RACE, ETHNICITY, AND LANGUAGE DATA
STANDARDIZATION FOR HEALTH CARE QUALITY IMPROVEMENT

RECOMMENDATIONS

Recommendation 3-1: An entity collecting data from individuals for purposes related to health and health care should:

- Collect data on granular ethnicity using categories that are applicable to the populations it serves or studies. Categories should be selected from a national standard list (see Recommendation 6-1a) on the basis of health and health care quality issues, evidence or likelihood of disparities, or size of subgroups within the population. The selection of categories should also be informed by analysis of relevant data (e.g., Census data) on the service or study population. In addition, an open-ended option of "Other, please specify:__" should be provided for persons whose granular ethnicity is not listed as a response option.
- Elicit categorical responses consistent with the current OMB standard race and Hispanic ethnicity categories, with the addition of a response option of "Some other race" for persons who do not identify with the OMB race categories.

Recommendation 3-2: Any entity collecting data from individuals for purposes related to health and health care should collect granular ethnicity data in addition to data in the OMB race and Hispanic ethnicity categories and should select the granular ethnicity categories to be used from a national standard set. When respondents do not self-identify as one of the OMB race categories or do not respond to the Hispanic ethnicity question, a national scheme should be used to roll up the granular ethnicity categories to the applicable broad OMB race and Hispanic ethnicity categories to the extent feasible.

Recommendation 3-3: To determine the utility for health and health care purposes, HHS should pursue studies on different ways of framing the questions and related response categories for collecting race and ethnicity data at the level of the OMB categories, focusing on completeness and accuracy of response among all groups.

- Issues addressed should include use of the one- or two-question format for race and Hispanic ethnicity, whether all individuals understand and identify with the OMB race and Hispanic ethnicity categories, and the increasing size of populations identifying with "Some other race."
- The results of such studies, together with parallel studies by the Census Bureau and other agencies, may reveal the need for an OMB review across all agencies to determine the best format for improving response among all groups.

Recommendation 4-1: To assess patient/consumer language and communication needs, all entities collecting data from individuals for purposes related to health and health care should:

- At a minimum, collect data on an individual's assessment of his/her level of English proficiency and on the preferred spoken language needed for effective communication with health care providers. For health care purposes, a rating of spoken English-language proficiency of less than very well is considered limited English proficiency.
- Where possible and applicable, additionally collect data on the language spoken by the individual at home and the language in which he/she prefers to receive written materials.

Recommendation 4-2: The choice of response categories for spoken and written language questions should be informed by analysis of relevant data on the service area (e.g., Census data) or service population, and any response list should include an option of "Other, please specify:__" for persons whose language is not listed.

Recommendation 4-3: When any health care entity collects language data, the languages used as response options or categories for analysis should be selected from a national standard set of languages in use in the United States. The national standard set should include sign language(s) for spoken language and Braille for written language.

Recommendation 5-1: Where directly collected race and ethnicity data are not available, entities should use indirect estimation to aid in the analysis of racial and ethnic disparities and in the development of targeted quality improvement strategies, recognizing the probabilistic and fallible nature of such indirectly estimated identifications.

- Race and ethnicity identifications based on indirect estimation should be distinguished from self-reports in data systems, and if feasible, should be accompanied by probabilities.
- Interventions and communications in which race and ethnicity identifications are based on indirect estimation may be better suited to population-level interventions and communications and less well suited to use in individual-level interactions.
- An indirectly estimated probability of an individual's race and ethnicity should never be placed in a medical record or used in clinical decision making.
- Analyses using indirectly estimated race and ethnicity should employ statistically valid methods that deal with probabilistic identifications.

Recommendation 6-1a: HHS should develop and make available national standard lists of granular ethnicity categories and spoken and written languages, with accompanying unique codes and rules for rollup procedures.

- HHS should adopt a process for routine updating of those lists and procedures as necessary. Sign languages should be included in national lists of spoken languages and Braille in lists of written languages.
- HHS should ensure that any national hierarchy used to roll up granular ethnicity categories to the broad OMB race and Hispanic ethnicity categories takes into account responses that do not correspond to one of the OMB categories.

Recommendation 6-1b: HHS and the Office of the National Coordinator for Health Information Technology (ONC) should adopt standards for including in electronic health records the variables of race, Hispanic ethnicity, granular ethnicity, and language need identified in this report.

Recommendation 6-1c: HHS and ONC should develop standards for electronic data transmission among health care providers and plans that support data exchange and possible aggregation of race, Hispanic ethnicity, granular ethnicity, and language need data across entities to minimize redundancy in data collection.

Recommendation 6-1d: The Centers for Medicare and Medicaid Services (CMS), as well as others sponsoring payment incentive programs, should ensure that the awarding of such incentives takes into account collection of the recommended data on race, Hispanic ethnicity, granular ethnicity, and language need so these data can be used to identify and address disparities in care.

Recommendation 6-1e: HHS should issue guidance that recipients of HHS funding (e.g., Medicare, the Children's Health Insurance Program [CHIP], Medicaid, community health centers) include data on race, Hispanic ethnicity, granular ethnicity, and language need in individual health records so these data can be used to stratify quality performance metrics, organize quality improvement and disparity reduction initiatives, and report on progress.

Recommendation 6-2: HHS, the Department of Veterans Affairs, and the Department of Defense should coordinate their efforts to ensure that all federally funded health care delivery systems collect the variables of race, Hispanic ethnicity, granular ethnicity, and language need as outlined in this report, and include these data in the health records of individuals for use in stratifying quality performance metrics, organizing quality improvement and disparity reduction initiatives, and reporting on progress.

Recommendation 6-3: Accreditation and standards-setting organizations should incorporate the variables of race, Hispanic ethnicity, granular ethnicity, and language need outlined in this report and associated categories (as updated by HHS) as part of their accreditation standards and performance measure endorsements.

- The Joint Commission, NCQA, and URAC should ensure collection in individual health records of the variables of race, Hispanic ethnicity, granular ethnicity, and language need as outlined in this report so these data can be used to stratify quality performance metrics, organize quality improvement and disparity reduction initiatives, and report on progress.
- NQF should review and amend its recommendations on the collection and use of data on race, Hispanic ethnicity, granular ethnicity, and language need to accord with the categories and procedures outlined in this report.
- Medical societies and medical boards should review and endorse the variables, categories, and procedures outlined in this report and educate their members on their use for quality improvement.

Recommendation 6-4: Through their certification, regulation, and monitoring of health care providers and organizations within their jurisdiction, states should require the collection of data on the race, Hispanic ethnicity, granular ethnicity, and language need variables as outlined in this report so these data can be used to stratify quality performance metrics, organize quality improvement and disparity reduction initiatives, and report on progress.