James L. Madara, MD
CEO, EXECUTIVE VICE PRESIDENT
james.madara@ama-assn.org

April 11, 2023

Shalanda Young
Director
Office of Management and Budget
725 17th Street, NW
Washington, DC 20503

Re: Initial Proposals for Updating OMB's Race and Ethnicity Statistical Standards [OMB-2023-0001]

Dear Director Young:

On behalf of the physician and medical student members of the American Medical Association (AMA), I appreciate the opportunity to offer our comments to the Office of Management and Budget (OMB) on the initial proposals from the Federal Interagency Technical Working Group on Race and Ethnicity Standards (Working Group) for revising OMB's 1997 Statistical Policy Directive No. 15: Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity (SPD 15).[1]

Overall, the AMA supports the proposed revisions and offers additional considerations that will enhance race and ethnicity data standards:

- We support the general approach of combining race and ethnicity and starting with a small number of pre-defined granular categories under each broader "minimum" category since this method is more likely to accurately represent the population.
- Additional standard categories should be offered on the granular list as an optional add-on for implementation with different geographies, contexts, or populations.
- We support the addition of the Middle Eastern and North African category since it is aligned with the work that the AMA has been pursuing in collaboration with other institutions.
- Decisions around naming conventions and burdens of data collection (e.g., privacy, legal jeopardy) should be made in close consultation with people from the populations that data are being collected from. We offer additional naming conventions to consider based on current publications.
- Additional resources should be allocated to support those who would be most burdened by the new data collection standards and those who historically have had the fewest resources.
- The OMB should create a timeline for regular review of race and ethnicity standards and adjust definitions based on new research and changing population trends.

[1] https://www.govinfo.gov/content/pkg/FR-1997-10-30/pdf/97-28653.pdf.

1. Collect Race and Ethnicity Information Using One Combined Question

The Working Group has proposed that SPD 15 move from the two separate questions format to a single combined question as the required design for self-reported race and ethnicity information collections.

1a. Please provide links or references to relevant studies that examine or test any impacts of collecting race and ethnicity information using separate questions compared to a combined question.

Research conducted by the U.S. Census Bureau[2] as well as an independent analysis of American Community Survey data by the Kaiser Family Foundation[3] suggests that the current separate race and ethnicity questions result in a higher ratio of respondents identifying as "Some Other Race" (SOR) and therefore, remaining uncategorized. This has led the U.S. Census Bureau to conclude that a single combined question for race and ethnicity will yield a "more accurate portrait of how the U.S. population self-identifies, especially for people who self-identify as multiracial or multiethnic."[4] Moreover, a study concerning the grouping of race and ethnicity data during COVID-19 showed that when racial and ethnic groups were not mutually exclusive, the percentage of individuals that were able to be counted increased.[5]

Therefore, **the AMA supports the proposal to offer one combined question since it will lessen confusion for individuals self-reporting who believe race and ethnicity are similar, or the same concepts, and will reduce SOR reporting by Hispanic/Latino respondents.[6]**

1b. To what extent would a combined race and ethnicity question that allows for the selection of one or more categories impact people's ability to self-report all aspects of their identity?

"The category of 'Some Other Race' alone or in combination grew to represent more than 15 percent of the U.S. population in the 2020 Census, becoming the country's second-largest racial group."[7] Research conducted by the U.S. Census Bureau suggests that moving from separate questions to a combined question will decrease the proportion of respondents who remain uncategorized, or within the SOR category, from about 10 percent to one to two percent. This change will be particularly impactful among respondents that identify as Hispanic since about three million people of Hispanic or Latino origin were classified as SOR in 2020.[8,9] As such, **the AMA supports the proposal to provide instructions for respondents to "Select all categories that apply" to allow people to self-report all aspects of their identity.**

---

[2] Improved Race, Ethnicity Measures Show U.S. is More Multiracial (census.gov).
[3] Proposed Changes to Federal Standards for Collecting and Reporting Race/Ethnicity Data: What Are They and Why do they Matter? | KFF.
[4] https://www.census.gov/library/stories/2021/08/improved-race-ethnicity-measures-reveal-united-states-population-much-more-multiracial.html.
[5] https://www.cdc.gov/mmwr/volumes/70/wr/mm7032a2.htm.
[6] https://www.urban.org/urban-wire/separating-race-ethnicity-surveys-risks-inaccurate-picture-latinx-community.
[7] https://tnsdc.utk.edu/2023/03/23/federal-race-hispanic-ethnicity-data-collection-proposals-up-for-comment/.
[8] 2015 National Content Test: Race and Ethnicity Analysis Report (census.gov).
[9] https://tnsdc.utk.edu/2023/03/23/federal-race-hispanic-ethnicity-data-collection-proposals-up-for-comment/.

2. Add "Middle Eastern or North African" (MENA) as a new minimum category.

The Working Group proposes that "Middle Eastern or North African" be added to SPD 15 as a new minimum reporting category distinct from all other reporting categories. The definition of the current "White" reporting category would be edited to remove MENA from its definition.

2a. Given the particular context of answering questionnaires in the U.S. (e.g., decennial census, Federal surveys, public benefit forms), is the term "Middle Eastern or North African (MENA)" likely to continue to be understood and accepted by those in this community? Further, would the term be consistently understood and acceptable among those with different experiences, i.e., those born in the U.S., those who immigrated but have lived for an extensive period of time in the U.S., and those who have more recently immigrated to the U.S.?

Racial and ethnic categories are socially constructed, differ between countries, and vary significantly over time.[10] "Race was constructed as a hierarchal human-grouping system, generating racial classifications to identify, distinguish and marginalize some groups across nations, regions, and the world. Race divides human populations into groups often based on physical appearance, social factors and cultural backgrounds" and varies over time, place, and context.[11] Therefore, the AMA recognizes that race is a social construct and is distinct from ethnicity, genetic ancestry, or biology.[12]

Over 700 participants from across the U.S. in 67 focus groups led the U.S. Census Bureau to conclude that it was "inaccurate" to count the MENA population within the "White" category.[13] The current lack of official data renders "invisible the unique challenges faced by Arab/MENA populations" leaving "the MENA population… undercounted and disadvantaged in terms of acquiring services that could benefit this group."[14,15] While alternate terminologies may emerge over time, MENA remains generally understood and accepted by these communities. Ongoing research, such as focus groups, should be conducted to better understand how terminology may change in these communities and provide data to OMB to inform future updates of the Race and Ethnicity Standards.

The AMA strongly supports the addition of "Middle Eastern or North African (MENA)" as a separate reporting category since it will improve disaggregation of data by race and increase the accuracy of the reporting surrounding the diversity of our patients, physicians, and medical students.

2b. Do these proposed nationality and ethnic group examples adequately represent the MENA category? If not, what characteristics or group examples would make the definition more representative?

---

[10] Strmic-Pawl H, Jackson BA, Garner S. Race counts: racial and ethnic data on the U.S. Census and the implications for tracking inequality. Sociology of Race and Ethnicity. 2018;4(1):1-13.
[11] https://www.genome.gov/genetics-glossary/Race.
[12] https://policysearch.ama-assn.org/policyfinder/detail/race%20and%20ethnicity?uri=%2FAMADoc%2FHOD.xml-H-65.953.xml.
[13] Compton E, Bentley M, Ennis S, Rastogi S. 2010 Census Race and Hispanic Origin Alternative Questionnaire Experiment. Washington, DC: U.S. Census Bureau; 2010.
[14] Awad GH, Abuelezam NN, Ajrouch KJ, Stiffler MJ. Lack of Arab or Middle Eastern and North African Health Data Undermines Assessment of Health Disparities. Am J Public Health. 2022;112(2):209-212.
[15] Kayyali R. U.S. Census classifications and Arab Americans: contestations and definitions of identity markers. Journal of Ethnic and Migration Studies. 2013;39(8):1299-1318.

Based upon experience in the medical education context, including two years of piloting a MENA category on the Association of American Medical Colleges (AAMC) Matriculating Student Questionnaire, there has been significant support for using MENA as a category with subcategories including: Afghan, Arab, Armenian, Egyptian, Iranian, Lebanese, Syrian, and Other Middle Eastern or North African (please specify).[16] Within the MENA category, noting the nationalities that currently have the largest populations within the U.S. is reasonable, and supported by AMA policy. However, the relative amount of these nationalities will shift over time and will need to be regularly adjusted. Moreover, to strengthen the proposed changes, in accordance with the American Community Survey (ACS), there should be options for "Middle Eastern," "Arab," and "Arabic" within the OMB categories.

Furthermore, it is important to recognize that the list of relevant specific MENA populations vary by location and other contextual factors. For example, within Chicago, IL, an area with one of the largest Arab ancestry populations in the U.S., "Palestinian" was a top five self-reported ancestry category on the ACS.[17] Though the current list of proposed nationalities provides a mix of geographies, languages, and religions it will be important to provide flexibility that will allow for regional and cultural differences. As such, **the AMA supports the goal of the Working Group's proposal to add MENA as a new minimum category, however, we believe that Afghan, Arab, Armenian, and Palestinian should be included as part of a longer optional list to draw from for specific regions, or within other contexts, to provide additional relevant self-identification opportunities.**

2c. Would this proposed definition allow the generation of statistics necessary to track the experience and wellbeing of the MENA population?

It is likely that the proposed definition will allow for the generation of more accurate statistics surrounding the wellbeing of the MENA population since similar local efforts that disaggregated data for Asian American and Latinx communities discovered previously underrecognized disparities for these groups.[18,19]

Currently, since the OMB does not have a MENA category, researchers utilize definitions of MENA similar to the one proposed by the OMB to investigate health disparities and differences in health behaviors.[20] However, individuals from Somalia are often categorized as MENA or Arab as opposed to African.[21] **As a result, the AMA suggests that the OMB continue conversations with the Somali American community and categorize this population in accordance with their recommendations. Otherwise, the AMA supports the proposed definition and believes it will generate statistics that allow for meaningful disaggregation and highlight obfuscated disparities.**

---

[16] https://www.aamc.org/media/33911/download?attachment (PDF page 21-22).
[17] Arab American Heritage | A Community Portrait | LiveStories.
[18] Health of Asians and Pacific Islanders in New York City (nyc.gov).
[19] Health of Latinos in New York City (nyc.gov).
[20] Abuelezam NN, El-Sayed AM, Galea S. The Health of Arab Americans in the United States: An Updated Comprehensive Literature Review. Front Public Health. 2018 Sep 11;6:262. doi: 10.3389/fpubh.2018.00262. PMID: 30255009; PMCID: PMC6141804.
[21] Abuelezam NN, El-Sayed AM, Galea S. The Health of Arab Americans in the United States: An Updated Comprehensive Literature Review. Front Public Health. 2018 Sep 11;6:262. doi: 10.3389/fpubh.2018.00262. PMID: 30255009; PMCID: PMC6141804.

3. <u>Require the Collection of Detailed Race and Ethnicity Categories by Default.</u>

The Working Group proposes that SPD 15 require data collection on race and ethnicity at the detailed category levels, unless an agency determines that the potential benefit of the detailed data would not justify the additional burden to the agency and the public or the additional risk to privacy or confidentiality. In those cases, agencies must at least use the SPD 15's minimum categories. In any circumstance, agencies are encouraged to collect and provide more granular data than the minimum categories.

3a. Is the example design inclusive such that all individuals are represented?

The provision of the nationalities with the largest populations currently in the U.S. is reasonable, however the composition of the population of the U.S. will shift over time and so this question will need to be updated regularly. Therefore, **the AMA believes that an optional supplemental longer list of checkboxes that can be used in specific geographies, contexts, or populations, should be provided. Furthermore, to avoid confusion, country of birth should be a separate question from race and ethnicity since the answer to these questions can be the same or different.**

In addition, it is noteworthy that "American Indian or Alaska Native" is the only category without checkboxes since this may disadvantage this group in being counted. As such, **the AMA strongly suggests consulting with experts, such as the National Congress of American Indians (NCAI), to determine if this community wishes to be the only group without granular categories, such as tribal nation, pre-specified.**

3b. The example design seen in Figure 2 collects additional detail primarily by country of origin. What other potential types of detail would create useful data or help respondents to identify themselves?

Religious groups have at times been racialized or considered as ethnicities (for example, Jewish or Muslim).[22,23,24] However, the U.S. Census Bureau can only collect data on religious affiliation on a voluntary basis and currently only asks this question in some surveys. Therefore, it is important to further study the interplay of identification by religious affiliation and its connection with race or ethnicity (or national origin) to determine if this will create useful data or help respondents to identify themselves.[25]

3c. Some Federal information collections are able to use open-ended write-in fields to collect detailed racial and ethnic responses, while some collections must use a residual closed-ended category (e.g., "Another Asian Group"). What are the impacts of using a closed-ended category without collecting further detail through open-ended written responses?

Self-identification is considered not only a gold-standard for data collection but also an important principle of respect and dignity for the individuals responding to race and ethnicity questions.[26] The

[22] The Racial Identity of U.S. Jews | Religion and Public Life at Harvard Divinity School.
[23] The Racialization of Muslims in the Post-9/11 United States | The Oxford Handbook of Philosophy and Race | Oxford Academic (oup.com).
[24] Between Anti-Black Racism and Islamophobia (degruyter.com).
[25] 2015 National Content Test: Race and Ethnicity Analysis Report (census.gov).
[26] Soucie J, Buckley BO, Albanese K, Harrington R, Hudson Scholle SH. Current Health Plan Approaches to Race and Ethnicity Data Collection and Recommendations for Future Improvements. (Lacourciere. J, ed.). National Committee for Quality Assurance; 2023.

correlation is weak between race or ethnicity self-identified or "socially assigned" by others and varies by group, with the largest health gaps correlated with socially assigned rather than self-identified race or ethnicity.[27] Conversely, using a close-ended category minimizes the opportunity for individuals to self-identify.

Moreover, using a close-ended category, such as "other" or "another," risks missing markedly minoritized or marginalized categories and emerging trends in small, but rapidly growing, categories. As such, the current prompt, noted in Figure 2 in the proposal, to amend the language to "Enter" without the use of the word "Other" is more welcoming and therefore a superior alternative.

3d. What should agencies consider when weighing the benefits and burdens of collecting or providing more granular data than the minimum categories?

The primary benefit of collecting more granular data is being able to identify root causes, as well as gaps in resources and outcomes, since these issues can only be addressed and tracked, if they are measured and defined. Factors including nativity, "immigrant status, language, socioeconomic status, and experiences with structural and interpersonal racism all significantly influence health outcomes, and that nuance is lost once data are…aggregated into five or six broad categories of race and ethnicity."[28] Moreover, "that lack of systematic disaggregation—whether it be at the collection, analysis, or reporting phase—limits the health and social services fields' ability to target their resources where they are most needed, to all communities experiencing significant disparities."[29]

For example, "Asian Americans are one of the most economically divided racial groups in the U.S. Asian Americans demonstrate a bimodal distribution, being overrepresented relative to the total U.S. population in both the top and bottom 10 percent of the income spectrum. Conversely, the aggregation of income data among Asian Americans misleadingly suggests that they, as a singular population, are actually thriving."[30] Another example supporting the importance of more granular categories, is a "yearslong study…about the inequities Arab Americans face in the Chicago area" which shows that "fewer Arab Americans own their residences and more of them are 'housing cost burdened…than white residents."[31] The study also shows that Arab Americans have "lower median household incomes and higher unemployment and uninsured rates than white residents."[32] In cases where data have been disaggregated, such as New York City, it has shed light on large variations among specific populations, including health insurance and premature mortality among Latino groups,[33] and income and suicide rates among Asian and Pacific Islander groups.[34] As such, more granular data can shed light on the heterogeneity of specific populations, bring focus to health issues that might otherwise be missed or underestimated, and highlight positive outcomes of subgroups.[35] Furthermore, granular data on demographics of patients and communities can allow for analysis of how outcomes are impacted by the degree of concordance with demographics of physicians and other health care workers. In practical application, the information gained from more granular data can help create pathways for care grounded in cultural humility,

---

[27] Using "Socially Assigned Race" to Probe White Advantages in Health Status on JSTOR
[28] https://www.healthaffairs.org/do/10.1377/forefront.20211123.426054.
[29] https://link.springer.com/article/10.1007/s11113-020-09631-6.
[30] The Critical Role of Racial/Ethnic Data Disaggregation for Health Equity | SpringerLink.
[31] Chicago area Arab Americans face stereotypes, invisibility | WBEZ Chicago.
[32] Chicago area Arab Americans face stereotypes, invisibility | WBEZ Chicago.
[33] Health of Latinos in New York City (nyc.gov).
[34] Health of Asians and Pacific Islanders in New York City (nyc.gov).
[35] Black Heterogeneity in Cancer Mortality: US-blacks, Haitians, and Jamaicans - PMC (nih.gov).

particularly in areas of mental health, vaccinations and screening, and chronic disease. Moreover, it can aid in a multifaceted approach in health care settings and promote personalized patient-centered care.[36]

However, there are additional potential burdens associated with the collection and provision of more granular data, along with practical constraints related to the collection, analysis, and reporting of data across agencies. These impediments are particularly poignant for institutions that are smaller or have less resources such as smaller physician practices and tribal and regional intertribal data collection entities.[37] As such, implementation of standards may need dedicated funding, or absent funding, allowable exceptions, for smaller organizations with fewer resources, to avoid further straining these service providers which may already be teetering toward closure.

Beyond funding, practical constraints can include implementing changes in electronic health records, workflows, and setting aside staff time and resources for training, collection, analysis, and reporting. Therefore, organizations contracted or required to implement more granular data strategies should be supported with the resources and technical assistance needed to engage the communities providing the data, implement these changes, and analyze and report their data.[38]

Moreover, during implementation, it is important to consider that many smaller practices rely on less resourced or niche health information technology (health IT) vendors, e.g., electronic health record (EHR) companies, to support the bulk of their technological needs. Changes in federal policies that would require modifications to EHR products, e.g., new race and ethnicity data capture/reporting requirements, must understand the impact and timelines needed for health IT vendors to update, test, and implement new capabilities. As such, it would help smaller organizations with fewer resources to be at the tail end of the implementation cycle, as part of a phased approach.

Furthermore, ensuring privacy for participants' personal information is paramount when collecting more granular data. This is because "[t]he collection of more meaningfully disaggregated racial/ethnic data is uncommon, but the analysis and reporting of those data are even more rare…. There are numerous methodological challenges to doing this well. For example, disaggregating data in research studies, even after more nuanced data have been collected, can result in small sample sizes, which risks generalizability and creates data privacy concerns."[39] In a recent survey of 1000 patients, nearly 75 percent said they were concerned about protecting the privacy of their personal information. Six in ten patients were worried about their information being used by companies to discriminate against them or their loved ones or to exclude them from opportunities to find housing, gain employment, or receive other benefits. The survey also identified that over 50 percent of patients were "very" or "extremely" concerned that unnecessary access to their information could result in negative repercussions related to insurance coverage, employment, or opportunities for health care.[40] As such, while many patients recognize the importance of accurately capturing race and ethnicity information, they are worried about the consequences of their information being misused.

Finally, the burden of providing more granular data collection can fall on respondents who may not have the time or capacity to respond and may be wary of how such data have been used in the past or may be

---

[36] Improving Patient Race and Ethnicity Data Capture to Address Health Disparities: A Case Study From a Large Urban Health System - PMC (nih.gov).
[37] DataDisaggregationAIAN-report_5_2018.pdf (ncai.org).
[38] Collection of Race and Ethnicity Data for Use by Health Plans to Advance Health Equity | Urban Institute.
[39] The Critical Role of Racial/Ethnic Data Disaggregation for Health Equity | SpringerLink.
[40] https://www.ama-assn.org/system/files/ama-patient-data-privacy-survey-results.pdf.

used in the future. For example, Arab Americans, who would be more visible with implementation of a MENA category, have been subject to both invisibility in the data and hypervisibility when targeted as a group for discrimination.[41] Data repositories like the census, despite safeguards, have been used in the past to target various groups.

Operational and Technical Considerations for the Collection of Race/Ethnicity Data on Health Care Forms/Transactions

With the increased collection of more granular race and ethnicity data, many workflow processes will be new which will be an added burden, especially on smaller physician practices and other health care providers. Therefore, **the AMA asks the Working Group to consider the following technical and operational aspects of data collection on health care forms and revenue cycle transactions:**

- Physicians and other health care providers must understand the importance of collecting the data and reporting it in applicable health care transactions. Without understanding the importance, practices and other organizations will likely view the activity as another regulatory burden, which will lessen the likelihood of collection.

- Technical solutions must be employed to support the collection and reporting of race and ethnicity data. Most physician practices and other health care providers rely on health IT vendors to support their data needs. Vendor products that minimize the work effort will be critical to the success of collecting and reporting these data. In addition, standardized collection tools must be in place to ensure these efforts result in meaningful, consistent, and interoperable data.

- Make the collection and reporting of race and ethnicity data in health care transactions voluntary for patients. While we have heard that some minoritized organizations are frustrated that race and ethnicity data are not currently routinely collected, we are also concerned that some minoritized or marginalized patients may fear that sharing this information will lead to further stigmatization or prejudice in the health care system.

- Electronic data transmission standards must be updated to capture race and ethnicity standards. The health care system continues to use a patchwork of paper forms and electronic transaction standards that exchange data between physicians, health care providers, health plans, and relevant organizations. Since both paper and electronic formats are frequently used, both need to accommodate the race and ethnicity data.

- Policy protections must be put in place to ensure that race and ethnicity data are only used to improve patient health outcomes and support public health initiatives. Regulators and other policymakers must develop guardrails to prevent the misuse of data. For example, health plans could use race and ethnicity data to limit plan options or coverage in certain geographic locations to avoid insuring high-risk, costlier patient populations.

Given these challenges, granular data should only be collected and provided when there is a clear plan for its successful use, including analysis and reporting, that ensures ethical and equitable accountability to those providing data. As such, federal policies must ensure that data are protected

[41] UIC report examines experiences, racial justice for Chicagoland's Arab Americans | UIC Today.

and privacy protections should emphasize equity and justice, contemplate individuals' rights, establish guardrails on data use, and be enforceable.

3e. Is it appropriate for agencies to collect detailed data even though those data may not be published or may require combining multiple years of data due to small sample sizes?

It is appropriate to collect more detailed data if the people providing it are meaningfully consulted, and there are adequate protections against the data being used to create stigmatizing, exclusionary, or punitive policies. Furthermore, it is valuable to collect data with the goal of merging multiple years when there are small sample sizes that do not allow annual reporting due to privacy, confidentiality, or statistical concerns.

Lack of data about communities, especially due to small sample size, has negative implications for federal data users' ability to be better informed about demographic shifts, upcoming needs, and budgets for smaller populations. As such, a small sample size analysis may be necessary to ensure adequate and timely reporting of data.

3f. What guidance should be included in SPD 15 or elsewhere to help agencies identify different collection and tabulation options for more disaggregated data than the minimum categories? Should the standards establish a preferred approach to collecting additional detail within the minimum categories, or encourage agencies to collect additional information while granting flexibility as to the kind of information and level of detail?

Guidance should provide a tiered approach to data collection, using only the broader ("minimum") categories under special circumstances. Standardization should be encouraged across this tiered approach to maximize the opportunities for comparison between jurisdictions and increase the ability to merge data across years or "rolling data up" from granular to broader ("minimum") categories when needed.

4. Update Terminology in SPD 15.

The Working Group proposes to make several changes to the terminologies used within minimum categories; discontinuing the terms majority/minority; and questions stem and instructions.

4a. What term (such as "transnational") should be used to describe people who identify with groups that cross national borders (e.g., "Bantu," "Hmong," or "Roma")?

Given the long history of human migration, and the regular revision of politically defined geographies, most groups exist across national borders. The idea that ethnicity and national origin are synonymous is socially constructed and historically inaccurate, contributing to abuses ranging from targeted statelessness to genocide. Terminology should comport with what people identifying with a particular group endorse, rather than labels created by others not in that group. There is not a clear and stable distinction between nationality, national origin, ancestry, and ethnicity (or race for that matter) so it may be best to leave this ambiguity in place and avoid arbitrary divisions between national and transnational. However, due to the unique needs of communities, such as Roma, Bantu, and Hmong, it is important to give these individuals a space to self-identify.

1. *If a combined race and ethnicity question is implemented, what term should be used for respondents who select more than one category? For example, is the preferred term "multiracial," "multiethnic," or something else?*

The AMA recommends that the OMB use "more than one race or ethnicity" or "multiple races or ethnicities" to reflect individuals who identify with multiple categories. However, it should be noted that having a category that identifies multiple races or ethnicities will not be useful without the context that is provided when people can self-identify with more than one race or ethnicity.

2. *Are these draft definitions (Referring to Section D, Previously Tested Definitions of Minimum Categories):*

i. Comprehensive in coverage of all racial and ethnic identities within the U.S.?

Since definitions change over space and time, the categories at any given moment will never be totally comprehensive. However, based on recent research by the U.S. Census Bureau and others, the proposed definitions of minimum categories seems to be the best set of categories for the current moment in history. Though these categories are not mutually exclusive, this is addressed to some degree by allowing multiple selections.

Furthermore, the use of "other" as an option is recommended to better understand which groups or social identities are not fully captured in the updated minimum standards. The AMA recognizes that this will likely require additional data cleaning and collection; however, the benefits outweigh the costs given that this category provides insights into emerging race and ethnicity considerations.

ii. Using equivalent criteria?

These categories are not all of the same type, which is why they cannot be mutually exclusive, but it is unlikely a suitable alternative exostsat this time. Some of the categories are racialized identities (White), while others suggest politically defined geographies (Middle Eastern or North African), and still others have a false pretense of geographies (most of Asia is on the same tectonic plate as Europe and the definition of continents has changed over time).[42,43] These definitions are further confused by the granular categories which are more consistent with nationalities.

Moreover, some of the race and ethnicity categories use "or" as two alternative designations for the same category (Black or African American), whereas other categories use "or" to join to adjacent geographic areas (Middle Eastern or North African), so this may need some explanation, or a different joining word or symbol for one of the two scenarios.

Furthermore, in compiling this Public Comment Letter to OMB, we became aware of a related comment period from the Centers for Disease Control and Prevention (CDC) about its CDC Race and Ethnicity 2022 Code System.[44] Our understanding is that the OMB SPD 15 and CDC Code System both include the minimum race and ethnicity categories defined by OMB, but CDC provides a more detailed set of race and ethnicity categories developed and maintained by the U.S. Bureau of the Census (BC). However, there are fundamental differences in terminology between OMB's draft SPD 15 and CDC's draft 2022 Code System, and with parallel comment periods occurring simultaneously, we want to stress the

[42] Tectonic Plates of the Earth | U.S. Geological Survey (usgs.gov).
[43] The Myth of Continents (nytimes.com).
[44] https://phinvads.cdc.gov/vads/DownloadHotTopicDetailFile.action?filename=B1399CAC-1FB2-ED11-81C2-005056ABE2F0.

importance of reconciling the terminology between the two data sets so that there is harmonization across federal race and ethnicity data collection, classification, and publication.

In addition, both data sets are noted as applicable vocabulary standards in the United States Core Data for Interoperability (USCDI) Version 1.[45] Unless there is consistency in terminology across both data sets, we are concerned that there will be confusion about which terminology to use. We encourage OMB to ensure that the final versions of OMB SPD 15 and the CDC Race and Ethnicity 2022 Code System are broadly consistent in the terminology that they use, and it is clear to the public and health care community stakeholders which terminology should be used in which instances.

iii. Reflective of meaningful distinctions?

These categories are reflective of meaningful distinctions insofar as the vast majority of the time respondents are able to select something they identify with, which is supported by the U.S. Census Bureau's research. The distinctions will change over time and space and need to be regularly updated based on how respondents see themselves and minimize uncategorized responses.

iv. Easy to understand?

The U.S. Census Bureau research suggests that these categories are widely understood.

v. Respectful of how people refer to themselves? Please suggest any alternative language that you feel would improve the definitions.

The people who select these identities are those best positioned to answer whether the categories are respectful. Of note, the MENA category is a sought-after improvement by people identifying with this category and it will be important to study how self-identification changes over time.

Moreover, there are alternative political geographies such as the Caribbean or Central American and alternative labels such as Indigenous that are absent and stakeholder input from people identifying with these categories should be explicitly and purposefully solicited if it has not already been.[46] For political geography, this relates to making available additional options for "rolling up data" from granular categories, and for alternate labels this relates to maximizing the number of people who see themselves in the categories provided.

Furthermore, explicitly mentioning Central American and Brazilian ethnicities within the Hispanic or Latino category would help clarify that this term includes those from Latin America who are often invisible or who may not have Spanish as their predominant language.

4b. As seen in Figure 2, based on the Working Group's initial proposal, the question stem asks "What is your race or ethnicity?" Do you prefer a different question stem such as: "What is your race and/or ethnicity?", "What is your race/ethnicity?", "How do you identify?", etc.? If so, please explain.

[45] https://www.healthit.gov/isa/united-states-core-data-interoperability-uscdi.
[46] Article: Caribbean Immigrants in the United States | migrationpolicy.org.

The question stem as proposed is a reasonable option leaving open any combination of the two. However, the use of a singular verb ("is") may suggest to respondents that they should choose only one. Alternatives that welcome selecting more options include:

- Please describe your race/ethnicity.
- What races and/or ethnicities do you identify with?[47]
- How would you describe your race and ethnicity?[48]

5. <u>Guidance is necessary to implement SPD 15 revisions on Federal information collections.</u>

The Working Group proposes that SPD 15 and its related documents be placed online in a central location and include implementation guidance on certain topics.

5b. With the proposals to use a combined race and ethnicity question and to add MENA as a minimum category, what specific bridging concerns do Federal data users have? Please submit any research on bridging techniques that may be helpful to the Working Group. Bridging refers to making data collected using one set of categories (e.g., two questions without MENA), consistent with data collected using a different set of categories (e.g., one question with MENA).

The Wisconsin Department of Public Instruction (DPI) offers a helpful summary of the concerns with various approaches when more categories are newly available and need to be bridged with historic data with fewer categories. Three common methodologies for the assignment of individuals who select two or more categories include:

1. Assignment to the smallest category of those selected.
2. Assignment to the largest category of those selected.
3. Assignment to largest, non-White category of those selected.

Bridging to the smallest category would have a greater chance of creating statistical anomalies in trend analysis. Bridging to the largest category might cause individuals who may have identified with minoritized groups in the past to now be counted in the White category, further exacerbating underrepresentation of minoritized groups. Bridging to the largest non-White category, based on the individual's response, preserves reporting in a minoritized category while minimizing the statistical impact of adding individuals to a category.[49]

More complex bridging methods have been proposed. The National Cancer Institute, related to tobacco, documented a multiple imputation method only for multiple races/ethnicities which yielded satisfactory results. However, small samples for some groups introduce inaccuracy of the probabilities and rapid changes in demographics make bridging more inaccurate over time.[50] Methods used by the National

[47] Rethinking the design of the race and ethnicity question on surveys | by Sara Clayton | UX Collective (uxdesign.cc).
[48] Estimates of Multiracial Adults and Other Racial and Ethnic Groups Across Various Question Formats | Pew Research Center.
[49] dpi_race_ethnicity_bridge.doc.
[50] Bridging Estimates by Race for the Tobacco Use Supplement to the Current Population Survey – (TUS-CPS) (cancer.gov).

Center for Health Statistics (NCHS), and adapted by others, involving regression and public-use micro data may be the most accurate but depend on geographic data that may not always be available.[51] The Census Bureau outlined their approaches but it is unclear if they could be operationalized broadly.[52]

It is noteworthy that bridging techniques are often used to assign multiple races/ethnicities a primary race/ethnicity. For bridging techniques to assess differential changes in data collection methods and changes in the self-reporting of one's identity, similar methods of assignment to the non-White category should be selected.

American Indian or Alaska Native, Hispanic or Latino, and Native Hawaiian or Pacific Islander are the primary groups whose race or ethnicity self-identification may shift over multiple reporting periods.[53] Additionally, with the option of Middle Eastern and North African there is an expected reduction in the White category. When greater specificity is provided, bridging should default to self-identification in the most recent and more expansive categories. There are multiple reasons why self-identification may change in self-reporting including greater awareness of one's identity, greater specificity in response options, and to gain specific benefits assigned to a particular group.[54] Federal data users should recognize that six to seven percent of self-report responses may shift, and apply similar limitations for the data used.

Furthermore, recent recommendations from the National Committee for Quality Assurance propose that getting at the truth of different or conflicting race and ethnicity responses for individuals requires an established process, with input from various interested parties, that account for identifying the most valid data source. This ensures a more transparent process and minimizes bias in arriving at an identity.[55]

5c. What guidance on bridging should be provided for agencies to implement potential revisions to SPD 15?

Bridging to the largest, non-White category based on the individual's response preserves reporting in a minoritized category while minimizing the statistical impact of adding individuals to a category. This option was utilized by Wisconsin DPI and the U.S. Department of Housing and Urban Development since it avoids the challenges associated with implementing a probabilistic bridging method. As such, this may be the best option in the presence of variable access to resources or geographic data. However, the regression method used by NCHS, and adapted by others, or alternatives offered by the Census Bureau may be preferable if all parties have access to resources and geographic data.

Being explicit in the bridging methodology and noting limitations regarding the socially constructed race and ethnicity terminology is an important component for the reporting of data. Additionally, the most impacted groups in demographic shifts based on the bridging methodology should be paid attention to and explicitly noted.

---

[51] A Practical Approach to Using Multiple-Race Response Data: A Bridging Method for Public-Use Microdata - PMC (nih.gov).
[52] PowerPoint Presentation (census.gov).
[53] America's Churning Races: Race and Ethnic Response Changes between Census 2000 and the 2010 Census - PMC (nih.gov).
[54] Millions of Americans changed their racial or ethnic identity from one census to the next | Pew Research Center.
[55] Soucie J, Buckley BO, Albanese K, Harrington R, Hudson Scholle SH. Current Health Plan Approaches to Race and Ethnicity Data Collection and Recommendations for Future Improvements. (Lacourciere. J, ed.). National Committee for Quality Assurance; 2023.

5d. How should race and ethnicity be collected when some method other than respondent self-identification is necessary (e.g., by proxy or observation)?

It is reasonable to collect data from a legal representative or designee of the person when the person is unable to answer. In cases of death where next of kin are unavailable, past records with self-identification should be prioritized over observation.

Several studies have documented misclassification of race and ethnicity as an issue of concern for minoritized groups including American Indian or Alaska Native, Asian or Asian American, Hispanic or Latino, Middle Eastern, and Native Hawaiian and Pacific Islander communities.[56,57] Misclassification rates can range from seven percent to 30 percent depending on the particular minoritized population, and undermines the ability of health care and public health systems to adequately address health inequities. While missing race and ethnicity data are an ongoing consideration, when observation is the method used to report the information, the high likelihood of misclassification raises concerns among populations that are already undercounted. This misclassification has been most notably studied in electronic health and/or mortality records. Several prominent examples that emerged during COVID-19 illustrated how misclassification can minimize the true impact of inequities on minoritized communities.

5e. What guidance should be provided for the collection and reporting of race and ethnicity data in situations where self-identification is unavailable?

People collecting and reporting on data should avoid data collected by observation, even if the tradeoff is missing data. It should be noted where possible, the degree to which data are drawn from asking a legal representative or designee of the person or accessing a past record with self-identification, when the person was unable to answer.

6. Comments on Any Additional Topics and Future Research.

6a. SPD 15 does not dictate the order in which the minimum categories should be displayed on Federal information collections. Agencies generally order alphabetically or by population size; however, both approaches have received criticism. What order, alphabetical or by population size, do you prefer and why? Or what alternative approach would you recommend?

With any approach, the order of categories will change over time due to shifts in populations or naming conventions. Alphabetical is preferred so the order does not reinforce power dynamics based on the relative size of populations.

6b. The current minimum categories are termed. Do you have suggestions for different terms for any of these categories?

The AMA believes that an alphabetized version of the proposal for updated categories is preferable to the current broad ("minimum") categories.

- American Indian or Alaska Native

[56] State of Racial Justice Reports | Institute for Research on Race and Public Policy | University of Illinois Chicago (uic.edu).
[57] Racial Misclassification of American Indians and Alaska Natives by Indian Health Service Contract Health Service Delivery Area - PMC (nih.gov).

- Asian
- Black or African American
- Hispanic or Latino
- Middle Eastern or North African
- Native Hawaiian or Pacific Islander
- White

"American Indian, Indian, Native American, or Native are acceptable and often used interchangeably in the United States. However, Native Peoples often have individual preferences on how they would like to be addressed…. In Mexico, Central America, and South America, the direct translation for Indian can have negative connotations. As a result, they prefer the Spanish word *indígena* (Indigenous), *comunidad* (community), and *pueblo* (people)."[58] Therefore, regarding the American Indian or Alaska Native category, the AMA is interested in the opinions of those identifying with this category as to whether Native American should be offered, or the word Indigenous should also be offered (or if this is considered confusing since the category is meant to be specific to the Americas and was not meant to include people who identify as Indigenous in other parts of the world). In terms of the Hispanic or Latino category, additional considerations include the emerging expansion of how this group self-identifies. This includes Latina, Latiné, Latino, Latinx, or Hispanic or of Spanish origin [LHS+].[59] Regarding the White category, we are interested in the opinions of those identifying with this category as to whether the word European should also be offered (or if this is considered confusing for people who identify as White in other parts of the world).

In addition to the alphabetical order of broad "minimum" categories, some additional language considerations could include:

- American Indian, Alaska Native, Indigenous, or Native American
- Asian or Asian American
- Black or African American
- Hispanic or Latina/Latiné/Latino/Latinx
- Middle Eastern or North African
- Native Hawaiian or Pacific Islander
- White
- Other

Finally, we agree with the other working group suggestions as outlined at the beginning of Section 4.

6c. *How can Federal surveys or forms collect data related to descent from enslaved peoples originally from the African continent? For example, when collecting and coding responses, what term best describes this population group (e.g., is the preferred term "American Descendants of Slavery," "American Freedmen," or something else)? How should this group be defined? Should it be collected as a detailed group within the "Black or African American" minimum category, or through a separate question or other approach?*

[58] The Impact of Words and Tips for Using Appropriate Terminology | Helpful Handout Educator Resource.
[59] Latinx Race and Ethnicity Data Gaps: The HACER Campaign and a Call to Action | AJPH | Vol. 112 Issue 10 (aphapublications.org).

We would defer to people who select these identities as those best positioned to suggest the terms for these categories. Several options in use to consider include:

- Descendants of enslaved (Black or African) (Americans or people) in the United States.[60]
- African Americans and those of African descent in the United States who are descendants of slavery.
- Descendants of enslaved people of African descent in the United States.

It is important to keep the question about descendants of slavery broad because about 57 percent of Black adults say that their ancestors were enslaved either in the U.S., outside the U.S., or both in the U.S. and abroad.[61] However, about 34 percent of Black Americans are unsure if their ancestors were enslaved.[62]

Finally, the history of enslavement of Native Americans is underrecognized and still being reconstructed.[63] Based on this observation, the best course may be to use a term such as "descendants of people enslaved in the United States" in a question separate from the race or ethnicity question, allowing for a fuller picture of enslavement to be constructed.

6d. The proposals in this FRN represent the Working Group's initial suggestions for revisions to SPD 15 to improve the accuracy and usefulness of Federal race and ethnicity data. The Working Group and OMB welcome comments and suggestions on any other ways that SPD 15 could be revised to produce more accurate and useful race and ethnicity data.

Under Hispanic or Latino, we suggest including Brazilian as an ethnicity, even though it does not align with the population size prioritization, to clarify who identifies as Latino even if they do not identify as Hispanic.

Additionally, we suggest offering two alternative spellings, Chamorro and Chamoru, under Native Hawaiian or Pacific Islander.[64]

Conclusion

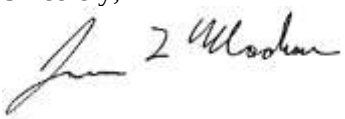
The AMA supports the adoption of racial and ethnic demographic data collection practices that allow for self-identification and the designation of one or more racial categories. Moreover, the AMA supports reporting demographic data in categories of race and ethnicity whereby Latino, Hispanic, and other identified ethnicities are categories, irrespective of race. Therefore, we encourage the OMB to adopt demographic data reporting practices that permit disaggregation of individuals who have chosen multiple categories of race so as to distinguish each category of individuals' demographics alone or in combination

[60] California becomes first state to break down Black employee data by lineage: NPR.
[61] https://www.pewresearch.org/race-ethnicity/2022/04/14/black-americans-family-history-slavery-and-knowledge-of-black-history/.
[62] https://www.pewresearch.org/race-ethnicity/2022/04/14/black-americans-family-history-slavery-and-knowledge-of-black-history/.
[63] Colonial enslavement of Native Americans included those who surrendered, too | Brown University.
[64] Chamorro vs. Chamoru (guampedia.com).

with any other racial and ethnic category and to develop standardized processes and identify strategies to improve the accurate collection, disclosure, and reporting of racial and ethnic data.[65]

Thank you for considering the AMA's comments. If you have any questions, please contact Margaret Garikes, Vice President of Federal Affairs, at margaret.garikes@ama-assn.org or 202-789-7409.

Sincerely,



James L. Madara, MD

[65] https://policysearch.ama-assn.org/policyfinder/detail/race%20and%20ethnicity?uri=%2FAMADoc%2Fdirectives.xml-0-1916.xml.