.census.gov
Census2000: www.census.gov/dmd/www.census.gov/CMS/www/
National Center for Health Statistics: www.cdc.gov/nchs
National Vital Statistics System: www.cdc.gov/nchs/nvss.htm
Births data: www.cdc.gov/nchs/births.htm
Deaths data: www.cdc.gov/nchs/about/major/dvs/mortdata.htm
National Death Index: www.cdc.gov/nchs/r&d/ndi.htm
Office of Management and Budget: www.whitehouse.gov/omb

University School of Public Health and Health Services and the Metropolitan Washington Public Health Assessment Center held a workshop on July 6–7, 2000, for discussion of these issues and of short- and long-term research needs. The workshop brought together key staff members from the Federal agencies, academic experts, and State officials with responsibilities for population estimates and vital statistics. The discussions included a review of the technical challenges that the transition to the revised Federal data standards pose for the calculation of postcensal population estimates and vital rates and consideration of research that could help improve the quality of those data.

Background

The 1997 revision of the OMB standards for data on race and ethnicity increased from four to five the minimum set of categories to be used by Federal agencies for identification of race. As in the past these categories represent a social-political construct and are not anthropologically or biologically based. The five categories for race are:

1. American Indian or Alaska Native
2. Asian
3. Black or African American
4. Native Hawaiian or Other Pacific Islander
5. White

The 1997 standards continue to call for use, when possible, of a separate question on Hispanic or Latino ethnicity and specify that the ethnicity question should appear before the question on race. Research suggested that this question order could help reduce the proportion of Hispanic respondents who report ''other race'' or who give no response to the separate question on race (2,3). Collection of additional detail on race or ethnicity is permitted so long as the additional categories can be aggregated into the minimum categories.

The revised standards also add a requirement to allow reporting of more than one race for an individual. Overall, the level of multiple-race reporting is expected to be low, but it is likely to vary among population groups (3–7). Data from the National Health Interview Survey (NHIS) for 1993–95 show that 1.6 percent of respondents reported more than one race (5). However, among the respondents who reported American Indian or Alaska Native race, more than 50 percent also reported at least one other race. In the 1996 Race and Ethnicity Targeted Test (RAETT), multiple-race reporting ranged from about 1.4 percent in the sample targeted to white respondents to about 10 percent in the sample targeted to Asian and Pacific Islander respondents (3). Although multiple-race reporting in the NHIS has remained relatively consistent since 1982, an analysis of birth certificate data from 1971 to 1995 shows an increase from 1.2 percent to 4.6 percent in the percent of births to parents of different races, suggesting that multiple-race reporting in censuses and surveys may increase in the future (5, 8).

The level of multiple-race responses can also be expected to vary across States and communities. In California 5.5 percent of respondents in the 1998 Sacramento dress rehearsal for Census 2000 selected more than one race (9). As of January 2000 California birth registration captures and codes reports of up to three races for each parent, and a preliminary analysis of reports for the first few months of 2000 shows, using the five OMB race categories, that 1.8 percent of mothers reported multiple races (10). But, as reported at the workshop, 16 percent of births show a multiple-race background when the race of both parents was considered, reflecting a combination of births with a least one multiple-race parent and births with parents of different races. The Hawaii Health Survey (HHS) gives respondents the opportunity to report up to four races for each household member but uses more race categories than the five specified by the OMB standards. Tabulations from the 1998 HHS showed that 33 percent of that State's population had multiple-race backgrounds on the basis of the larger set of race categories (11).

The introduction of multiple-race reporting is producing new challenges in the tabulation of data and in comparisons with earlier data produced under the prior standards. By April 2001 the first data from census 2000 will be available at the national, State, and local levels for 63 categories of race in each of the 2 categories of ethnicity (Hispanic or Latino; Not Hispanic or Latino). However, routine use of all 126 race-ethnicity categories in most data products, including population estimates and vital statistics, will not be practical because the small number of cases in many of the categories will not satisfy requirements for data quality or protection of confidentiality. Current tabulation guidance directs Federal agencies to provide the total number of persons identified as of more than one race and to include as much detail as possible on multiple-race responses, consistent with criteria for data quality and confidentiality (5).

Two basic tabulation options are available for combining multiple-race responses into a smaller set of categories. ''Exclusive'' tabulation assigns each response to one category in a set of mutually exclusive race categories (for example, the five standard single-race categories plus the most common multiple-race combinations and aggregated categories like ''three or more races''). ''Inclusive'' tabulation assigns a multiple-race response to each of the single-race categories mentioned; for example, an ''Asian White'' response would be counted in the Asian category and in the white category. An exclusive tabulation produces an unduplicated count of respondents, while an inclusive tabulation produces a count of the total number of race responses. The choice of tabulation categories and the tabulation approach will depend on the intended uses of the data.

Population Estimates and Vital Statistics

The workshop focused on special concerns related to the application of the 1997 OMB data standards to population estimates and vital

statistics. Calculation of population estimates depends on population counts from decennial censuses and on data on the components of demographic change, which include vital statistics data on births and deaths plus estimates of migration. Vital rates—birth and death rates, specifically—are calculated from numerator data obtained from reports of births and deaths and denominator data on population size that are generally from population estimates.

Although data from censuses and vital statistics must regularly be used together, there are some key differences between them in the way information on race and ethnicity is collected. The decennial census relies primarily on self-identification of race and ethnicity by at least one adult respondent in a household, and on proxy reports on the race and ethnicity of children and other adults in the household. See figure I for the census 2000 ethnicity and race questions. Birth records include information, usually obtained by self-report, on the race and ethnicity of the parents of newborns (although one parent might report for the other), but race and ethnicity are generally not reported for newborns. Currently, birth data are usually tabulated according to the race and ethnicity of the mother, regardless of the race and ethnicity of the father. Displayed in figures II and III are the race and ethnicity questions as they appear on the current standard birth certificate and as planned for the revised standard certificate.

For deaths, race and ethnicity are reported by an informant or observer. This person may be a family member of the deceased, but may also be a funeral director or other informant with little or no knowledge of the deceased's background. Figures IV and V show the race and ethnicity questions appearing on the current standard death certificate and as planned for the revised standard certificate. An assessment of the quality of death rates found that rates for the white population and the black population are generally reliable but overstated (by 1 percent and 5 percent, respectively), primarily because of population undercounts affecting the denominators (12). For the Asian and Pacific Islander population and the American Indian or Alaska Native population, misclassification of race on death certificates produces a substantial understatement of death rates: about 11 percent for Asians and Pacific Islanders and about 21 percent for American Indians and Alaska Natives (12).

Decennial census and vital statistics data also differ in that collection of the census data is a constitutionally mandated responsibility of the Federal Government, whereas authority for the collection of vital statistics rests with the States. The States voluntarily work with NCHS to establish a set of common practices regarding vital records, and those practices are reflected in standard birth and death certificates produced by NCHS. States, however, retain the authority to collect and tabulate data in ways that may differ from Federal guidelines. California, for example, produces State population estimates and vital statistics by race using a set of mutually exclusive tabulation categories that includes "Hispanic" along with the race categories specified by OMB. For vital statistics, the State combines information from the separate questions on ethnicity and race on birth and death certificates to tabulate data using categories of race that match those for the State population data. Persons of Hispanic ethnicity are assigned to the Hispanic category regardless of race. For Federal reporting purposes, California recodes as white those persons who report their race as Mexican or other designations considered to be of Hispanic ethnicity.

Of specific concern for the near term is, as noted earlier, that while census 2000 followed the 1997 OMB standards in collecting population data on race and ethnicity, all but one or two States are still following the prior standards in collecting data on births and deaths. The revised standards for race and ethnicity data will be in place for vital statistics, and reflected in new standard birth and death certificates, later in the decade. During the transition to full implementation of the 1997 OMB standards and to population estimates based on census 2000, the Census Bureau and NCHS will be studying how the data are changing and ways to facilitate the transition. During the transition period NCHS and the Census Bureau must also work together to develop the appropriate detailed population estimates for use as denominators in the calculation of vital rates.

NCHS hopes to be able to provide assistance to States to help them update electronic birth certificate systems to accommodate the elements of the new standard certificates, including the revised data standards for race and ethnicity. With work just beginning on electronic death certificate systems, it should be possible to design them to be consistent with the revised data standards from the start. As these systems are updated or developed, some States may be able to provide NCHS with additional detail on race and ethnicity as it is originally recorded on birth and death certificates. The added detail from the certificates may allow NCHS to recode data to match either the old or revised reporting standards. Under newly established agreements with NCHS, the States of Alaska, California, Hawaii, and Washington will begin submitting all terms found in the section of the certificate where race is reported (currently only the first listed race is transmitted to NCHS). As of August 2000 California began submitting data for 2000, and the other three States began developing the necessary coding and reporting processes to be able to submit their data. With access to these data, NCHS will be able to test computational procedures and presentation formats for multiple-race data before the revised standard certificates are implemented.

Primary Concerns

The workshop discussions emphasized the fundamental need for vital statistics and population data that can meet requirements for accuracy, consistency, meaningfulness, and acceptability. But data may be needed in different forms for different purposes. For example, distributions based on inclusive tabulations of

multiple-race responses, which count a response in each single-race category reported and total to more than 100 percent of the actual population, may be valuable for analyses of certain aspects of population and social change. Program planning and administration, however, are more likely to require data from exclusive tabulations that use a set of mutually exclusive race categories and count each person only once. It may also prove difficult to produce data that can satisfy all requirements equally well. Population estimates with some detail on race and ethnicity are needed and used at the county and community level even though the error in such estimates is likely to be greater than for larger geographic units.

The various issues raised at the workshop were summarized as four basic concerns toward which new research efforts might be directed. These four concerns are:

1. Producing valid and reliable population *estimates* at the most feasible levels of detail for race and ethnicity and geography for the period while population and vital statistics data are being collected under different standards for classification of race and ethnicity, and for subsequent years when data collection standards become consistent but classification differences related to self-identification of race and ethnicity compared with informant/observer identification persist.
2. Similarly, producing valid and reliable vital *rates* at the most feasible levels of detail for race and ethnicity and geography for the transition period and for subsequent years.
3. Understanding and measuring error related to race and ethnicity in vital rates and population estimates to guide assessments of data quality for publication and use of vital rates and population estimates.
4. Gaining a better understanding of the phenomenon of racial and ethnic identity in terms of self-identification and identification of others.

The first three concerns are principally methodological, and addressing them calls for studies of matters such as quantifying the level of consistency in racial and ethnic identification over time or determining appropriate tabulation and adjustment factors between separate data sources. The fourth concern regarding racial and ethnic identification is not specifically a methodological matter, but a better understanding of the factors that influence racial and ethnic self-identification and identification of others would help guide the collection, tabulation, and analysis of data.

Research Strategies

Workshop participants identified several research strategies that could be used to address the technical and methodological issues that the 1997 OMB standards for data on race and ethnicity raise for vital statistics and population estimates, and to address some of the more conceptual and perceptual aspects of race and ethnicity that influence reporting in censuses, surveys, and vital records.

Matching Studies

There was considerable interest in using matching studies to learn more about the correspondence between reports of race and ethnicity in census 2000 and various surveys compared with vital records. Such studies could help determine if there are systematic differences between these data sources in reporting of race or ethnicity for specific race groups, age groups, or geographic areas. If so, it might be possible to construct adjustment factors that would improve the consistency of numerator and denominator data for vital statistics, and by extension the accuracy of population estimates. Matching studies might also help in determining how best to bridge from the series of population estimates based on the 1990 census data and the prior reporting standards for race and ethnicity data to a new series based on Census 2000 data and the new single- and multiple-race categories that will result from use of the revised reporting standards. Workshop participants felt that census matching studies would be valuable for many such reasons, and they suggested planning for these studies following every decennial census.

Several approaches to matching studies were discussed and are summarized below. Because the population in some race groups is small relative to the total population and multiple-race reporting is relatively infrequent, census-based matching studies were thought to be more promising than matches to the smaller study populations in surveys. But matching studies based on surveys, such as the NHIS, the American Community Survey (ACS), or the Current Population Survey (CPS), could be valuable for pilot studies of matching procedures or for continuing prior studies and learning about changes over time in reporting of race and ethnicity.

Census 2000 matching studies—In such studies, the race and ethnicity reported on birth or death certificates could be compared with the race and ethnicity reported for those individuals in census 2000. Studies would involve matching birth records from a period immediately preceding the census (for example, July 1999–March 2000) or death records from a period immediately following the census (for example, May–December 2000) with the corresponding census 2000 record.

Workshop participants considered a study to match death records and census 2000 records a high priority. There is as yet little basis for judging how persons who report multiple races in the census will be identified by the informants or observers who provide information on race for death certificates. About 800,000 to 1 million deaths would be expected from May through December 2000. It was estimated that about 75 percent of these records could be matched to census 2000 records. Achieving a more complete match may not be cost-effective because the census records do not include social security numbers, which are a valuable (but not foolproof) key for record linkages. The National Death Index (NDI) provides a

mechanism for identifying the original death certificates filed in each State. Finding the State records would be valuable because those records sometimes contain multiple-race information that has generally not been reported to NCHS.

The proposed analysis is similar to a classic study by Kitagawa and Hauser, in which death records were matched with 1960 census records to study socioeconomic differentials in mortality (13). With a large enough set of matched records in a contemporary study, it might be possible to use data from the census 2000 long-form responses to learn more about socioeconomic correlates of single- or multiple-race responses. Such data would also provide valuable information for studies of racial and ethnic identity, one of the other research priorities identified by workshop participants (and discussed later in this report). In addition, the linked data would make it possible to study racial and ethnic differentials in mortality in greater detail than is possible with existing data sources.

A study to match birth records and census 2000 records would also be valuable. Data from birth records are generally tabulated according to the race and ethnicity of the mother without regard to the father's race and ethnicity, and little is known about how the maternal classification corresponds to the race and ethnicity, subsequently, attributed to the child in the decennial census. Of particular interest would be the race and ethnicity reported in census 2000 for children whose birth records show parents of different races or ethnicities or for whom information on paternal race and ethnicity is missing. Original State birth records, such as the California records mentioned previously, would provide additional detail, but the lack of a national birth index comparable to the NDI would complicate the process of identifying those records. Because birth records are available only for children born in the United States (and its territories), foreign-born children would remain unmatched in such studies. However, the contribution of foreign-born children to multiple-race reporting in the census may be small.

Although there was general agreement that studies to match birth or death records with census 2000 records would be valuable, several constraints were noted. In the period following the census, Census Bureau staff must produce many products under demanding deadlines and so will have limited time to devote to the additional work that matching studies would require. A delay in conducting a matching study will, however, limit its usefulness for improving the consistency of vital statistics and population data during the transition to reporting under the revised race and ethnicity data standards. Consideration would also have to be given to the appropriate level of investment in reconciling unmatched and mismatched records. Larger numbers of unmatched records might result in underestimates of reporting discrepancies between the census and vital records. It would also be necessary to ensure that State confidentiality requirements with respect to birth and death records were met.

Survey-based matching studies— Several surveys also might be used for matching studies.

American Community Survey (ACS). The ACS is designed to produce community-level demographic and socioeconomic data more frequently than can be done by the decennial census, and beginning in 1999, ACS questions on race and ethnicity were the same as those used in census 2000. When fully implemented, the ACS will include 3 million households annually, making it much larger than other survey samples and increasing the feasibility of studies involving the matching of birth or death records to compare reporting on race and ethnicity. In the near term, however, studies involving the smaller scale samples being used to develop and test the ACS might also be useful. For example, it might be possible to track deaths in the 1999 or 2000 cohort of ACS respondents to assess reports of race and ethnicity on death certificates before and after full implementation of multiple-race reporting for vital records. The 1998 ACS sample incorporated two South Carolina counties that were also census 2000 dress-rehearsal sites in 1998, making it possible to compare responses to the old and the new formats for the race and ethnicity questions. The analysis could be extended by matching those records to reports from census 2000 as well.

The ACS samples might also be used for pilot studies to test matching procedures and algorithms. For example, since neither ACS nor census 2000 records include social security numbers, other matching techniques could be tested. The inclusion of two California counties (San Francisco and Tulare) in the ACS sites for 1999 also presents an opportunity for preliminary comparisons of race reporting in vital records (beginning in 2000) and census-type records when both offer a multiple-race option. With the planned continuation of ACS data collection at the California test sites through 2002, the comparisons could include matches with either birth or death records.

Current Population Survey and National Longitudinal Mortality Study The Current Population Survey (CPS), a monthly survey of about 50,000 households, is conducted primarily to collect data on labor force characteristics of the population but is also a regular source of other demographic and socioeconomic information. The National Longitudinal Mortality Study (NLMS), a collaborative project by the National Institutes of Health, the Census Bureau, and NCHS, has matched selected CPS records for 1973–85 to death certificates for 1979–89 through the NDI (12). Plans call for updating the death certificate match through the 1990s. As noted above, these matched records have been used, among other purposes, to study the quality of race data reported on death certificates.

Because of its age, the current NLMS database cannot provide information on multiple-race reporting, but it offers an existing and tested framework within which new CPS cohorts could be added and matched to death certificates in order to study changes in race reporting and the consistency between survey and death certificate reporting during the transition

from the old to the revised standards for race and ethnicity data. The NLMS is best suited to studies at the national level because too few deaths could be linked to CPS cohorts to produce reliable results at the State or community level for multiple-race or small single-race groups.

National Health Interview Survey. The National Health Interview Survey (NHIS), an annual survey covering about 43,000 households, is a primary source of information about the health of the population. NHIS cohorts are being used to establish a database of survey records matched with death certificates, and NHIS records have the advantage over CPS records of including, since 1976, multiple-race responses that are accompanied by a response on a primary-race identification. That information makes it possible to study patterns of self-identification in single- and multiple-race categories, and with matches to death certificates, corresponding observer assignment into single-race categories. For NCHS and other entities interested in the relationship between health disparities and race and ethnicity, an NHIS-based matching study would also have the advantage of providing access to data on health status and health risk factors, which are not available from CPS records. Workshop participants also noted that a subset of the 1993 NHIS respondents (10,847 women aged 15–44) were interviewed for the 1995 cycle of the National Survey of Family Growth (NSFG) (14), potentially offering access to additional information about the characteristics of some multiple-race respondents.

Other matching studies—Other matching study approaches or data sources were also discussed at the workshop.

National Mortality Followback Survey. The National Mortality Followback Study (NMFS), which is conducted periodically by NCHS, involves drawing a sample of death reports and contacting the death certificate informant. This kind of study could be used to obtain supplemental information and to replicate information reported on the death certificate. The NMFS approach can provide information on consistency of reporting from a single source but not on consistency across reporting systems. A new round of the NMFS might, however, ask informants to respond using the multiple-race reporting format as well as the single-race reporting format of the current death certificate. An informant-followback component might also be included in a study based on matching census 2000 records and death reports. The size of the NMFS sample would have to be increased from past rounds (22,957 decedents in the 1993 round) to provide reliable data about smaller race groups or multiple-race reporting.

Immigration and Naturalization Service data. Workshop participants discussed the possibility of adding Immigration and Naturalization Service (INS) data to matching studies to improve data on the foreign-born population, particularly those of Hispanic origin. The Hispanic population is among those undercounted in the census. In addition, the Hispanic population may be undercounted in the intercensal estimates of documented and undocumented immigration. Also, misclassification to non-Hispanic ethnicity on death certificates contributes to inaccuracies in death rates for this segment of the population. Workshop participants thought that immigrants were likely to be making only a small contribution to the multiple-race population. Studies matching INS records with census reports or vital records might, however, provide better information on immigrant fertility and the contribution of immigration to population growth and to racial and ethnic diversity in the United States.

Despite the interest in trying to use INS data, little was known about the technical feasibility of matching those records to decennial census records or death certificates, or about the confidentiality constraints that might exist. Several limitations of such data were noted. In particular, INS files cover only persons who have permanent resident status. They do not account for emigrants, persons in the United States on a temporary basis (for example, students), or most importantly, unauthorized immigrants.

Vital Records Studies

Studies focusing specifically on vital records, particularly birth certificates, were also discussed at the workshop. Although NCHS works with the States to define a standard set of data on births and deaths to be reported to the Federal level, the original data collection process is established and managed by each individual State. As a result, a few States are already collecting additional data on race and ethnicity that could be used to study reporting in vital records and to explore how reporting may change under the revised OMB standards. Specific ideas for vital records studies included the following topics:

Race of child. Although data on births are currently tabulated by the race and ethnicity of the mother, Washington State birth records also have separate questions on the race and Hispanic origin of the child as reported by the mother or other informant. This information, which was recently provided to NCHS but has not been widely analyzed, presents an opportunity to assess the comparability of race and ethnicity for mothers and their children. States have generally discouraged proposals to add routine collection of separate data on race and Hispanic origin of the child. Washington State might serve as a case study for exploring whether special difficulties exist in collecting or tabulating such data. With more information about Washington State's experience, other States might be willing to test the collection of data on child's race and ethnicity.

Regional Study Program—Data from vital statistics or health surveys in Alaska, California, Hawaii, and Washington provide a collection of resources that might become a base for a regionally focused set of studies to explore application of the revised OMB standards for data on race and ethnicity. As noted above, NCHS already established agreements with these four States to begin providing multiple-race data from vital records as soon as 2000

or 2001. NCHS assistance might also make it possible for this set of States or others to code and tabulate additional data for submission to NCHS, such as the data on race of child in Washington State.

Compilation of State Vital Statistics reporting mandates and practices—As the workshop discussions around data from Alaska, California, Hawaii, and Washington State demonstrated, the availability of data on race and ethnicity that is more detailed than the data called for by the standard birth and death certificates presents important opportunities to study how the revised OMB standards will affect vital rates and population estimates. To know the full extent of such State-level data resources, better information is needed about the States' specific vital records data collection mandates and practices. A formal compilation of such information was proposed.

Studies of Error

The opportunity for multiple-race reporting under the revised OMB standards will add new challenges to producing valid and reliable population estimates, especially for smaller race groups and at subnational levels. Two research strategies were proposed to address measurement error and uncertainty in population estimates.

Identifying sources of error—A careful study of the procedures used to produce population estimates in terms of the potential sources of error (bias and random variation) that would be introduced into the data would provide valuable guidance to the producers and the users of those estimates. Some of the sources of error include the census undercount and undercount adjustments, missing data and imputation procedures, and misreporting and misclassification of data. The revised OMB standards will not change the types of error in the data but will introduce specific new forms of error, such as misreporting or misclassification between single- and multiple-race categories that need to be understood. The proposed matching studies would help identify and quantify some types of bias in the data. Techniques like multiple imputation can be used to assess variance. Because the types and levels of error can differ among various demographic and geographic subpopulations and various data sources (for example, census records, birth records, and death records), studies should examine the total error and the net effect of offsetting error.

Workshop participants noted the importance of reporting information about error—and quantifying it when possible, perhaps with confidence intervals—when population estimates are produced. Also important is providing data users with information about the assumptions on which the estimates are based.

Testing the accuracy of estimates and projections—The Census Bureau produces population estimates and projections for States and communities, but States also produce independent estimates of their populations, which may differ from the Census Bureau estimates. For example, at the time of the workshop, California's State population estimate was almost 1 million more than the Census Bureau estimate. Studies are needed to compare the accuracy of Census Bureau and State estimates and to identify the sources of discrepancy. The Census Bureau conducts studies of the "error of closure" between the census population and population estimates derived from the demographic components of change in conjunction with decennial censuses, but such studies are needed more frequently to account for State-level variations in population dynamics.

California data from census 2000, the census dress-rehearsal, and birth and death records might be used to test estimation procedures with data that follow the revised OMB standards. California also has plans to study the State estimates and projections for racial and ethnic subgroups to assist in bridging its data series between the old and revised OMB data standards.

Studies on Racial and Ethnic Identity

Proposals for other studies emphasized opportunities to improve data on race and ethnicity by gaining a better understanding of the meaning of racial and ethnic identities in society and of the factors that influence self-identification and identification of others. Studies might use routinely conducted surveys like the CPS or the NHIS to ask about the race of respondents' parents and other ancestors to learn more about the racial and ethnic backgrounds of respondents who report multiple races. A better understanding might also be gained of factors associated with single-race reporting for respondents with multiple-race backgrounds. An analysis of 1993–95 data from the NHIS found, for example, that although 3.1 percent of children in two-parent households had parents of different races, fewer than one-half of those children were reported to be of more than one race (15).

Another study proposal pointed to the opportunity to use existing NCHS vital statistics records dating back to 1970 to compile data on births with differing maternal and paternal races. These aggregate data on births with multiple-race backgrounds—classified by age in 2000, sex, and possibly other characteristics—could be compared to levels of multiple-race reporting in census 2000 to gain a general sense of potential levels of multiple-race reporting compared with levels actually seen in the census. This approach would avoid the difficulties related to matching individual records and might point to factors associated with differences in identification with multiple races. But because these birth record data include only a single race for each parent, they would miss a multiple-race heritage dating back to an earlier generation. Also proposed was a study of race and ethnicity as reported in census 2000 for people who were children in multiple-race households in the 1980 or 1990 censuses. However, linking records across censuses for such a longitudinal study may not be feasible.

In addition, further consideration could be given to the correspondence between the revised OMB data standards and public perceptions of race and ethnicity. For example, the use of separate questions on Hispanic origin and race may conflict with the evolving

racial identity of the Hispanic population. Workshop participants noted that Californians generally consider Mexicans and other Hispanics to be "people of color," and that many California Hispanics agree that they are not white. Under present practices, however, Hispanics who do not choose one of the standard race categories are generally classified as white in vital statistics.

Other Information Needs

In addition to research strategies, the workshop participants saw a need for information products that would help researchers and the public make better use of data on race and ethnicity from vital statistics, decennial censuses, and population estimates and projections. Many users may not understand the nature of the changes being made as a result of the revised OMB standards or the underlying differences between data based on self-identification and data based on reports from informants or observers.

The research community will be interested in access to detailed data on race and ethnicity in public-use files for census 2000 and other data sets. Complete detail on responses on race and ethnicity, rather than recoded data for a more limited set of categories, will be of greatest use in studying changes in reporting over time and among population groups. Such data products must also ensure that the confidentiality of individual respondents is protected. Whether they are using microdata or other data products, researchers will need good technical information about the quality and the reliability of the data.

Efforts should also be made to reach other audiences with more general information about the changes being made under the revised OMB standards and what those changes mean for the data that will be reported in the future. Potential audiences for such information include Congress, the general public, interest groups, and the broader research community. Information should be presented in forms tailored to meet each group's needs.

References

1. Office of Management and Budget. Revisions to the Standards for Classification of Federal Data on Race and Ethnicity. Federal Register 62(210):58781–90. October 30, 1997. www.whitehouse.gov/OMB/fedreg/1997.html
2. Bureau of the Census. Findings on questions on Race and Hispanic Origin Tested in the 1996 National Content Survey. Staff of Special Population Statistics, Population Division. Population Division Working Paper No. 16. 1996. www.census.gov/population/www/documentation/twps0016/twps0016.html
3. Bureau of the Census. Results of the 1996 Race and Ethnic Targeted Test. Population Division Working Paper No. 18. 1997. www.census.gov/population/www/documentation/twps0018/twps0018.html
4. Office of Management and Budget. 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. Federal Register 62(131):36873–946. July 9, 1997. www.whitehouse.gov/OMB/fedreg/1997.html
5. Tabulation Working Group, Interagency Committee for the Review of Standards for Data on Race and Ethnicity. Provisional Guidance on the Implementation of the 1997 Standards for Federal Data on Race and Ethnicity. December 15, 2000. www.whitehouse.gov/omb/inforeg/r&e_guidance2000update.pdf
6. Goldstein JR, Morning AJ. The multiple-race population of the United States: Issues and Estimates. Proceedings of the National Academy of Sciences 97(11):6230–5. 2000.
7. Hirschman C, Alba R, Farley R. The meaning and measurement of race in the U.S. census: Glimpses into the future. Demography 37(3):381–93. 2000.
8. Ventura SJ, Martin JA, Curtin SA, Mathews TJ. Report of final natality statistics, 1995. Monthly vital statistics report; vol 45 no 11, supp. Hyattsville, Maryland: National Center for Health Statistics. 1997.
9. Dodds DJ. Tabulating diversity: Race and ethnicity categories in census 2000. Applied Demography 13(2):1–6. 2000.
10. Williams R. California's Multiple Race Reporting from AVSS: Preliminary Data. 2000. AVSS Project, Institute for Social, Behavioral, and Economic Research, University of California, Santa Barbara. www.avss.ucsb.edu/OVR/RACE2Kb.htm
11. Baker KK, Onaka AT, Reyes-Salvail F, Horiuchi B, Dannemiller J, Tom T, Zhang J, Tanaka C. Multi-Race Health Statistics: A State Perspective. Hawaii Health Survey (HHS) 1998. 1999. www.hawaii.gov/doh/stats/surveys/nchs.html
12. Rosenberg HM, Maurer JD, Sorlie PD, Johnson NJ, et al. Quality of death rates by race and Hispanic origin: A summary of current research, 1999. National Center for Health Statistics. Vital Health Stat 2(128). 1999. www.cdc.gov/nchs/
13. Kitagawa E, Hauser PM. Differential mortality in the United States: A Study in Socioeconomic Epidemiology. Cambridge, Mass.: Harvard University Press. 1973.
14. Kelly JE, Mosher WD, Duffer AP, Kinsey SH. Plan and operation of the 1995 National Survey of Family Growth. Vital Health Stat 1(36). 1997.
15. Parker JD, Lucas JB. Multiple race reporting for children in a national health survey. Ethnicity and Disease. Forthcoming.

Appendix I. OMB Federal Register Notice, October 30, 1997: Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity

OFFICE OF MANAGEMENT AND BUDGET

Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity

AGENCY: Executive Office of the President, Office of Management and Budget (OMB), Office of Information and Regulatory Affairs

ACTION: Notice of decision.

SUMMARY: By this Notice, OMB is announcing its decision concerning the revision of Statistical Policy Directive No. 15, Race and Ethnic Standards for Federal Statistics and Administrative Reporting. OMB is accepting the recommendations of the Interagency Committee for the Review of the Racial and Ethnic Standards with the following two modifications: (1) the Asian or Pacific Islander category will be separated into two categories -- "Asian" and "Native Hawaiian or Other Pacific Islander," and (2) the term "Hispanic" will be changed to "Hispanic or Latino."

The revised standards will have five minimum categories for data on race: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White. There will be two categories for data on ethnicity: "Hispanic or Latino" and "Not Hispanic or Latino."

The Supplementary Information in this Notice provides background information on the standards (Section A); a summary of the comprehensive review process that began in July 1993 (Section B); a brief synopsis of the public comments OMB received on the recommendations for changes to the standards in response to the July 9, 1997, **Federal Register** Notice (Section C); OMB's decisions on the specific recommendations of the Interagency Committee (Section D); and information on the work that is underway on tabulation issues associated with the reporting of multiple race responses (Section E).

The revised standards for the classification of Federal data on race and ethnicity are presented at the end of this notice; they replace and supersede Statistical Policy Directive No. 15.

EFFECTIVE DATE: The new standards will be used by the Bureau of the Census in the 2000 decennial census. Other Federal programs should adopt the standards as soon as possible, but not later than January 1, 2003, for use in household surveys, administrative forms and records, and other data collections. In addition, OMB has approved the use of the new standards by the Bureau of the Census in the "Dress Rehearsal" for Census 2000 scheduled to be conducted in March 1998.

ADDRESSES: Please send correspondence about OMB's decision to: Katherine K. Wallman, Chief Statistician, Office of Information and Regulatory Affairs, Office of Management and Budget, Room 10201 New Executive Office Building, 725 17th Street, N.W., Washington, D.C. 20503; fax: (202) 395-7245.

ELECTRONIC AVAILABILITY AND ADDRESSES: This **Federal Register** Notice and the related OMB Notices of June 9, 1994, August 28, 1995, and July 9, 1997, are available electronically from the OMB Homepage on the World Wide Web:

<</OMB/fedreg/>>.

Federal Register Notices are also available electronically from the U.S. Government Printing Office web site: <<http:/www.access.gpo.gov/su_docs/aces/aces140.html>>. Questions about accessing the **Federal Register** online via **GPO Access** may be directed to telephone (202) 512-1530 or toll free at (888) 293-6498; to fax (202) 512-1262; or to E-mail <<gpoaccess@gpo.gov>>.

This Notice is available in paper copy from the OMB Publications Office, 725 17th Street, NW, NEOB, Room 2200, Washington, D.C. 20503; telephone (202) 395-7332; fax (202) 395-6137.

FOR FURTHER INFORMATION CONTACT: Suzann Evinger, Statistical Policy Office, Office of Information and Regulatory Affairs, Office of Management and Budget, NEOB, Room 10201, 725 17th Street, N.W., Washington, D.C. 20503; telephone: (202) 395-3093; fax (202) 395-7245.

SUPPLEMENTARY INFORMATION:

A. Background

For more than 20 years, the current standards in OMB's Statistical Policy Directive No. 15 have provided a common language to promote uniformity and comparability for data on race and ethnicity for the population groups specified in the Directive. They were developed in cooperation with Federal agencies to provide consistent data on race and ethnicity throughout the Federal Government. Development of the data standards stemmed in large measure from new responsibilities to enforce civil rights laws. Data were needed to monitor equal access in housing, education, employment, and other areas, for populations that historically had experienced discrimination and differential treatment because of their race or ethnicity. The standards are used not only in the decennial census (which provides the data for the "denominator" for many measures), but also in household surveys, on administrative forms (e.g., school registration and mortgage lending applications), and in medical and other research. The categories represent a social-political construct designed for collecting data on the race and ethnicity of broad population groups in this country, and are not anthropologically or scientifically based.

B. Comprehensive Review Process

Particularly since the 1990 census, the standards have come under increasing criticism from those who believe that the minimum categories set forth in Directive No. 15 do not reflect the increasing diversity of our Nation's population that has resulted primarily from growth in immigration and in interracial marriages. In response to the criticisms, OMB announced in July 1993 that it would undertake a comprehensive review of the current categories for data on race and ethnicity.

This review has been conducted over the last four years in collaboration with the Interagency Committee for the Review of the Racial and Ethnic Standards, which OMB established in March 1994 to facilitate the participation of Federal agencies in the review. The members of the Interagency Committee, from more than 30 agencies, represent the many and diverse Federal needs for data on race and ethnicity, including statutory requirements for such data. The Interagency Committee developed the following principles to govern the

review process:

1. The racial and ethnic categories set forth in the standards should not be interpreted as being primarily biological or genetic in reference. Race and ethnicity may be thought of in terms of social and cultural characteristics as well as ancestry.

2. Respect for individual dignity should guide the processes and methods for collecting data on race and ethnicity; ideally, respondent self-identification should be facilitated to the greatest extent possible, recognizing that in some data collection systems observer identification is more practical.

3. To the extent practicable, the concepts and terminology should reflect clear and generally understood definitions that can achieve broad public acceptance. To assure they are reliable, meaningful, and understood by respondents and observers, the racial and ethnic categories set forth in the standard should be developed using appropriate scientific methodologies, including the social sciences.

4. The racial and ethnic categories should be comprehensive in coverage and produce compatible, nonduplicative, exchangeable data across Federal agencies.

5. Foremost consideration should be given to data aggregations by race and ethnicity that are useful for statistical analysis and program administration and assessment, bearing in mind that the standards are not intended to be used to establish eligibility for participation in any federal program.

6. The standards should be developed to meet, at a minimum, Federal legislative and programmatic requirements. Consideration should also be given to needs at the State and local government levels, including American Indian tribal and Alaska Native village governments, as well as to general societal needs for these data.

7. The categories should set forth a minimum standard; additional categories should be permitted provided they can be aggregated to the standard categories. The number of standard categories should be kept to a manageable size, determined by statistical concerns and data needs.

8. A revised set of categories should be operationally feasible in terms of burden placed upon respondents; public and private costs to implement the revisions should be a factor in the decision.

9. Any changes in the categories should be based on sound methodological research and should include evaluations of the impact of any changes not only on the usefulness of the resulting data but also on the comparability of any new categories with the existing ones.

10. Any revision to the categories should provide for a crosswalk at the time of adoption between the old and the new categories so that historical data series can be statistically adjusted and comparisons can be made.

11. Because of the many and varied needs and strong interdependence of Federal agencies for racial and ethnic data, any changes to the existing categories should be the product of an interagency collaborative effort.

12. Time will be allowed to phase in any new categories. Agencies will not be required to update historical records.

13. The new directive should be applicable throughout the U.S. Federal statistical system. The standard or standards must be usable for the decennial census, current surveys, and administrative records, including those using observer identification.

The principal objective of the review has been to enhance the accuracy of the demographic information collected by the Federal Government. The starting point for the review was the minimum set of categories for data on race and ethnicity that have provided information for more than 20 years for a variety of purposes, and the recognition of the importance of being able to maintain this historical continuity. The review process has had two major elements: (1) public comment on the present standards, which helped to identify concerns and provided numerous suggestions for changing the standards; and (2) research and testing related to assessing the possible effects of suggested changes on the quality and usefulness of the resulting data.

Public input, the first element of the review process, was sought through a variety of means: (1) During 1993, Congressman Thomas C. Sawyer, then Chairman of the House Subcommittee on Census, Statistics, and Postal Personnel, held four hearings that included 27 witnesses, focusing particularly on the use of the categories in the 2000 census. (2) At the request of OMB, the National Academy of Sciences' Committee on National Statistics (CNSTAT) conducted a workshop in February 1994 to articulate issues surrounding a review of the categories. The workshop included representatives of Federal agencies, academia, social science research institutions, interest groups, private industry, and a local school district. (A summary of the workshop, **Spotlight on Heterogeneity: The Federal Standards for Racial and Ethnic Classification**, is available from CNSTAT, 2101 Constitution Avenue, N.W., Washington, D.C. 20418.) (3) On June 9, 1994, OMB published a **Federal Register** (59 FR 29831-29835) Notice that contained background information on the development of the current standards and requested public comment on: the adequacy of current racial and ethnic categories; the principles that should govern any proposed revisions to the standards; and specific suggestions for change that had been offered by individuals and interested groups over a period of several years. In response, OMB received nearly 800 letters. As part of this comment period and to bring the review closer to the public, OMB also heard testimony from 94 witnesses at hearings held during July 1994 in Boston, Denver, San Francisco, and Honolulu. (4) In an August 28, 1995, **Federal Register** (60 FR 44674-44693) Notice, OMB provided an interim report on the review process, including a summary of the comments on the June 1994 **Federal Register** Notice, and offered a final opportunity for comment on the research to be conducted during 1996. (5) OMB staff have also discussed the review process with various interested groups and have made presentations at numerous meetings.

The second element of the review process involved research and testing of various proposed changes. The categories in OMB's Directive No. 15 are used not only to produce data on the demographic characteristics of the population, but also to monitor civil rights enforcement and program implementation. Research was undertaken to provide an objective assessment of the data quality issues associated with various approaches to collecting data on race and ethnicity. To that end, the Interagency Committee's Research Working Group, co-chaired by the Bureau of the Census and the Bureau of Labor Statistics, reviewed the various criticisms and suggestions for changing the current categories, and developed a research agenda for some of the more significant issues that had been identified. These issues included how to collect data on persons who identify themselves as "multiracial"; whether to combine race and Hispanic origin in one question or have separate questions on race and Hispanic origin; whether to combine the concepts of race, ethnicity, and ancestry; whether to change the terminology used for particular categories; and whether to add new categories to the current minimum set.

Because the mode of data collection can have an effect on how a person responds, the research agenda proposed studies both in surveys using in-person or telephone interviews and in self-administered

questionnaires, such as the decennial census, which are filled out by the respondent and mailed back. Cognitive interviews were conducted with various groups to provide guidance on the wording of the questions and the instructions for the tests and studies.

The research agenda included several major national tests, the results of which are discussed throughout the Interagency Committee's **Report to the Office of Management and Budget on the Review of Statistical Policy Directive No. 15**: (1) In May 1995, the Bureau of Labor Statistics (BLS) sponsored a Supplement on Race and Ethnicity to the Current Population Survey (CPS). The findings were made available in a 1996 report, **Testing Methods of Collecting Racial and Ethnic Information: Results of the Current Population Survey Supplement on Race and Ethnicity**, available from BLS, 2 Massachusetts Avenue, N.E., Room 4915, Postal Square Building, Washington, D.C. 20212, or by calling 202-606-7375. The results were also summarized in an October 26, 1995, news release, which is available electronically at <<http://stats.bls.gov/news.release/ethnic.toc.htm>>. (2) The Bureau of the Census, as part of its research for the 2000 census, tested alternative approaches to collecting data on race and ethnicity in the March 1996 National Content Survey (NCS). The Census Bureau published the results in a December 1996 report, **Findings on Questions on Race and Hispanic Origin Tested in the 1996 National Content Survey**; highlights of the report are available at <<http://www.census.gov/population/www/socdemo/96natcontentsurvey.html>>. (3) In June 1996, the Census Bureau conducted the Race and Ethnic Targeted Test (RAETT), which was designed to permit assessments of the effects of possible changes on smaller populations not reliably measured in national samples, including American Indians, Alaska Natives, detailed Asian and Pacific Islander groups (such as Chinese and Hawaiians), and detailed Hispanic groups (such as Puerto Ricans and Cubans). The Census Bureau released the results in a May 1997 report, **Results of the 1996 Race and Ethnic Targeted Test**; highlights of the report are available at <<http://www/census.gov/population/www/documentation/twps-0018.html>>. Single copies (paper) of the NCS and RAETT reports may be obtained from the Population Division, U.S. Bureau of the Census, Washington, D.C. 20233; telephone 301-457-2402.

In addition to these three major tests, the National Center for Education Statistics (NCES) and the Office for Civil Rights in the Department of Education jointly conducted a survey of 1,000 public schools to determine how schools collect data on the race and ethnicity of their students and how the administrative records containing these data are maintained to meet statutory requirements for reporting aggregate information to the Federal Government. NCES published the results in a March 1996 report, **Racial and Ethnic Classifications Used by Public Schools** (NCES 96-092). The report is available electronically at <<http://nces.ed.gov/pubs/96092.html>>. Single paper copies may be obtained from NCES, 555 New Jersey, NW, Washington, D.C. 20208-5574, or by calling 202-219-1442.

The research agenda also included studies conducted by the National Center for Health Statistics, the Office of the Assistant Secretary for Health, and the Centers for Disease Control and Prevention to evaluate the procedures used and the quality of the information on race and ethnicity in administrative records such as that reported on birth certificates and recorded on death certificates.

On July 9, 1997, OMB published a **Federal Register** Notice (62 FR 36874 - 36946) containing the Interagency Committee's **Report to the Office of Management and Budget on the Review of Statistical Policy Directive No. 15**. The Notice made available for comment the Interagency Committee's recommendations for how OMB should revise Directive No. 15. The report consists of six chapters. Chapter 1 provides a brief history of Directive No. 15, a summary of the issues considered by the Interagency Committee, a review of the research activities, and a discussion of the criteria used in conducting the evaluation. Chapter 2 discusses a number of general concerns that need to be addressed when considering any changes to the current standards. Chapters 3 through 5 report the results of the research as they bear on

the more significant suggestions OMB received for changes to Directive No. 15. Chapter 6 gives the Interagency's Committee's recommendations concerning the various suggested changes based on a review of public comments and testimony and the research results.

C. Summary of Comments Received on the Interagency Committee's Recommendations

In response to the July 9, 1997, **Federal Register** Notice, OMB received approximately 300 letters (many of them hand written) on a variety of issues, plus approximately 7000 individually signed and mailed, preprinted postcards on the issue of classifying data on Native Hawaiians, and about 500 individually signed form letters from members of the Hapa Issues Forum in support of adopting the recommendation for multiple race reporting. Some of the 300 letters focused on a single recommendation of particular interest to the writer, while other letters addressed a number of the recommendations. The preponderance of the comments were from individuals. Each comment was considered in preparing OMB's decision.

1. **Comments on Recommendations Concerning Reporting More Than One Race**

The Interagency Committee recommended that, when self-identification is used, respondents who wish to identify their mixed racial heritage should be able to mark or select more than one of the racial categories originally specified in Directive No. 15, but that there should not be a "multiracial" category. This recommendation to report multiple races was favorably received by most of those commenting on it, including associations and organizations such as the American Medical Association, the National Education Association, the National Council of La Raza, and the National Committee on Vital and Health Statistics, as well as all Federal agencies that responded. Comments from some organizations, such as the NAACP Legal Defense and Educational Fund, the Lawyers' Committee for Civil Rights Under Law, and the Equal Employment Advisory Council, were receptive to the recommendation on multiple race responses, but expressed reservations pending development of tabulation methods to ensure the utility of these data. The recommendation was also supported by many of the advocacy groups that had earlier supported a "multiracial" (box) category, such as the Association of MultiEthnic Americans and its affiliates nationwide. Several individuals wrote in support of "multiple race" reporting, basing their comments on a September 1997 article, "What Race Am I?" in **Mademoiselle** magazine, which urged its readers "to express an opinion on whether or not a 'Multiracial' category should be included in all federal record keeping, including the 2000 census." A few comments specifically favoring multiple race responses suggested that respondents should also be asked to indicate their primary racial affiliation in order to facilitate the tabulation of responses. A handful of comments on multiple race reporting suggested that individuals with both Hispanic and non-Hispanic heritages be permitted to mark or select both categories (see discussion below).

A few comments, in particular some from state agencies and legislatures, opposed any multiple race reporting because of possible increased costs to collect the information and implementation problems. Comments from the American Indian tribal governments also were opposed to the recommendation concerning reporting more than one race. A number of the comments that supported multiple race responses also expressed concern about the cost and burden of collecting the information to meet Federal reporting requirements, the schedule for implementation, and how the data would be tabulated to meet the requirements of legislative redistricting and enforcement of the Voting Rights Act. A few comments expressed support for categories called "human," or "American"; several proposed that there be no collection of data on race.

2. **Comments on Recommendation for Classification of Data on Native Hawaiians**

The Interagency Committee recommended that data on Native Hawaiians continue to be classified in the

Asian or Pacific Islander category. This recommendation was opposed by the Hawaiian congressional delegation, the 7,000 individuals who signed and sent preprinted yellow postcards, the State of Hawaii departments and legislature, Hawaiian organizations, and other individuals who commented on this recommendation. Instead, the comments from these individuals supported reclassifying Native Hawaiians in the American Indian or Alaska Native category, which they view as an "indigenous peoples" category (although this category has not been considered or portrayed in this manner in the standards). Native Hawaiians, as the descendants of the original inhabitants of what is now the State of Hawaii, believe that as indigenous people they should be classified in the same category as American Indians and Alaska Natives. On the other hand, the American Indian tribal governments have opposed such a reclassification, primarily because they view the data obtained from that category as being essential for administering Federal programs for American Indians. Comments from the Native Hawaiians also noted the Asian or Pacific Islander category provides inadequate data for monitoring the social and economic conditions of Native Hawaiians and other Pacific Islander groups. Because the Interagency Committee had recommended against adding categories to the minimum set of categories, requesting a separate category for Native Hawaiians was not viewed as an option by those who commented.

3. **Comments on Recommendation Concerning Classification of Data on Central and South American Indians**

The Interagency Committee recommended that data for Central and South American Indians be included in the American Indian or Alaska Native category. Several comments from the American Indian community opposed this recommendation. Moreover, comments from some Native Hawaiians pointed out what they believed to be an inconsistency in the Interagency Committee's recommendation to include in the American Indian or Alaska Native category descendants of Central and South American Indians -- persons who are not original peoples of the United States -- if Native Hawaiians were not to be included.

4. **Comments on Recommendation Not to Add an Arab or Middle Eastern Ethnic Category**

The Interagency Committee recommended that an Arab or Middle Eastern ethnic category should not be added to the minimum standards for all reporting of Federal data on race and ethnicity. Several comments were received in support of having a separate category in order to have data viewed as necessary to monitor discrimination against this population.

5. **Comments on Recommendations for Terminology**

Comments on terminology largely supported the Interagency Committee's recommendations to retain the term "American Indian," to change "Hawaiian" to "Native Hawaiian," and to change "Black" to "Black or African American." There were a few requests to include "Latino" in the category name for the Hispanic population.

D. OMB's Decisions

This section of the Notice provides information on the decisions taken by OMB on the recommendations that were proposed by the Interagency Committee. The Committee's recommendations addressed options for reporting by respondents, formats of questions, and several aspects of specific categories, including possible additions, revised terminology, and changes in definitions. In reviewing OMB's decisions on the recommendations for collecting data on race and ethnicity, it is useful to remember that these decisions:

retain the concept that the standards provide a **minimum** set of categories for data on race and ethnicity;

permit the collection of more detailed information on population groups provided that any additional categories can be aggregated into the minimum standard set of categories;

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);

do **not** identify or designate certain population groups as "minority groups";

continue the policy that the categories are **not** to be used for determining the eligibility of population groups for participation in any Federal programs;

do **not** establish criteria or qualifications (such as blood quantum levels) that are to be used in determining a particular individual's racial or ethnic classification; and

do **not** tell an individual who he or she is, or specify how an individual should classify himself or herself.

In arriving at its decisions, OMB took into account not only the public comment on the recommendations published in the **Federal Register** on July 9, 1997, but also the considerable amount of information provided during the four years of this review process, including public comments gathered from hearings and responses to two earlier OMB Notices (on June 9, 1994, and August 28, 1995). The OMB decisions benefited greatly from the participation of the public that served as a constant reminder that there are real people represented by the data on race and ethnicity and that this is for many a deeply personal issue. In addition, the OMB decisions benefited from the results of the research and testing on how individuals identify themselves that was undertaken as part of this review process. This research, including several national tests of alternative approaches to collecting data on race and ethnicity, was developed and conducted by the professional statisticians and analysts at several Federal agencies. They are to be commended for their perseverance, dedication, and professional commitment to this challenging project.

OMB also considered in reaching its decisions the extent to which the recommendations were consistent with the set of principles (see Section B of the Supplementary Information) developed by the Interagency Committee to guide the review of this sensitive and substantively complex issue. OMB believes that the Interagency Committee's recommendations took into account the principles and achieved a reasonable balance with respect to statistical issues, data needs, social concerns, and the personal dimensions of racial and ethnic identification. OMB also finds that the Committee's recommendations are consistent with the principal objective of the review, which is to enhance the accuracy of the demographic information collected by the Federal Government by having categories for data on race and ethnicity that will enable the capture of information about the increasing diversity of our Nation's population while at the same time respecting each individual's dignity.

As indicated in detail below, OMB accepts the Interagency Committee's recommendations concerning reporting more than one race, including the recommendation that there be no category called "multiracial," the formats and sequencing of the questions on race and Hispanic origin, and most of the changes to terminology.

OMB does not accept the Interagency Committee's recommendations concerning the classification of data on the Native Hawaiian population and the terminology for Hispanics, and it has instead decided to make the

changes that follow.

Native Hawaiian classification.--OMB does not accept the recommendation concerning the continued classification of Hawaiians in the Asian or Pacific Islander category. **Instead, OMB has decided to break apart the Asian or Pacific Islander category into two categories -- one called "Asian" and the other called "Native Hawaiian or Other Pacific Islander."** As a result, there will be five categories in the minimum set for data on race.

The "Native Hawaiian or Other Pacific Islander" category will be defined as "A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands." (The term "Native Hawaiian" does not include individuals who are native to the State of Hawaii by virtue of being born there.) In addition to Native Hawaiians, Guamanians, and Samoans, this category would include the following Pacific Islander groups reported in the 1990 census: Carolinian, Fijian, Kosraean, Melanesian, Micronesian, Northern Mariana Islander, Palauan, Papua New Guinean, Ponapean (Pohnpelan), Polynesian, Solomon Islander, Tahitian, Tarawa Islander, Tokelauan, Tongan, Trukese (Chuukese), and Yapese.

The "Asian" category will be defined as "A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam."

The Native Hawaiians presented compelling arguments that the standards must facilitate the production of data to describe their social and economic situation and to monitor discrimination against Native Hawaiians in housing, education, employment, and other areas. Under the current standards for data on race and ethnicity, Native Hawaiians comprise about three percent of the Asian and Pacific Islander population. By creating separate categories, the data on the Native Hawaiians and other Pacific Islander groups will no longer be overwhelmed by the aggregate data of the much larger Asian groups. Native Hawaiians will comprise about 60 percent of the new category.

The Asian, Native Hawaiian, and Pacific Islander population groups are well defined; moreover, there has been experience with reporting in separate categories for the Native Hawaiian and Pacific Islander population groups. The 1990 census included "Hawaiian," "Samoan," and "Guamanian" as response categories to the race question. In addition, two of the major tests conducted as part of the current review (the NCS and the RAETT) used "Hawaiian" and/or "Native Hawaiian," "Samoan," "Guamanian," and "Guamanian or Chamorro" as response options to the race question. These factors facilitate breaking apart the current category.

Terminology for Hispanics.--OMB does not accept the recommendation to retain the single term "Hispanic." **Instead, OMB has decided that the term should be "Hispanic or Latino."** Because regional usage of the terms differs -- Hispanic is commonly used in the eastern portion of the United States, whereas Latino is commonly used in the western portion -- this change may contribute to improved response rates.

The OMB decisions on the Interagency Committee's specific recommendations are presented below:

(1) OMB accepts the following recommendations concerning reporting more than one race:

When self-identification is used, a method for reporting more than one race should be adopted.

The method for respondents to report more than one race should take the form of multiple responses

to a single question and not a "multiracial" category.

When a list of races is provided to respondents, the list should not contain a "multiracial" category.

Based on research conducted so far, two recommended forms for the instruction accompanying the multiple response question are "Mark one or more ..." and "Select one or more...."

If the criteria for data quality and confidentiality are met, provision should be made to report, at a minimum, the number of individuals identifying with more than one race. Data producers are encouraged to provide greater detail about the distribution of multiple responses.

The new standards will be used in the decennial census, and other data producers should conform as soon as possible, but not later than January 1, 2003.

(2) OMB accepts the following recommendations concerning a combined race and Hispanic ethnicity question:

When self-identification is used, the two question format should be used, with the race question allowing the reporting of more than one race.

When self-identification is not feasible or appropriate, a combined question can be used and should include a separate Hispanic category co-equal with the other categories.

When the combined question is used, an attempt should be made, when appropriate, to record ethnicity and race or multiple races, but the option to indicate only one category is acceptable.

(3) OMB accepts the following recommendations concerning the retention of both reporting formats:

The two question format should be used in all cases involving self-identification.

The current combined question format should be changed and replaced with a new format which includes a co-equal Hispanic category for use, if necessary, in observer identification.

(4) OMB accepts the following recommendation concerning the ordering of the Hispanic origin and race questions:

When the two question format is used, the Hispanic origin question should precede the race question.

(5) OMB accepts the following recommendation concerning adding Cape Verdean as an ethnic category:

A Cape Verdean ethnic category should not be added to the minimum data collection standards.

(6) OMB accepts the following recommendation concerning the addition of an Arab or Middle Eastern ethnic category:

An Arab or Middle Eastern ethnic category should not be added to the minimum data standards.

(7) OMB interprets the recommendation not to add any other categories to mean the expansion of the minimum set to include new population groups. The OMB decision to break apart the "Asian or Pacific Islander" category does not create a category for a new population group.

(8) OMB accepts the following recommendation concerning changing the term "American Indian" to "Native American":

The term American Indian should not be changed to Native American.

(9) OMB accepts the following recommendation concerning changing the term "Hawaiian" to "Native Hawaiian":

The term "Hawaiian" should be changed to "Native Hawaiian."

(10) OMB does not accept the recommendation concerning the continued classification of Native Hawaiians in the Asian or Pacific Islander category.

OMB has decided to break apart the Asian or Pacific Islander category into two categories -- one called "Asian" and the other called "Native Hawaiian or Other Pacific Islander." As a result, there are five categories in the minimum set for data on race.

The "Native Hawaiian or Other Pacific Islander" category is defined as "A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands."

The "Asian" category is defined as "A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam."

(11) OMB accepts the following recommendations concerning the use of "Alaska Native" instead of "Eskimo" and "Aleut":

"Alaska Native" should replace the term "Alaskan Native."

Alaska Native should be used instead of Eskimo and Aleut.

The Alaska Native response option should be accompanied by a request for tribal affiliation when possible.

(12) OMB accepts the following recommendations concerning the classification of Central and South American Indians:

Central and South American Indians should be classified as American Indian.

The definition of the "American Indian or Alaska Native" category should be modified to include the original peoples from Central and South America.

In addition, OMB has decided to make the definition for the American Indian or Alaska Native

category more consistent with the definitions of the other categories.

(13) OMB accepts the following recommendations concerning the term or terms to be used for the name of the Black category:

The name of the Black category should be changed to "Black or African American."

The category definition should remain unchanged.

Additional terms, such as Haitian or Negro, can be used if desired.

(14) OMB decided to modify the recommendations concerning the term or terms to be used for Hispanic:

The term used should be "Hispanic or Latino."

The definition of the category should remain unchanged.

In addition, the term "Spanish Origin," can be used if desired.

Accordingly, the Office of Management and Budget adopts and issues the revised minimum standards for Federal data on race and ethnicity for major population groups in the United States which are set forth at the end of this Notice.

Topics for further research

There are two areas where OMB accepts the Interagency Committee's recommendations but believes that further research is needed: (1) multiple responses to the Hispanic origin question and (2) an ethnic category for Arabs/Middle Easterners.

Multiple Responses to the Hispanic Origin Question.--The Interagency Committee recommended that respondents to Federal data collections should be permitted to report more than one race. During the most recent public comment process, a few comments suggested that the concept of "marking more than one box" should be extended to the Hispanic origin question. Respondents are now asked to indicate if they are "of Hispanic origin" **or** "not of Hispanic origin." Allowing individuals to select more than one response to the ethnicity question would provide the opportunity to indicate ethnic heritage that is both Hispanic and non-Hispanic.

The term "Hispanic" refers to persons who trace their origin or descent to Mexico, Puerto Rico, Cuba, Central and South America, and other Spanish cultures. While there has been considerable public concern about the need to review Directive No. 15 with respect to classifying individuals of mixed racial heritage, there has been little comment on reporting both an Hispanic and a non-Hispanic origin. On many Federal forms, Hispanics can also express a racial identity on a separate race question. In the decennial census, individuals who consider themselves part Hispanic can also indicate additional heritages in the ancestry question.

On one hand, it can be argued that allowing individuals to mark both categories in the Hispanic origin question would parallel the instruction "to mark (or select) one or more" racial categories. Individuals would

not have to choose between their parents' ethnic heritages, and movement toward an increasingly diverse society would be recognized.

On the other hand, because the matter of multiple responses to the Hispanic ethnicity question was not raised in the early phases of the public comment process, no explicit provisions were made for testing this approach in the research conducted to inform the review of Directive No. 15. While a considerable amount of research was focused on how to improve the response rate to the Hispanic origin question, it is unclear whether and to what extent explicitly permitting multiple responses to the Hispanic origin question would affect nonresponse to the race question or hamper obtaining more detailed data on Hispanic population groups.

Information on the possible impact of any changes on the quality of the data has been an essential element of the review. While the effects of changes in the Hispanic origin question are unknown, they could conceivably be substantial. Thus, OMB has decided **not** to include a provision in the standards that would explicitly permit respondents to select both "Hispanic origin" and "Not of Hispanic Origin" options. OMB believes that this is an item for future research. In the meantime, the ancestry question on the decennial census long form does provide respondents who consider themselves part Hispanic to write in additional heritages.

Research on an Arab/Middle Easterner category.--During the public comment process, OMB received a number of requests to add an ethnic category for Arabs/Middle Easterners so that data could be obtained that could be useful in monitoring discrimination. The public comment process indicated, however, that there was no agreement on a definition for this category. The combined race, Hispanic origin, and ancestry question in the RAETT, which was designed to address requests that were received from groups for establishing separate categories, did not provide a solution.

While OMB accepted the Interagency's Committee recommendation not to create a new category for this population group, OMB believes that further research should be done to determine the best way to improve data on this population group. Meanwhile, the write-ins to the ancestry question on the decennial census long form will continue to provide information on the number of individuals who identify their heritage as Arab or Middle Easterner.

E. Tabulation Issues

The revised standards retain the concept of a minimum set of categories for Federal data on race and ethnicity and make possible at the same time the collection of data to reflect the diversity of our Nation's population. Since the Interagency Committee's recommendation concerning the reporting of more than one race was made available for public comment, the focus of attention has been largely on how the data would be tabulated. Because of the concerns expressed about tabulation methods and our own view of the importance of this issue, OMB committed to accelerate the work on tabulation issues when it testified in July 1997 on the Interagency Committee's recommendations.

A group of statistical and policy analysts drawn from the Federal agencies that generate or use these data has spent the past few months considering the tabulation issues. Although this work is still in its early stages, some preliminary guidance can be shared at this time. In general, OMB believes that, consistent with criteria for confidentiality and data quality, the tabulation procedures used by the agencies should result in the production of as much detailed information on race and ethnicity as possible.

Guidelines for tabulation ultimately must meet the needs of at least two groups within the Federal Government, with the overriding objective of providing the most accurate and informative body of data. The

first group is composed of those government officials charged with carrying out constitutional and legislative mandates, such as redistricting legislatures, enforcing civil rights laws, and monitoring progress in anti-discrimination programs. (The legislative redistricting file produced by the Bureau of the Census, also known as the Public Law 94-171 file, is an example of a file meeting such legislative needs.) The second group consists of the staff of statistical agencies producing and analyzing data that are used to monitor economic and social conditions and trends.

Many of the needs of the first group can be met with an initial tabulation that provides, consistent with standards for data quality and confidentiality, the full detail of racial reporting; that is, the number of people reporting in each single race category and the number reporting each of the possible combinations of races, which would add to the total population. Depending on the judgment of users, the combinations of multiple responses could be collapsed. One method would be to provide separate totals for those reporting in the most common multiple race combinations and to collapse the data for other less frequently reported combinations. The specifics of the collapsed distributions must await the results of particular data collections. A second method would be to report the total selecting each particular race, whether alone or in combination with other races. These totals would represent upper bounds on the size of the populations who identified with each of the racial categories. In some cases, this latter method could be used for comparing data collected under the old standards with data collected under the new standards. It is important that users with the same or closely related responsibilities adopt the same tabulation method. Regardless of the method chosen for collapsing multiple race responses, the total number reporting more than one race must be made available, if confidentiality and data quality requirements can be met, in order to ensure that any changes in response patterns resulting from the new standards can be monitored over time.

Meeting the needs of the second group (those producing and analyzing statistical data to monitor economic and social conditions and trends), as well as some additional needs of the first group, may require different tabulation procedures. More research must be completed before guidelines that will meet the requirements of these users can be developed. A group of statistical and policy experts will review a number of alternative procedures and provide recommendations to OMB concerning these tabulation requirements by Spring 1998. Four of the areas in which further exploration is needed are outlined below.

Equal employment opportunity and other anti-discrimination programs have traditionally provided the numbers of people in the population by selected characteristics, including racial categories, for business, academic, and government organizations to use in evaluating conformance with program objectives. Because of the potentially large number of categories that may result from application of the new standards, many with very small numbers, it is not clear how this need for data will be best satisfied in the future.

The numbers of people in distinct groups based on decennial census results are used in developing sample designs and survey controls for major demographic surveys. For example, the National Health Interview Survey uses census data to increase samples for certain population groups, adjust for survey non-response, and provide weights for estimating health outcomes at the national level. The impact of having data for many small population groups with multiple racial heritages must be explored.

Vital statistics data include birth and death rates for various population groups. Typically the numerator (number of births or deaths) is derived from administrative records, while the denominator comes from intercensal population estimates. Birth certificate data on race are likely to have been self reported by the mother. Over time, these data may become comparable to data collected under the new standards. Death certificate data, however, frequently are filled out by an observer, such as a mortician, physician, or funeral director. These data, particularly for the population with multiple racial heritages, are likely to be quite different from the information obtained when respondents report about themselves. Research to define

comparable categories to be used in both numerators and denominators is needed to assure that vital statistics are as accurate and useful as possible.

More generally, statistical indicators are often used to measure change over time. Procedures that will permit meaningful comparisons of data collected under the previous standards with those that will be collected under the new standards need to be developed.

The methodology for tabulating data on race and ethnicity must be carefully developed and coordinated among the statistical agencies and other Federal data users. Moreover, just as OMB's review and decision processes have benefited during the past four years from extensive public participation, we expect to discuss tabulation methods with data users within and outside the Federal Government. OMB expects to issue additional guidance with respect to tabulating data on race and ethnicity by Fall 1998.

Sally Katzen

Administrator, Office of Information and Regulatory Affairs.

Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity

This classification provides a minimum standard for maintaining, collecting, and presenting data on race and ethnicity for all Federal reporting purposes. The categories in this classification are social-political constructs and should not be interpreted as being scientific or anthropological in nature. They are not to be used as determinants of eligibility for participation in any Federal program. The standards have been developed to provide a common language for uniformity and comparability in the collection and use of data on race and ethnicity by Federal agencies.

The standards have five categories for data on race: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White. There are two categories for data on ethnicity: "Hispanic or Latino," and "Not Hispanic or Latino."

1. Categories and Definitions

The minimum categories for data on race and ethnicity for Federal statistics, program administrative reporting, and civil rights compliance reporting are defined as follows:

-- **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 tribal affiliation or community attachment.

-- **Asian.** A person having origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent including, for example, Cambodia, China, India, Japan, Korea, Malaysia, Pakistan, the Philippine Islands, Thailand, and Vietnam.

-- **Black or African American.** A person having origins in any of the black racial groups of Africa. Terms such as "Haitian" or "Negro" can be used in addition to "Black or African American."

-- **Hispanic or Latino.** A person of Cuban, Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race. The term, "Spanish origin," can be used in addition to "Hispanic or Latino."

-- **Native Hawaiian or Other Pacific Islander.** A person having origins in any of the original peoples of Hawaii, Guam, Samoa, or other Pacific Islands.

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

Respondents shall be offered the option of selecting one or more racial designations. Recommended forms for the instruction accompanying the multiple response question are "Mark one or more" and "Select one or more."

2. Data Formats

The standards provide two formats that may be used for data on race and ethnicity. Self-reporting or self-identification using two separate questions is the preferred method for collecting data on race and ethnicity. In situations where self-reporting is not practicable or feasible, the combined format may be used.

In no case shall the provisions of the standards be construed to limit the collection of data to the categories described above. The collection of greater detail is encouraged; however, any collection that 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.

With respect to tabulation, the procedures used by Federal agencies shall result in the production of as much detailed information on race and ethnicity as possible. However, Federal agencies shall not present data on detailed categories if doing so would compromise data quality or confidentiality standards.

a. Two-question format

To provide flexibility and ensure data quality, separate questions shall be used wherever feasible for reporting race and ethnicity. When race and ethnicity are collected separately, ethnicity shall be collected first. If race and ethnicity are collected separately, the minimum designations are:

Race:

-- American Indian or Alaska Native

-- Asian

-- Black or African American

-- Native Hawaiian or Other Pacific Islander

-- White

Ethnicity:

-- Hispanic or Latino

-- Not Hispanic or Latino

When data on race and ethnicity are collected separately, provision shall be made to report the number of respondents in each racial category who are Hispanic or Latino.

When aggregate data are presented, data producers shall provide the number of respondents who marked (or selected) only one category, separately for each of the five racial categories. In addition to these numbers, data producers are strongly encouraged to provide the detailed distributions, including all possible combinations, of multiple responses to the race question. If data on multiple responses are collapsed, at a minimum the total number of respondents reporting "more than one race" shall be made available.

b. Combined format

The combined format may be used, if necessary, for observer-collected data on race and ethnicity. Both race (including multiple responses) and ethnicity shall be collected when appropriate and feasible, although the selection of one category in the combined format is acceptable. If a combined format is used, there are six minimum categories:

-- American Indian or Alaska Native

-- Asian

-- Black or African American

-- Hispanic or Latino

-- Native Hawaiian or Other Pacific Islander

-- White

When aggregate data are presented, data producers shall provide the number of respondents who marked (or selected) only one category, separately for each of the six categories. In addition to these numbers, data producers are strongly encouraged to provide the detailed distributions, including all possible combinations, of multiple responses. In cases where data on multiple responses are collapsed, the total number of respondents reporting "Hispanic or Latino and one or more races" and the total number of respondents reporting "more than one race" (regardless of ethnicity) shall be provided.

3. Use of the Standards for Record Keeping and Reporting

The minimum standard categories shall be used for reporting as follows:

a. Statistical reporting

These standards shall be used at a minimum for all federally sponsored statistical data collections that include data on race and/or ethnicity, except when the collection involves a sample of such size that the data

on the smaller categories would be unreliable, or when the collection effort focuses on a specific racial or ethnic group. Any other variation will have to be specifically authorized by the Office of Management and Budget (OMB) through the information collection clearance process. In those cases where the data collection is not subject to the information collection clearance process, a direct request for a variance shall be made to OMB.

b. General program administrative and grant reporting

These standards shall be used for all Federal administrative reporting or record keeping requirements that include data on race and ethnicity. Agencies that cannot follow these standards must request a variance from OMB. Variances will be considered if the agency can demonstrate that it is not reasonable for the primary reporter to determine racial or ethnic background in terms of the specified categories, that determination of racial or ethnic background is not critical to the administration of the program in question, or that the specific program is directed to only one or a limited number of racial or ethnic groups.

c. Civil rights and other compliance reporting

These standards shall be used by all Federal agencies in either the separate or combined format for civil rights and other compliance reporting from the public and private sectors and all levels of government. Any variation requiring less detailed data or data which cannot be aggregated into the basic categories must be specifically approved by OMB for executive agencies. More detailed reporting which can be aggregated to the basic categories may be used at the agencies' discretion.

4. Presentation of Data on Race and Ethnicity

Displays of statistical, administrative, and compliance data on race and ethnicity shall use the categories listed above. The term "nonwhite" is not acceptable for use in the presentation of Federal Government data. It shall not be used in any publication or in the text of any report.

In cases where the standard categories are considered inappropriate for presentation of data on particular programs or for particular regional areas, the sponsoring agency may use:

a. The designations "Black or African American and Other Races" or "All Other Races" as collective descriptions of minority races when the most summary distinction between the majority and minority races is appropriate;

b. The designations "White," "Black or African American," and "All Other Races" when the distinction among the majority race, the principal minority race, and other races is appropriate; or

c. The designation of a particular minority race or races, and the inclusion of "Whites" with "All Other Races" when such a collective description is appropriate.

In displaying detailed information that represents a combination of race and ethnicity, the description of the data being displayed shall clearly indicate that both bases of classification are being used.

When the primary focus of a report is on two or more specific identifiable groups in the population, one or more of which is racial or ethnic, it is acceptable to display data for each of the particular groups separately and to describe data relating to the remainder of the population by an appropriate collective description.

Appendix II. Workshop Participants

Robert N. Anderson, Ph.D.
Statistician
Division of Vital Statistics
National Center for Health Statistics

Frank D. Bean, Ph.D.
Professor
Department of Sociology
University of California, Irvine

Lester R. Curtin, Ph.D.
Acting Associate Director for Research and Methodology
National Center for Health Statistics

Jane S. Durch, M.A.
The George Washington University School of Public Health and Health Services

Irma T. Elo, Ph.D.
Assistant Professor, Department of Sociology
Research Associate, Population Studies Center
University of Pennsylvania

Suzann K. Evinger, M.A.
Policy Analyst
Office of Management and Budget

Reynolds Farley, Ph.D.
Population Studies Center
Institute for Social Research
University of Michigan

Linda Gage
California State Demographer
Demographic Research Unit
California Department of Finance

Nancy M. Gordon, Ph.D.
Associate Director for Demographic Programs
U.S. Census Bureau

Charles Hirschman, Ph.D.
Boeing International Professor
Department of Sociology
University of Washington

Frederick W. Hollmann, Ph.D.
Demographer
Population Projections Branch
Population Division
U.S. Census Bureau

Nancy Kirkendall, Ph.D.
Senior Mathematical Statistician
Energy Information Administration
Department of Energy

John F. Long, Ph.D.
Chief, Population Division
U.S. Census Bureau

Jennifer H. Madans, Ph.D. Associate Director for Science
National Center for Health Statistics

C. Jane McKendry, M.Ed.
Chief, Vital Statistics Section
Center for Health Statistics
California Department of Health Services

Stella U. Ogunwole, Ph.D.
Demographer
Racial Statistics Branch
Population Division
U.S. Census Bureau

Alvin T. Onaka, Ph.D.
State Registrar and Chief, Office of Health Status Monitoring
Hawaii Department of Health

Jennifer D. Parker, Ph.D.
Senior Staff Fellow
Infant and Child Health Studies Branch
National Center for Health Statistics

Harry M. Rosenberg, Ph.D.
Special Assistant to the Director
Division of Vital Statistics
National Center for Health Statistics

Susan Schechter, M.A.
Senior Statistician
Office of Management and Budget

Nat Schenker, Ph.D.
Senior Scientist
Office of Research and Methodology
National Center for Health Statistics

Robert Schoen, Ph.D.
Hoffman Professor of Family Sociology and Demography
Department of Sociology, and Population Research Institute
Pennsylvania State University

Stanley K. Smith, Ph.D.
Professor and Program Director
Bureau of Economic & Business Research
University of Florida

Michael A. Stoto, Ph.D.
Professor and Chair
Department of Epidemiology and Biostatistics
The George Washington University School of Public Health and Health Services

Stephanie J. Ventura, M.A.
Chief, Reproductive Statistics Branch
Division of Vital Statistics
National Center for Health Statistics

Katherine K. Wallman
Chief Statistician
Office of Management and Budget

Signe I. Wetrogan
Assistant Chief for Population Estimates and Projections
Population Division
U.S. Census Bureau

Appendix III. Questions on Race and Hispanic Origin

→ **NOTE: Please answer BOTH Questions 5 and 6.**

5. Is this person Spanish/Hispanic/Latino? *Mark* ☒ *the* ***"No"*** *box if* ***not*** *Spanish/Hispanic/Latino.*

- ☐ **No,** not Spanish/Hispanic/Latino
- ☐ Yes, Puerto Rican
- ☐ Yes, Mexican, Mexican Am., Chicano
- ☐ Yes, Cuban
- ☐ Yes, other Spanish/Hispanic/Latino — *Print group.*

6. What is this person's race? ***Mark*** ☒ ***one or more races*** *to indicate what this person considers himself/herself to be.*

- ☐ White
- ☐ Black, African Am., or Negro
- ☐ American Indian or Alaska Native — *Print name of enrolled or principal tribe.*

- ☐ Asian Indian
- ☐ Japanese
- ☐ Native Hawaiian
- ☐ Chinese
- ☐ Korean
- ☐ Guamanian or Chamorro
- ☐ Filipino
- ☐ Vietnamese
- ☐ Samoan
- ☐ Other Asian — *Print race.*
- ☐ Other Pacific Islander — *Print race.*

- ☐ Some other race — *Print race.*

Source: U.S. Census Bureau, Census 2000 questionnaire.

Figure I. Questions on race and Hispanic origin from Census 2000

25. OF HISPANIC ORIGIN? (Specify No or Yes—If yes, specify Cuban, Mexican, Puerto Rican, etc.)	26. RACE—American Indian, Black, White, etc. *(Specify below)*

Figure II. Questions on Hispanic origin and race on the current U.S. standard certificate of live birth

21. MOTHER OF HISPANIC ORIGIN? (Check the box that best describes whether the mother is Spanish/Hispanic/Latina. Check the "No" box if mother is not Spanish/Hispanic/Latina) ☐ No, not Spanish/Hispanic/Latina ☐ Yes, Mexican, Mexican American, Chicana ☐ Yes, Puerto Rican ☐ Yes, Cuban ☐ Yes, other Spanish/Hispanic/Latina (Specify)_____	22. MOTHER'S RACE (Check one or more races to indicate what the mother considers herself to be) ☐ White ☐ Black or African American ☐ American Indian or Alaska Native (Name of the enrolled or principal tribe)_____ ☐ Asian Indian ☐ Chinese ☐ Filipino ☐ Japanese ☐ Korean ☐ Vietnamese ☐ Other Asian (Specify)_____ ☐ Native Hawaiian ☐ Guamanian or Chamorro ☐ Samoan ☐ Other Pacific Islander (Specify)_____ ☐ Other (Specify)_____
24. FATHER OF HISPANIC ORIGIN? (Check the box that best describes whether the father is Spanish/Hispanic/Latino. Check the "No" box if father is not Spanish/Hispanic/Latino) ☐ No, not Spanish/Hispanic/Latino ☐ Yes, Mexican, Mexican American, Chicano ☐ Yes, Puerto Rican ☐ Yes, Cuban ☐ Yes, other Spanish/Hispanic/Latino (Specify)_____	25. FATHER'S RACE (Check one or more races to indicate what the father considers himself to be) ☐ White ☐ Black or African American ☐ American Indian or Alaska Native (Name of the enrolled or principal tribe)_____ ☐ Asian Indian ☐ Chinese ☐ Filipino ☐ Japanese ☐ Korean ☐ Vietnamese ☐ Other Asian (Specify)_____ ☐ Native Hawaiian ☐ Guamanian or Chamorro ☐ Samoan ☐ Other Pacific Islander (Specify)_____ ☐ Other (Specify)_____

Figure III. Questions on Hispanic origin and race on the draft revised U.S. standard certificate of live birth

14. WAS DECEDENT OF HISPANIC ORIGIN? (Specify No or Yes – If yes, specify Cuban, Mexican, Puerto Rican, etc.) ☐ No ☐ Yes *Specify*	15. RACE – American Indian, Black, White, etc. *(Specify)*

Figure IV. Questions on Hispanic origin and race on the current U.S. standard certificate of death

52. DECEDENT OF HISPANIC ORIGIN? Check the box that best describes whether the decedent is Spanish/Hispanic/Latino. Check the "No" box if decedent is not Spanish/Hispanic/Latino.	53. DECEDENT'S RACE (Check one or more races to indicate what the decedent considers himself or herself to be)
☐ No, not Spanish/Hispanic/Latino	☐ White
☐ Yes, Mexican, Mexican American, Chicano	☐ Black or African American
☐ Yes, Puerto Rican	☐ American Indian or Alaska Native (Name of the enrolled or principal tribe) _____
☐ Yes, Cuban	☐ Asian Indian
☐ Yes, other Spanish/Hispanic/Latino	☐ Chinese
Specify _____	☐ Filipino
	☐ Japanese
	☐ Korean
	☐ Vietnamese
	☐ Other Asian-Specify_____
	☐ Native Hawaiian
	☐ Guamanian or Chamorro
	☐ Samoan
	☐ Other Pacific Islander-specify_____
	☐ Other (specify)_____

Figure V. Questions on Hispanic origin and race on the draft revised U.S. standard certificate of death

Vital and Health Statistics series descriptions

SERIES 1. **Programs and Collection Procedures**—These reports describe the data collection programs of the National Center for Health Statistics. They include descriptions of the methods used to collect and process the data, definitions, and other material necessary for understanding the data.

SERIES 2. **Data Evaluation and Methods Research**—These reports are studies of new statistical methods and include analytical techniques, objective evaluations of reliability of collected data, and contributions to statistical theory. These studies also include experimental tests of new survey methods and comparisons of U.S. methodology with those of other countries.

SERIES 3. **Analytical and Epidemiological Studies**—These reports present analytical or interpretive studies based on vital and health statistics. These reports carry the analyses further than the expository types of reports in the other series.

SERIES 4. **Documents and Committee Reports**—These are final reports of major committees concerned with vital and health statistics and documents such as recommended model vital registration laws and revised birth and death certificates.

SERIES 5. **International Vital and Health Statistics Reports**—These reports are analytical or descriptive reports that compare U.S. vital and health statistics with those of other countries or present other international data of relevance to the health statistics system of the United States.

SERIES 6. **Cognition and Survey Measurement**—These reports are from the National Laboratory for Collaborative Research in Cognition and Survey Measurement. They use methods of cognitive science to design, evaluate, and test survey instruments.

SERIES 10. **Data From the National Health Interview Survey**—These reports contain statistics on illness; unintentional injuries; disability; use of hospital, medical, and other health services; and a wide range of special current health topics covering many aspects of health behaviors, health status, and health care utilization. They are based on data collected in a continuing national household interview survey.

SERIES 11. **Data From the National Health Examination Survey, the National Health and Nutrition Examination Surveys, and the Hispanic Health and Nutrition Examination Survey**—Data from direct examination, testing, and measurement on representative samples of the civilian noninstitutionalized population provide the basis for (1) medically defined total prevalence of specific diseases or conditions in the United States and the distributions of the population with respect to physical, physiological, and psychological characteristics, and (2) analyses of trends and relationships among various measurements and between survey periods.

SERIES 12. **Data From the Institutionalized Population Surveys**—Discontinued in 1975. Reports from these surveys are included in Series 13.

SERIES 13. **Data From the National Health Care Survey**—These reports contain statistics on health resources and the public's use of health care resources including ambulatory, hospital, and long-term care services based on data collected directly from health care providers and provider records.

SERIES 14. **Data on Health Resources: Manpower and Facilities**—Discontinued in 1990. Reports on the numbers, geographic distribution, and characteristics of health resources are now included in Series 13.

SERIES 15. **Data From Special Surveys**—These reports contain statistics on health and health-related topics collected in special surveys that are not part of the continuing data systems of the National Center for Health Statistics.

SERIES 16. **Compilations of Advance Data From Vital and Health Statistics**—Advance Data Reports provide early release of information from the National Center for Health Statistics' health and demographic surveys. They are compiled in the order in which they are published. Some of these releases may be followed by detailed reports in Series 10–13.

SERIES 20. **Data on Mortality**—These reports contain statistics on mortality that are not included in regular, annual, or monthly reports. Special analyses by cause of death, age, other demographic variables, and geographic and trend analyses are included.

SERIES 21. **Data on Natality, Marriage, and Divorce**—These reports contain statistics on natality, marriage, and divorce that are not included in regular, annual, or monthly reports. Special analyses by health and demographic variables and geographic and trend analyses are included.

SERIES 22. **Data From the National Mortality and Natality Surveys**—Discontinued in 1975. Reports from these sample surveys, based on vital records, are now published in Series 20 or 21.

SERIES 23. **Data From the National Survey of Family Growth**—These reports contain statistics on factors that affect birth rates, including contraception, infertility, cohabitation, marriage, divorce, and remarriage; adoption; use of medical care for family planning and infertility; and related maternal and infant health topics. These statistics are based on national surveys of women of childbearing age.

SERIES 24. **Compilations of Data on Natality, Mortality, Marriage, Divorce, and Induced Terminations of Pregnancy**—These include advance reports of births, deaths, marriages, and divorces based on final data from the National Vital Statistics System that were published as supplements to the *Monthly Vital Statistics Report* (MVSR). These reports provide highlights and summaries of detailed data subsequently published in *Vital Statistics of the United States*. Other supplements to the MVSR published here provide selected findings based on final data from the National Vital Statistics System and may be followed by detailed reports in Series 20 or 21.

For answers to questions about this report or for a list of reports published in these series, contact:

Data Dissemination Branch
National Center for Health Statistics
Centers for Disease Control and Prevention
6525 Belcrest Road, Room 1064
Hyattsville, MD 20782-2003

(301) 458–4636
E-mail: nchsquery@cdc.gov
Internet: www.cdc.gov/nchs/