

Explanation of Data Standards for Race, Ethnicity, Sex, Primary Language, and Disability

HHS examined current Federal data collection standards, adequacy of prior testing, and quality of the data produced in prior surveys; consulted with statistical agencies and programs; and built on the Office of Management and Budget (OMB) data collection standards, the Institute of Medicine (IOM) Report *Race, Ethnicity, and Language Data Collection: Standardization for Health Care Quality Improvement*, and its members' experience with collecting and analyzing demographic data.^{i,ii} HHS also paid special attention to current data collection policies for major HHS surveys and those of the Census Bureau.

The following criteria guided the development of data standards for each of the five required variables:

1. Evidence-based and demonstrated to have worked well in practice for national survey data collection.
2. Represent a minimum data standard, with agencies permitted to collect as much additional detail as desired, provided that the additional detail could be aggregated back to the minimum standard.
3. Standards mandated by Office of Management and Budget (OMB) would serve as the starting point for any standard.
4. Standards would be for population surveys in which person-level data is collected via either self-report or from respondents who serve as a knowledgeable household representative.

Proposed Data Collection Standards for the Department of Health and Human Services

I and II. Race and Ethnicity

The starting place for the race and ethnicity data collection standards recommended to the Secretary is OMB's government-wide standards, originally issued in 1997 after a comprehensive public engagement process and extensive field testing. The principles underlying these government-wide standards are described below. The justifications for these principles are described by OMB in detail at http://www.whitehouse.gov/omb/fedreg_1997standards/.

- 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). The surveyor should not tell an individual who he or she is, or specify how an individual should classify himself or herself.
- To provide flexibility and ensure data quality, separate questions for race and ethnicity should be used wherever feasible. Specifically, when self-reporting or other self-identification approaches are used, first ask about ethnicity, and then ask about race. The standard acknowledges that this standard may not work in other contexts (e.g., administrative records.)
- The specified race and ethnicity categories provide a minimum set of categories 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.
 - The OMB minimum categories for race are: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White.
 - The minimum categories for ethnicity are: Hispanic or Latino and Not Hispanic or Latino.
- When self-reporting or other self-identification approaches are used, allow respondents who wish to identify their multi-racial heritage to choose more than one race; there is no multi-racial category.
- OMB encourages additional granularity where it is supported by sample size and as long as the additional detail can be aggregated back to the minimum standard set of race and ethnicity categories.
- 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.

The OMB standard is required in all federally sponsored data collections, including HHS data collection initiatives. HHS is not taking comment on the underlying OMB standards. However because additional granularity in the race and ethnicity categories is important for documenting

and tracking health disparities, large federal surveys such as the National Health Interview Survey (NHIS), Current Population Survey (CPS), and the ACS have implemented more granular strategy, particularly for Hispanic and Asian subpopulations.

Proposed Data Standard for Ethnicity

Are you Hispanic, Latino, or Spanish origin

- a. *No, not of Hispanic, Latino, or Spanish origin*
- b. *Yes, Mexican, Mexican American, Chicano†*
- c. *Yes, Puerto Rican†*
- d. *Yes, Cuban†*

These categories roll-up to the Hispanic or Latino category of the OMB standard

†These categories roll-up to the Hispanic or Latino category of the OMB standard

Proposed Data Standard for Race

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

- a. *White**
- b. *Black or African American**
- c. *American Indian or Alaska Native**

These categories are part of the OMB standard

- d. *Asian Indian***
- e. *Chinese***
- f. *Filipino***
- g. *Japanese ***
- h. *Korean***
- i. *Vietnamese***
- j. *Other Asian***

These categories roll-up to the Asian category of the OMB standard

- k. *Native Hawaiian****
- l. *Guamanian or Chamorro****
- m. *Samoan****
- n. *Other Pacific Islander****

These categories roll-up to the Native Hawaiian or Other Pacific Islander category of the OMB standard

* These categories are part of the OMB standard

** These categories roll-up to the Asian category of the OMB standard

*** These categories roll-up to the Native Hawaiian or Other Pacific Islander category of the OMB standard

Rationale for Race and Ethnicity Data Standard

The proposed HHS standards for race and ethnicity is an expansion of the OMB standard, based on that currently used in the American Community Survey (ACS) and the 2000 and 2010 Decennial Census.ⁱⁱⁱ The new data standard provides additional categories of race and ethnicity from which racial and ethnic differences in health and healthcare can be explored. The Census

Bureau field tested this additional granularity.^{iv} These more detailed race categories roll up to the OMB standard five categories: American Indian or Alaska Native, Asian, Black or African American, Native Hawaiian or Other Pacific Islander, and White. Respondents are also able to select more than one racial categories. These more detailed ethnicity categories roll up to the OMB standard categories: Hispanic or Latino and Not Hispanic or Latino.^{v,vi}

III. Sex

Proposed Data Standard for Sex

What is your sex?

- a. _____ *Male*
- b. _____ *Female*

Rationale for Sex Data Standard

The category of sex was defined as biologic sex. The demonstrated best practice for this measure is male and female.

IV. Primary Language

Proposed Data Standard for Primary Language

How well do you speak English?

- a. _____ *Very well*
- b. _____ *Well*
- c. _____ *Not well*
- d. _____ *Not at all*

Rationale for Primary Language Data Standard

The draft data standard for primary language is English proficiency measurement. The survey item selected for the minimum standard is currently used by the ACS and has been used to assess English proficiency by the Census Bureau since 1980.^{vii viii} Health disparities have been associated with limited English language proficiency.^{ix} For population based surveys, English proficiency is the most appropriate measure to assess language related health disparities, while English proficiency in combination with specific language spoken is more appropriate for the point of health care delivery.

The proposed data standard for primary language represents a minimum standard for population surveys with self-report data, assuming sample sizes are sufficient to support such estimates. Agencies may wish to collect data on the specific language spoken by survey respondents. For data collections for which specific language will be collected, the following survey items are recommended. These items are also currently being used by the ACS and have also been used by the Census Bureau to assess spoken language since 1980.

Proposed Data Collection for Spoken Language

Do you speak a language other than English at home?

Yes

No

For persons speaking a language other than English (answering yes to the question above):

What is this language?

Spanish

Other Language (Identify)

V. Disability Status

Proposed Data Standard for Disability Status

1. *Are you deaf or do you have serious difficulty hearing?*
 - a. *Yes*
 - b. *No*
2. *Are you blind or do you have serious difficulty seeing, even when wearing glasses?*
 - a. *Yes*
 - b. *No*
3. *Because of a physical, mental, or emotional condition, do you have serious difficulty concentrating, remembering, or making decisions? (5 years old or older)*
 - a. *Yes*
 - b. *No*
4. *Do you have serious difficulty walking or climbing stairs? (5 years old or older)*
 - a. *Yes*
 - b. *No*
5. *Do you have difficulty dressing or bathing? (5 years old or older)*
 - a. *Yes*
 - b. *No*
6. *Because of a physical, mental, or emotional condition, do you have difficulty doing errands alone such as visiting a doctor's office or shopping? (15 years old or older)*
 - a. *Yes*
 - b. *No*

Rationale for Disability Data Standard

The six item set of questions used by on ACS and other major surveys to characterize functional disability is proposed as the standard for collecting population survey data on disability. The question set was developed by a Federal interagency committee and reflects how disability is conceptualized consistent with the International Classification of Functioning, Disability, and Health. The question set went through several rounds of cognitive and field testing and has been adopted in most major federal data collection systems. ^x OMB has encouraged the use of this question set when Federal agencies conducting national population studies in order to promote a consistency in measurement and continuity in the dialogue.

ⁱ OMB (Office of Management and Budget). (1977). *Statistical Policy Directive No. 15, Race and Ethnic Standards for Federal Statistics and Administrative Reporting*.

<http://wonder.cdc.gov/wonder/help/populations/bridged-race/directive15.html>(accessed August 3, 2009).

ⁱⁱ IOM (Institute of Medicine). (2009). *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement*. Washington, DC: The National Academies Press.

ⁱⁱⁱ Humes, K. (2009). *2010 Alternative Questionnaire Experiment: Race and Hispanic Origin Treatments*. Final Report, 2009 National Census Test Analysis, U.S. Census Bureau).

^{iv} Alberti, N. (2006) *The 2005 National Census Test: Analysis of the Race and Ethnicity Questions*. Final Report, 2005 National Census Test Analysis. U.S. Census Bureau

^v Office of Management and Budget. (1997a) *Recommendation 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 Classification of Federal Data on Race and Ethnicity*, Federal Register: 62: 36873-36946, July 9.

^{vi} Office of Management and Budget. (1997b) *Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity*, Federal Register: 62: No.210, October 30.

^{vii} Shin, Hyon B. and R. Kominski. (2010). *Language Use in the United States: 2007*, American Community Survey Reports, ACS-12. U.S. Census Bureau, Washington, DC.

^{viii} Divi, C., R. Koss R. G., Schmaltz, S. P., and Loeb, J. M. (2007). *Language Proficiency and Adverse Events in US Hospitals: A Pilot Study*. International Journal for Quality in Health Care 19(2):60-67.

^{ix} Weinick, R. M. and N. A. Krauss. (2000). *Racial/Ethnic Differences in Children's Access to Care*. Journal of Public Health 90(110):1771-1774.

^x Brault, M, S. Stern, D. Raglin. (2007). *Evaluation Report Covering Disability*, American Community Survey Content Test Report P.4. U.S. Census Bureau, Washington, DC.