

Physician Data Initiative

Race and Ethnicity Data Collection and Reporting Standard

I. Purpose

The purpose of this standard is to help organizations in the healthcare community to collect and report race and ethnicity data in a way that is as thoughtful as possible and cognizant of the complexities of these data while trying to align with federal practice where applicable, including but not limited to the Office of Management and Budget and the Census Bureau. It is likely that no single data collection methodology will be viewed by everyone in the community as perfect. Moreover, there are different and sometimes conflicting perspectives about how to collect and report race and ethnicity data. Being socially and politically constructed concepts, race and ethnicity categories and subcategories can change over time and vary across countries. With this caveat, the Association of American Medical Colleges (AAMC), Accreditation Council for Graduate Medical Education (ACGME), and the American Medical Association (AMA) have tried to take diverse perspectives into account and have grounded their work in feedback from constituents about how individuals self-identify. All three organizations have agreed to use this standard for race and ethnicity data collection and reporting. Given ambiguities in the distinctions among race, ethnicity, culture, ancestry, geography, nationality, country of origin, and other concepts, the three organizations refer to race and ethnicity in this standard as a shorthand way to describe a holistic approach for allowing individuals to self-identify as the individuals find appropriate.



II. Standards

Data Collection

There are two versions of the data collection standard—a short version that includes the race and ethnicity categories and a long version that includes both the race and ethnicity categories and their associated subcategories.

Short Version

How do you self-identify? Please check as many categories as may apply:

- American Indian or Alaska Native
- Asian
- Black or African American
- Hispanic or Latino
- Middle Eastern or North African
- Native Hawaiian or Pacific Islander
- White
- Some other race or ethnicity: [write-in]

Long Version

How do you self-identify? Please check as many categories and subcategories as may apply:

- American Indian or Alaska Native
 - Tribal affiliation: [write-in]
- Asian
 - Bangladeshi
 - Cambodian
 - Chinese
 - Filipino
 - Indian
 - Indonesian
 - Japanese
 - Korean
 - Laotian
 - Pakistani
 - Taiwanese
 - Vietnamese
 - Some other Asian: [write-in]

- Black or African American
 - African
 - African American
 - Afro-Caribbean
 - Ethiopian
 - Haitian
 - Jamaican
 - Nigerian
 - Somali
 - Some other Black or African American: [write-in]

- Hispanic or Latino
 - Argentinean
 - Colombian
 - Cuban
 - Dominican
 - Mexican or Mexican American
 - Peruvian
 - Puerto Rican
 - Salvadoran
 - Some other Hispanic or Latino: [write-in]

- Middle Eastern or North African
 - Arab
 - Egyptian
 - Iranian
 - Israeli
 - Lebanese
 - Moroccan
 - Palestinian
 - Syrian
 - Some other Middle Eastern or North African: [write-in]

- Native Hawaiian or Pacific Islander
 - Chamorro
 - Fijian
 - Marshallese
 - Native Hawaiian
 - Samoan
 - Tongan
 - Some other Native Hawaiian or Pacific Islander: [write-in]

- White
 - English
 - French
 - German
 - Irish
 - Italian
 - Polish
 - Some other White: [write-in]
- Some other race or ethnicity: [write-in]

Basic principles governing the data collection standard are below:

- Utilization of the data collection standard by organizations within the healthcare community is voluntary. Similarly, responding to a race and ethnicity question must be voluntary and the data collection must contain language that clearly communicates the voluntary nature of the question to the individual responding to the data collection.
- Prior to selecting a race or ethnicity subcategory, an individual must select the corresponding race and ethnicity category.
- Prior to providing a write-in response about another race or ethnicity, an individual must select the corresponding race and ethnicity category and/or subcategory.
- All race and ethnicity categories and subcategories should appear at once, in the alphabetical order shown above, on a single page or screen. As space may dictate, the arrangement of race and ethnicity categories and subcategories may be modified from a vertical orientation to a horizontal orientation; however, the alphabetical ordering shown above must be retained.
- Use of the short version or the long version should be guided by how much detail seems reasonable and the purpose of the data collection.
- If the data may be reported by an authorized respondent on behalf of another individual (e.g., a student, trainee, faculty member, physician, or patient), then two additional race and ethnicity categories may be placed at the end of the listing: “Decline to answer” and “Unknown.” Declining to answer may be because of policy or practice of not sharing these data. This standard will aid in differentiating between a respondent who chooses not to answer the question from one who did not have the adequate information about the individual to respond. In such data collections, the race and ethnicity question should be adjusted to reflect that the data come from a third-party’s administrative records

rather than from the individual directly. As a result, the question should have the form of “How does XX self-identify? Please check as many categories and subcategories as may apply.”.

- Personally identified or identifiable race and ethnicity data should be treated as sensitive personal information as required under applicable law.

Data Reporting

The descriptions of each race and ethnicity data breakout in a report should be detailed enough to allow a reader to understand and replicate the information in the report. Given the many ways that race and ethnicity categories and subcategories can be broken out for reporting purposes, it is important to describe the breakouts including the duplicate and unique counts across each category to ensure accurate counts of the overall population.

III. Methodology for the Race & Ethnicity Data Collection and Reporting Standard

The standard is based on data from a variety of sources, including responses from individuals, including applicants and students to open-ended items and closed-ended items on race and ethnicity questions found in AAMC data collection tools for approximately 50 years. In addition, the standard reflects feedback from the ACGME and AMA based on their experiences, policies, and data expertise. The standard is also informed by research from the US Census Bureau about how to collect race and ethnicity data. Related publications are on the Census Bureau website [Research to Improve Data on Race and Ethnicity](#), including the February 29, 2017 publication [2015 National Content Test: Race and Ethnicity Analysis Report](#). Moreover, this standard factors in the work of the Office of Management and Budget (OMB) that led to the March 28, 2024 announcement of a revision to the 1997 statistical directive on how Federal agencies should collect and report race and ethnicity data. This statistical directive has influenced how many organizations address race and ethnicity data. The initial proposal for a revised statistical directive appeared in the [Federal Register](#) on January 27, 2023 and evolved out of the Federal Interagency Technical Working Group on Race and Ethnicity Standards. Details about the efforts of this Working Group as well as the final statistical directive can be found at [Statistical Policy Directive 15 Revision](#). While the proposed AAMC, ACGME, and AMA standard incorporates much of what the Federal Register contains, the standard departs from the Federal Register in some ways, such as by including as a minimum race/ethnicity category option for “Some other race or ethnicity” (along with a corresponding write-in option). Offering individuals the opportunity to self-identify outside of the explicit OMB race/ ethnicity categories respects individual preferences and enables on-going review of additional categories to be considered as part of an evolving standard.

In recognition that race and ethnicity data collection and data reporting practices have evolved, it seems reasonable to expect that these practices will continue to evolve. As a result, the collaborating organizations will review this standard every three years or as circumstances warrant, such as material changes in the regulatory or legal landscape. Updated standards will require approval from each of the collaborating organizations. At the time of the initial release of the race and ethnicity data collection and data reporting standard, the collaborating organizations were the AAMC, ACGME, and AMA.

Copyright Notice, Permission and Disclaimer and Attribution Statement

This content may be modified at any time upon written agreement of the Parties (email is sufficient).

Unless the instructions indicate otherwise, each use or disclosure of the Standards by any Party or third-party shall include the following placed in close proximity to the Standards:

Copyright Notice:

[*Original Standards – Physician Data Initiative*] © 2024. Accreditation Council for Graduate Medical Education (ACGME), American Medical Association (AMA), and Association of American Medical Colleges (AAMC). May be reproduced and distributed with attribution to the AAMC, ACGME, and AMA. Modifications are permitted with an acknowledgement of the modifications made.

Disclaimer:

DISCLAIMER. The Standards are not guidelines and have not been tested for all potential applications. The Standards, in their original or a modified form, are provided “as is” without express or implied warranties of any kind, including warranties of merchantability, fitness for a particular purpose, or non-infringement. The AAMC, ACGME, and AMA expressly disclaim and assume no liability for use of or reliance upon the Standards whether in their original or a modified form.

Attribution Statements:

Any reproduction or distribution of the Standards without modification shall include the following attribution statement:

“This [reproduction, etc.] was prepared using [Physician Data Initiative Standards] developed by the Accreditation Council for Graduate Medical Education (ACGME), American Medical Association (AMA), and Association of American Medical Colleges (AAMC). The content reflects the views of [name of user/author].”

Any reproduction or distribution of the Standards with modification shall include the following attribution statement:

“This [reproduction, etc.] was prepared using [Physician Data Initiative Standards] originally developed by the Accreditation Council for Graduate Medical Education (ACGME), American Medical Association (AMA), and Association of American Medical Colleges (AAMC) and modified as follows: [____]. The content reflects the views of [name of user/author].”