

April 5, 2023

Bob Sivinski
Chair
Interagency Technical Working Group on Race and Ethnicity Standards

Submitted electronically: regulations.gov

RE: Comments on OMB's Proposed Revised Race and Ethnicity Data Standards—AHIP Comments

Dear Bob Sivinski:

Thank you for the opportunity to provide input on the Office of Management and Budget's (OMB) proposed revised race and ethnicity data standards. AHIP appreciates the Administration's focus and leadership on health equity. Achieving health equity is also a priority for AHIP and our health insurance provider members. We take our collective responsibility to improve health equity seriously.

Robust, accurate, actionable, and standardized demographic patient data is fundamental to advancing health equity and improving health outcomes. Accurate, complete, and interoperable data is necessary to effectively and efficiently detect disparities, inform effective solutions that take cultural preferences and socioeconomic circumstances into account, and assess what interventions are most impactful. We applaud OMB for leading evidence-based and stakeholder-driven efforts to improve the accuracy and completeness of demographic data by revising race and ethnicity data standards. In the future, we urge OMB to lead similar efforts to develop standards for other demographic data elements that are important to identity (e.g., sexual orientation, gender, language, disability status, etc.) and align these standards across the health care ecosystem to promote interoperable data exchange.

Standardizing and improving demographic data has been a priority of the health insurance industry. Between 2020 and 2022, AHIP sought to develop improved demographic data standards that are more aligned, patient-centered, and actionable than the various standards that exist today. AHIP employed an evidence-based and stakeholder-driven process to conduct this work by convening diverse groups of health insurance providers and other stakeholders (e.g., patients representing different communities, providers, community-based organizations, and

others). In addition to race and ethnicity, our workgroup also [developed](#) recommended data standards for language, sexual orientation, gender, pronouns, relationship status, disability status, military experience, and spirituality.

We hope this work will lead to high-level data standardization and alignment across the health care ecosystem while allowing for local granular customization. This will ensure organizations can collect a reasonable amount of data that is actionable and relevant for their local communities while still aggregating and exchanging data seamlessly with appropriate entities. This work informed the following recommendations.

1. Collect race and ethnicity information using one combined question

OMB seeks feedback on moving from two separate questions on race and ethnicity to one combined race/ethnicity question.

Recommendation: We agree that combining race and ethnicity will help reduce confusion and improve accuracy of demographic data. Most people do not understand the difference between race and ethnicity or why Hispanic or Latino is the only ethnicity considered. We believe OMB’s proposed revised combined race/ethnicity question is a good starting point.

There are tradeoffs by moving from two separate race and ethnicity questions to one combined question. OMB highlights public concerns about how race data might be lost with a combined single race/ethnicity question. We heard similar concerns in our stakeholder engagement process. Moving from two questions to one combined question will also require changes to data systems, education for staff and the broader community, revisions to existing quality measures that are currently structured around OMB’s 1997 standards, and may impact data bridging and analysis. To address these concerns while still improving the accuracy of data by combining race/ethnicity into one question, we [developed](#) the following potential solution that we believe is the best approach:

- Continue with two separate questions for ethnicity and race—each of which would contain high-level reporting categories and more granular response options for each high-level category.
- On the race question, add a high-level response category for “I only identify as Hispanic or Latino/a.”

This way, respondents can note both a Hispanic or Latino/a ethnicity as well as additional races with which they may identify. Or, if someone only identifies as Hispanic or Latino/a, then they have the option to denote that on both an ethnicity question as well as a race question. This will help bridge data previously collected with data collected moving forward to permit longitudinal analysis while ensuring individuals have options with which they identify. Examples of how these questions could look are below:

Ethnicity: Do you identify as Hispanic or Latino/a?

- I am Hispanic or Latino/a [Include granular options]
- I am not Hispanic or Latino/a
- I choose not to respond
- I do not know

Race: Please tell us with which race(s) you identify. Select all that apply:

- I only identify as Hispanic or Latino/a
- American Indian or Alaska Native [Include granular options]
- Asian [Include granular options]
- Black or African American [Include granular options]
- Middle Eastern or North African [Include granular options]
- Native Hawaiian or Pacific Islander [Include granular options]
- White [Include granular options]
- I choose not to respond
- I do not know

2. Add “Middle Eastern or North African” (MENA) as a New Minimum Category.

OMB seeks input on adding a new minimum reporting category for “Middle Eastern or North African” and to remove Middle Eastern or North African from the definition of the “White” reporting category.

Recommendation: We support the addition of a minimum reporting category for “Middle Eastern or North African.” We believe this will help improve the accuracy of race/ethnicity data by providing a category for how people identify. While the term “Middle Eastern or North

African” may not be consistently understood and acceptable among those with different experiences (e.g., 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.), this could be true of any of the minimum reporting or granular reporting category terminologies. For example, a child born to a Nigerian family in the U.S. may identify as “African American,” as “Nigerian,” as “Igbo,” or as any combination of the three.

3. Detailed Race and Ethnicity Categories.

OMB seeks feedback on the granular response options and whether they are inclusive and appropriate. OMB also seeks input on whether these detailed categories should be required of all agencies or if flexibility should be granted and how data on populations with small numbers should be managed and reported.

Recommendation: We believe OMB’s proposed revised granular race/ethnicity data standards are a great starting point as it allows organizations to collect more detailed data that can help inform more culturally appropriate care and work that reduces disparities while rolling data up to higher-level categories. We also agree with the addition of the “Other” response option to capture identities not contained in the close-ended granular response options. We prefer an option for “Other” as opposed to residual close-ended categories such as “Another Asian Group” so that individuals have an opportunity to provide information on their identity.

Recommendation: Even though we believe OMB’s proposed granular race/ethnicity categories are a great starting point, we request OMB provide more information as to how federal agencies intend to use the more granular level data. The context would help determine if the use of the proposed granular categories is appropriate at the federal level. It may be more appropriate to use granular data at the local level to inform culturally appropriate care and work that aims to reduce disparities. For example, the higher-level categories may be more appropriate for certain consumer surveys required by federal programs to ensure sufficient sample size per category to support accurate calculations, protect privacy, and increase response rates. Given the different appropriate uses of data, we recommend OMB engage the public and afford adequate time for public input to gather feedback each time granular data is proposed or suggested to ensure appropriate uses of data.

Recommendation: Unless OMB and federal agencies have specific reasoning and intent to use the outlined set of granular-level response options as proposed at a federal-level, we recommend that the extent or types of granular data collection be customizable at the local level rather than required nationally. Health insurance providers, hospitals, health care providers, and other health care organizations will likely be required to collect OMB's data standards as part of regulatory or accreditation requirements from Centers for Medicare & Medicaid Services (CMS), the National Committee for Quality Assurance (NCQA), Utilization Review Accreditation Commission (URAC), or the Joint Commission. If a standardized set of granular categories are required nationally rather than allowing for local customization, this could lead to an increased data collection burden on organizations who will be required to collect granular data that may not be relevant for their local service area while not collecting data that would be more relevant for their local service area. **We recommend OMB create data standards for granular response options but grant flexibility to implementing organizations that allows them to select which granular categories to include in their own data collection efforts based on what would be most meaningful for their local service area. Organizations should also be allowed to add additional response options that may not be included in OMB's standardized set of granular categories.**

Customizable granular response choices can be based on Census data or other localized data as to which populations are most prevalent in particular areas—whether state, county, zip code, or census tract. Granular data could still be rolled up to the higher-level reporting categories. This would help reduce data burden and ensure organizations do not have to keep a burdensome amount of data variables while ensuring local organizations have data that is most relevant to their service area. **OMB should provide updated guidance and standardized hierarchy as to how organizations should roll-up granular response options into the higher-level categories to ensure there is consistent mapping of data across disparate sources. To ensure organizations do not use the flexibility as an “opt-out” to collecting granular data, OMB and/or federal agencies could require organizations to collect a certain number of granular categories (e.g., five) but allow organizations to select which ones to collect based on local needs.**

Recommendation: The detailed response categories in OMB's proposed standards are not comprehensively inclusive but are a good starting point. Ironically, in efforts to be inclusive of more identities, having a limited set of granular response categories may appear to be even more exclusive as it could lead to confusion as to why some were included while others were not. We recommend OMB reconsider whether six granular response categories for each high-level reporting category is appropriate. In some categories, fewer granular categories may be

warranted while in others, more. For example, we question whether six granular response categories are necessary for some high-level reporting categories (such as White) and would need to know how the data would be used to make a fair determination. **As opposed to requiring six detailed categories *equally* across all the higher-level reporting categories, we recommend OMB use an *equity* lens to determine the appropriate number of granular response options for each higher-level reporting category that would be helpful to inform culturally appropriate care and work that aims to reduce disparities.**

Country of origin may not be the most appropriate or effective way to capture identity as more people identify simply as “American.” Depending on why the information is being gathered or used, it may be best to ask about the ancestry of the individual’s family. Challenges with this approach are that individuals may not know their family ancestry nor align with how the individual identifies. There are also other aspects to identity that are important to consider, such as tribal affiliation, language, and spirituality, among others. Some individuals may not identify with country of origin but with tribal ancestry (e.g., Igbo), ethnicity (e.g., Pashtun), or cultural affiliation (e.g., Jewish). These could appear as an additional level of granularity beyond the detailed response options in OMB’s proposed revision that are based on country of origin. If federal agencies or organizations are concerned about the additional data burden, this kind of information could be collected with an open-ended question, such as “Are there things about your culture or cultural identity that you would like us to know?” **To better inform structured data standards moving forward, we recommend OMB continue research in this area to monitor trends in terminology used for identity and changes in populations.**

Alternatively, the only way to avoid not being fully inclusive is to bypass structured close-ended granular response options altogether and instead simply ask an open-ended question on “How do you identify?” Such a question allows individuals to focus on how they perceive themselves as opposed to technical definitions in structured granular categories. As with any option, there are tradeoffs with using a single open-ended question. Analyzing and aggregating open-ended response options from paper forms will be time- and labor-intensive and could prove difficult given misspellings and responses that may not align with current race/ethnicity categories. Open-ended data could more easily be collected on digital forms where an individual types and drop-down lists or response options emerge based on what is typed, but these drop-down lists need to be extensively comprehensive to account for a wide array of identities.

Recommendation: In addition to data standards, we request OMB provide implementation guidance to ensure information on race and ethnicity is collected in a trusted, patient-centered way. For example, **we recommend OMB provide guidance and offer model introductory**

language that could be voluntarily adopted to explain why this information is being collected, how it will be used, how it will not be used, how it will be protected, and how it might be shared with others to build trust and encourage response. Introductory language should explain why specific granular response options are presented rather than others to help dispel potential confusion (e.g., largest sub-populations nationally are presented).

Recommendation: In its implementation guidance, we also recommend that OMB direct organizations to include a “I choose not to respond” response choice so that questions on race/ethnicity can be mandatory while still honoring an individual’s agency in voluntarily providing this information. We also recommend including a response choice for “I do not know” as some people truly may not know their race/ethnicity. Including response options for “I choose not to respond” and “I do not know” will help with evaluation and analysis since they can each be coded differently from a skipped question. Having data on the number of individuals who choose not to respond to questions on identity will also inform engagement efforts and be important for quality measures to demonstrate that attempts have been made to collect data but that some individuals prefer not to share such personal information or do not know how to answer. **Introductory language should also assure individuals that they are not obligated to select one of the granular options if that is not how they identify or if they do not feel comfortable sharing that level of information.**

Recommendation: We agree with OMB that data with small population sizes should be handled carefully to avoid non-representative data, privacy and confidentiality risks, unreliable and biased statistical estimates, and unfair comparisons on reducing disparities. We recommend OMB establish minimum denominators for reporting and analysis that are statistically significant to ensure this data is protected and appropriately used.

4. Terminology and Definitions

OMB seeks input on specific terminology that should be used for framing the race/ethnicity question, terminology that should be used to refer to specific populations that cross national borders, and whether the definitions of the minimum categories are accurate, respectful, and reflective of how people identify.

Recommendation: We recommend OMB frame the race/ethnicity question as “Please tell us with which race(s) and ethnicities you identify. Select all that apply.” We believe it is important to include the aspect of “how one identifies” in questions on demographics. This

question format can extend to other demographic data elements (e.g., gender, sexual orientation, pronouns, etc.).

Recommendation: We do not recommend that terms such as “multiracial,” “multiethnic” be used to describe individuals who select more than one race or ethnicity. In our stakeholder engagement, we found that people do not tend to identify as “multiracial” or “multiethnic” but instead identify by their specific races and ethnicities (e.g., “I am Filipino and White” or “I am Black and Cuban”). Terms denoting “mixed race” can also have a negative connotation to some individuals. **Similarly, we do not recommend that terms such as “transnational” be used to describe groups that extend beyond country borders.** Country borders are fluid and ever-changing; we therefore recommend that individuals be referred as how they identify, whether “Bantu,” “Hmong,” “Kurd,” etc.

5. Guidance on Implementation

OMB seeks input on how best to collect race/ethnicity data when self-identification is unavailable and techniques for bridging data collected previously with data collected moving forward.

Recommendation: We believe that self-reported data is the gold standard and should be encouraged. However, **when self-identification is unavailable, we recommend that indirect data methods (e.g., Bayesian analyses) be allowed to estimate race/ethnicity based on surname and geocoded address.** We acknowledge that there are limitations and inaccuracies with indirect data approaches—while they can characterize populations, they are far less accurate in identifying individuals. However, we believe that these are the best and most objective methods when undertaking large-sized population-level analyses to fill in gaps in data when self-identified data is unavailable. NCQA allows indirect data collection and analysis methods as part of its Health Plan Accreditation and Health Equity Accreditation as organizations transition to 100% self-reported data. However, **OMB should work with programmatic users like NCQA to create an indicator that distinguishes between data that was derived from indirect methods versus self-reported data** so that consumers and users of the data have visibility into how the data was established to inform appropriate uses of the data. **We do not recommend that observation be used to fill in gaps in self-identified data** given the subjective nature of observation.

6. Comments on Any Additional Topics and Future Research

OMB seeks additional input on the display order of the minimum categories, terminology of the minimum categories, and how best to collect data related to descent from enslaved peoples originally from the African continent.

Recommendation: We recommend ordering the minimum categories alphabetically to dispel notions of majority and minority populations. Alphabetical ordering will also be helpful if/when OMB decides to expand data collection for other demographic data (such as sexual orientation and gender) to dispel notions of what is considered “normal” because it is listed first as the most prevalent.

Recommendation: We recommend OMB reconsider the term “American Indian” given its historical roots in colonization. As part of our work to improve demographic data standards, we engaged a small sample of individuals from the Navajo, Wampanoag, and Zia tribes on terms they use to identify themselves. None use the term “American Indian”—preferring the terms “Native American” or “Indigenous”. We acknowledge that the terms “Native American” and “Indigenous” could cause confusion with those who identify as “Native Hawaiian”, “Pacific Islander”, or indigenous populations from other geographic areas other than the United States (e.g., Indigenous Central American or Indigenous South American). **We recommend OMB conduct additional research and engage tribal communities to better understand what terms they prefer to use to identify themselves** while reducing confusion from an identification and reporting strategy.

Recommendation: We agree with the importance of collecting data to better distinguish individuals and populations who are descended from enslaved peoples originally from the African continent. This will be important to better understand the generational impact of slavery, Jim Crow, and other racist and discriminatory policies and how disparities and inequities resulting from these policies could be rectified. However, while we agree with the importance of distinguishing these populations, **we do not recommend including a response choice such as “American Freedmen” or “American Descendant of Slavery” on a question asking for race/ethnicity. Instead, we would recommend this information be asked in a separate question** that could arise from branched logic for those who select “African American” on the race/ethnicity question. Given the lack of agreement and understanding on terminology and definitions to use to distinguish these populations, including such a term or definition on a question on race/ethnicity could lead to confusion and inaccurate data. **We recommend OMB conduct rigorous research and engagement with communities to determine the most**

appropriate terms and definitions before adding this category or question to OMB standards. From our engagement process, the term “Descendent of Enslaved Africans” emerged as the preferred term to be used on a separate question because it seemed to be more well understood than “American Freedmen” and more grammatically correct than “American Descendant of Slavery.”

***Recommendation:* We recommend OMB work to improve, standardize, and align other demographic data elements that are important to identity. These could include but are not limited to language, sexual orientation, gender, pronouns, disability status, relationship status, spirituality, and veteran status or military experience.** Standardizing and aligning these demographic data elements across the government and health care industry will help organizations better understand the populations they serve, provide tailored solutions, and advance health equity efforts.

***Recommendation:* We recommend OMB work with other federal agencies to align these data elements across the entire ecosystem to promote interoperable data exchange.** Aligning demographic data content and exchange standards at the ecosystem level across public and private health care stakeholders would facilitate electronic data capture, sharing, aggregation, and analysis. Collecting the information once and sharing it seamlessly across trusted entities would lead to a more efficient, effective, collaborative, and patient-centered system. Under such a system, the burden on patients, providers, and health insurance providers would be minimized, as information regarding sensitive identity and socioeconomic questions would only need to be collected once. It would also permit a broader understanding of populations and geographic differences with larger sample sizes, as well as the creation of national benchmarks.

Specific actions that federal agencies and other accrediting and regulatory bodies can take to improve standardization, alignment, and interoperability of demographic data include but are not limited to:

- Changes to the Centers for Disease Control and Prevention Public Health Information Network Vocabulary Access and Distribution System (PHIN VADS) codes and value sets for race and ethnicity to fill in gaps—both at a high-level and at a granular-level,
- Adoption as part of the U.S. Core Data for Interoperability by the Office for the National Coordinator for Health Information Technology (ONC),
- Inclusion in the 21st Century Cures Act Information Sharing regulation by ONC,
- Collection by CMS for enrollment in original Medicare, Medicare Advantage and Marketplace plans,

- Collection by States at enrollment for Medicaid or Children's Health Insurance Programs,
- Adoption by the Health and Human Services Department Office of Civil Rights (OCR) for civil rights and non-discrimination tracking,
- Recommendations to NCQA, URAC, the Joint Commission, and other accrediting bodies to include these as allowable data standards as part of its Health Plan, Health Equity, and Health Equity Plus accreditations,
- Recommendations to the National Association of Insurance Commissioners (NAIC) to encourage state Insurance Commissioners to adopt these data standards, and
- Counting such data collection efforts as Quality Improvement Activities under the Medical Loss Ratio requirements regulated by CMS.

Conclusion

The COVID-19 pandemic illuminated more clearly than ever the importance and urgency for having accurate demographic data to better identify and reduce disparities. We believe improving demographic data standards and aligning at an ecosystem level through such policy changes is crucial to advance equity. An aligned and standardized approach to interoperable demographic data will empower the health care ecosystem to collaborate on shared health equity goals, measure progress towards those goals, and better serve individuals and communities. With consistent and interoperable data standards, great strides can be made in reducing inequities while improving outcomes and minimizing the data burden placed on individuals and on the larger health care ecosystem.

Thank you for the opportunity to provide feedback on such meaningful work. We look forward to collaborating with the Administration and other stakeholders to improve and expand demographic data collection to improve care and reduce disparities. Please do not hesitate to contact me with any questions at (202) 778-3246 or dlloyd@ahip.org.

Sincerely,

A handwritten signature in cursive script that reads "Danielle A. Lloyd".

Danielle A. Lloyd, MPH
SVP, Private Market Innovations & Quality Initiatives