Creating Safe Space II: Capturing Race and Ethnicity Data

RAISE Community Workshop 5

Thursday, March 16, 2023, 2 – 3 PM ET

Summary

Overview of RAISE Community Workshop V

During RAISE workshop V, Susan Winckler, CEO of the Reagan-Udall Foundation, and RDML Richardae Araojo, FDA Associate Commissioner for Minority Health and Director of the Office of Minority Health and Health Equity, welcomed everyone. During the session we heard three presentations. First, Dr. Carla Rodriguez-Watson, RAISE PI, summarized our previous RAISE workshops and their connection to workshop V. Then, Ronnie Tepp of Pyxis Partners presented a case study in community engagement and inclusion specifically focused on people of Middle Eastern or North African (MENA) descent. Finally, Dr. Susan Jenkins from the US Department of Health and Human Services discussed Statistical Policy Directive (SPD) 15, and the proposals to change SPD 15. The session closed with a discussion moderated by Sarah Greene, Senior Advisor to the National Academy of Medicine.

Connecting the Dots: A Deeper Dive into Meeting People Where They Are

Carla Rodriguez-Watson, PhD, MPH

Director of Research, Reagan-Udall Foundation for the FDA

A special thanks to Dr. Ann Marie Meyer from the Department of Epidemiology at the University of North Carolina Gillings School of Public Health. As part of the expert panel and the 'champion' for workshop five, she helped identify topics and experts for the conversation.

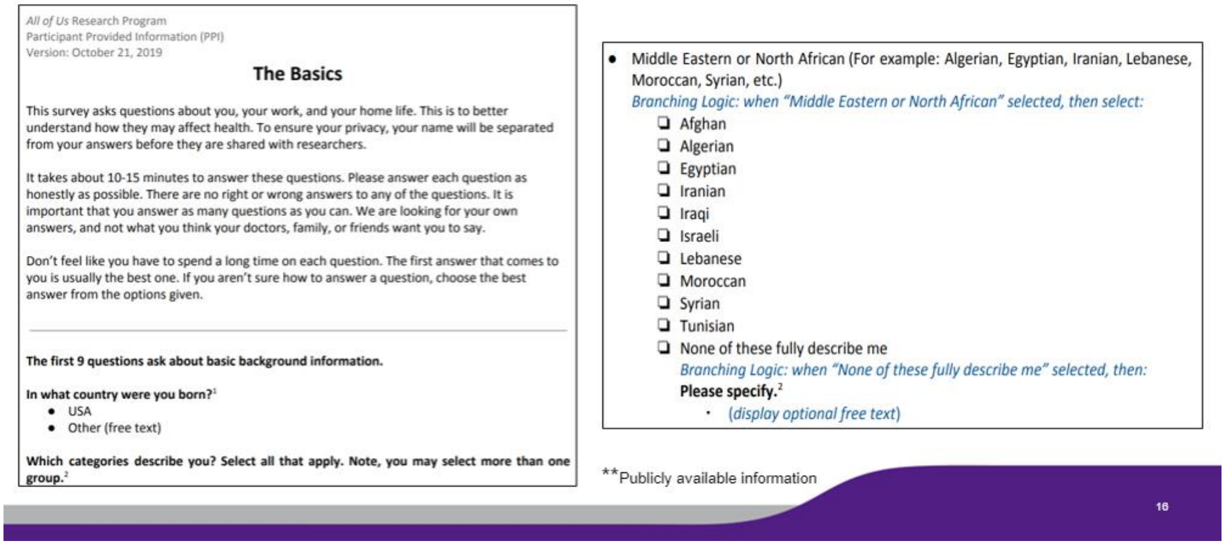
To level set, the RAISE project begins with the assumption that race and ethnicity (R&E) are critical for understanding population, health and the real-world utilization and performance and medical products across racialized groups. We acknowledge that R&E alone do not answer all the questions; but are critical to understanding population health and the real-world utilization and performance of medical products across racialized groups. As such, R&E are important to the FDA. The question of whether or when race or ethnicity is, or is not, the appropriate variable is not in the scope of the RAISE project. And though issues of access to care frame our conversation, addressing access to care is not the focus of RAISE. The focus of RAISE is the part of the data continuum that includes reporting collection, curation, and integration of R&E data (where the bolus of real-world data is generated/lost).

In workshop four, "Creating Safe Space I: Reporting Race 101", we discussed working with communities, to build technology that (1) enables data collection in the context of workflow and addresses the needs of individuals to see themselves in those categories and (2) allows communities to learn and engage in activities that are important to them. We were also introduced to how different standards for R&E are used, based on the population that is being served. For example, Epic systems shared how their clients on the west coast prefer more granular ethnic categories for Asian populations, while their East Coast clients may prefer more granular categories to describe the Hispanic and Latino ethnicity. In workshop five, we explored the topic of meeting people where they are through a case study of how to partner with communities to create categories that are meaningful to them. Then, we heard an overview of the current OMB categories and the rationale and process for revising them.

How Do You Check a Box When You Don’t See Yourself Represented? How Community Engagement Expanded Inclusivity in All of Us

Ronnie Tepp, Principal
Pyxis Partners

- Pyxis Partners works to reach, educate, and engage diverse communities in programs that advance health equity and improve health outcomes. This is accomplished by bringing program awareness and education to diverse communities through trusted messengers and creating and supporting networks of partners to reach community members to move them through their individual engagement journey.
- There is no separate check box for people of Middle Eastern or North African (MENA) descent on the US Census forms. As such, there is no direct way to produce a national account for MENA populations, and they are invisible in statistics that inform research, civil rights, protections, voting districts, housing policies, and medical research.
- There are currently an estimated 3.7 million Americans of MENA descent living in the United States. However, our understanding of this community's health is limited as the Federal Government characterizes people with MENA origins as White. Additionally, evidence suggests that many Arab Americans often select ‘other’ because they do not feel that White properly describes them. Either choice is a misclassification that impedes our understanding of health and disease risk in this population.
- Pyxis Partners joined the *All of Us* Research Program in 2016 as one of the first engagement awardees. Working closely alongside the Division of Engagement and Outreach, Pyxis Partners focused on building an engagement infrastructure through trusted messengers across the country in communities that have been historically underrepresented in biomedical research. One of the first partners in this network was the Arab Community Center for Economic and Social Services (ACCESS) in Dearborn, Michigan.
- As the *All of Us* program worked to develop the first set of participants surveys, ACCESS worked with the Pyxis team to advocate for inclusion of a MENA category in the survey. Their education and advocacy to the program and survey development team helped to make the case for inclusion of this category and alignment with the program's core values. The effort was successful, and the category was included in the basic survey (slide below).



All of Us Research Program
Participant Provided Information (PPI)
Version: October 21, 2019

The Basics

This survey asks questions about you, your work, and your home life. This is to better understand how they may affect health. To ensure your privacy, your name will be separated from your answers before they are shared with researchers.

It takes about 10-15 minutes to answer these questions. Please answer each question as honestly as possible. There are no right or wrong answers to any of the questions. It is important that you answer as many questions as you can. We are looking for your own answers, and not what you think your doctors, family, or friends want you to say.

Don’t feel like you have to spend a long time on each question. The first answer that comes to you is usually the best one. If you aren’t sure how to answer a question, choose the best answer from the options given.

The first 9 questions ask about basic background information.

In what country were you born?[1]
- USA
- Other (free text)

Which categories describe you? Select all that apply. Note, you may select more than one group.[2]

- Middle Eastern or North African (For example: Algerian, Egyptian, Iranian, Lebanese, Moroccan, Syrian, etc.)
  Branching Logic: when “Middle Eastern or North African” selected, then select:
  - ☐ Afghan
  - ☐ Algerian
  - ☐ Egyptian
  - ☐ Iranian
  - ☐ Iraqi
  - ☐ Israeli
  - ☐ Lebanese
  - ☐ Moroccan
  - ☐ Syrian
  - ☐ Tunisian
  - ☐ None of these fully describe me
    Branching Logic: when “None of these fully describe me” selected, then:
    Please specify.[2]
    - *(display optional free text)*

**Publicly available information

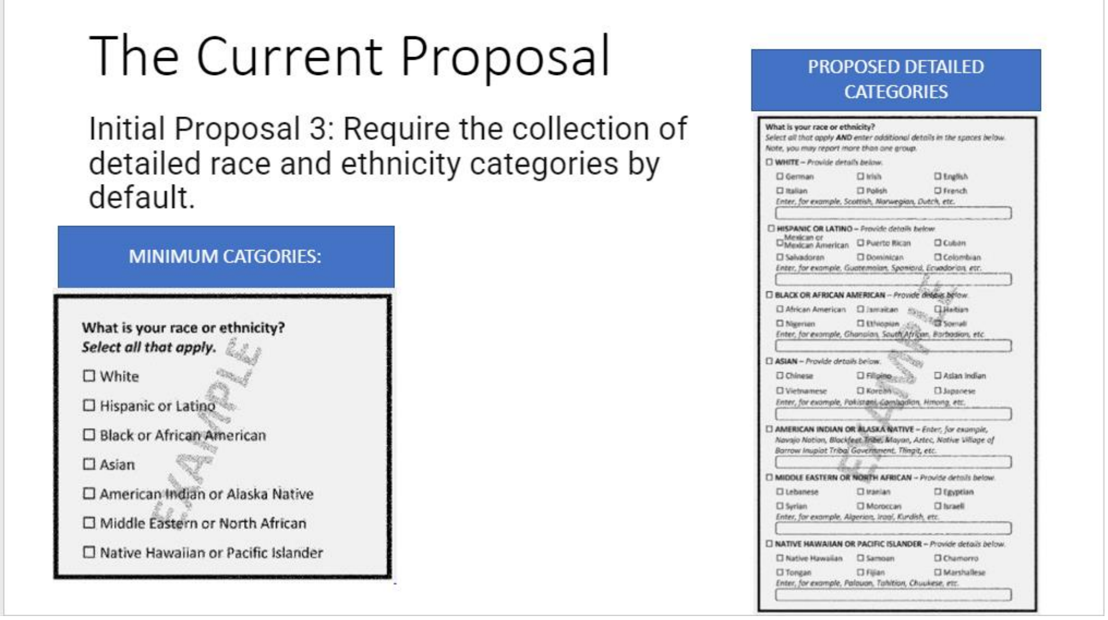
- This effort had an impact beyond simply securing a new category in a federal survey. It resulted in a community feeling seen and included:
    - The *All of Us* newsletter from March of 2021 features an article about an Arab American community member hearing about the inclusion of the MENA category in the survey and learning that ACCESS was involved with the program as a community partner. This community member then became interested in the program and ultimately joined as a participant, sharing her health data.
    - Another example of how this deepened advocacy across the MENA community is through social media posts. ACCESS was able to leverage their partnership with the *All of Us* Research Program to bring awareness to the program via Instagram, which led to increased enrollment by community members over time.
- This case study demonstrates how meaningful community engagement led a federal research program to create a new R&E category in a participant survey. This action enabled a community to feel seen and generated enthusiasm within the community for the program. This, in turn, led to increased and deeper engagement, broader awareness, and ultimately increased enrollment into the program. The increased enrollment of people of MENA descent led to population and community specific research, representing an important return of value to the community.
- When working with communities, it is critical to be able to demonstrate a return of value, and this 'value' can vary across people and communities. Individuals want to see themselves in the program and in the data. They want to see a box on a form that they can check that reflects themselves and their communities.
- To conclude, we should take this model and apply it to the conversations that are happening now, be it the 2030 census and other considerations, not only to make sure that communities see themselves reflected, but that R&E is captured in a way that ensures people are going to participate in research efforts.
- If interested, there is a publicly available *All of Us* research projects directory that now includes data from over 600,000 participants.

Defining and Redefining Federal Race and Ethnicity Categories: Insight from the HHS experience

Susan Jenkins, PhD, HHS Evaluation Officer
Director of the HHS Division of Evidence, Evaluation and Data Policy
US Department of Health and Human Services

- The Office of Management and Budget (OMB) Statistical Policy Directives identify minimum requirements for Federal principal statistical agencies when they engage in statistical activities. Directive No. 15: Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity The goal of having a standard is to make sure that there is comparability of R&E data across all federal data sets to link and compare data on specific populations to investigate questions, like disparities. The other purpose of SPD 15 is to maximize the quality of the data by ensuring comparability and commonality across the format, the language, the procedures for collecting the data, and that improves the quality of the data that are available.
- The OMB interagency technical working group (ITWG) that will evaluate relevant research, engage in a meaningful way with the American public and all impacted agencies, and develop a set of recommendations for revisions to Statistical Policy Directive 15. The group was initiated in June of 2022 by the chief statistician of the United States, and is working on a four-phase process:
    - The first phase in summer of 2022 was to set up the committee, the ITWG.

  - The second phase was to break up the committee into different working groups to look at different aspects of the SPD, and what questions that needed to be addressed in the potential proposal for a revision. This culminated in the publication of the Federal Register.
  - The third is the collection of information through that Federal register notice.
  - In the final phase, in the spring or summer of 2024, the statistician of the United States will determine if there will be a new statistical policy directive using the public comments and additional research.
- In making the proposals, OMB reviewed research on issues such as: combining the race and ethnicity questions into one question and creating a new minimum category for people who identify as Middle Eastern or North African (MENA) The Census Bureau has also done substantial research around what the R&E questions are, how people answer them, how people understand them, and whether the data are reflective of what is happening in the United States in terms of race and ethnicity and how people identify.
- The current proposal is comprised of five initial proposals:
  - Initial Proposal 1: Collect race and ethnicity information using one combined question. The Working Group proposes 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.
  - Initial Proposal 2: Add “Middle Eastern or North African” (MENA) as a new minimum category. The definition of the current “White” reporting category would be edited to remove MENA from its definition.
  - Initial Proposal 3: Require the collection of detailed race and ethnicity categories by default. The Working Group proposes that SPD 15 require data collection on race and ethnicity beyond the 7 detailed category levels, which include White, Black or African American, Asian, American Indian or Alaska Native, Native Hawaiian or Pacific Islander, the new Hispanic or Latino category, and MENA. The proposal is that detailed sub information to those categories would also be collected to get a better sense of race and ethnicity of individuals in the United States (figure below).



The Current Proposal

Initial Proposal 3: Require the collection of detailed race and ethnicity categories by default.

MINIMUM CATGORIES:

What is your race or ethnicity?
Select all that apply.
☐ White
☐ Hispanic or Latino
☐ Black or African American
☐ Asian
☐ American Indian or Alaska Native
☐ Middle Eastern or North African
☐ Native Hawaiian or Pacific Islander

PROPOSED DETAILED CATEGORIES

What is your race or ethnicity?
*Select all that apply **AND** enter additional details in the spaces below.*
Note, you may report more than one group.
☐ **WHITE** – *Provide details below.*
☐ German ☐ Irish ☐ English
☐ Italian ☐ Polish ☐ French
Enter, for example, Scottish, Norwegian, Dutch, etc.
☐ **HISPANIC OR LATINO** – *Provide details below*
☐ Mexican or Mexican American ☐ Puerto Rican ☐ Cuban
☐ Salvadoran ☐ Dominican ☐ Colombian
Enter, for example, Guatemalan, Spaniard, Ecuadorian, etc.
☐ **BLACK OR AFRICAN AMERICAN** – *Provide details below.*
☐ African American ☐ Jamaican ☐ Haitian
☐ Nigerian ☐ Ethiopian ☐ Somali
Enter, for example, Ghanaian, South African, Barbadian, etc.
☐ **ASIAN** – *Provide details below.*
☐ Chinese ☐ Filipino ☐ Asian Indian
☐ Vietnamese ☐ Korean ☐ Japanese
Enter, for example, Pakistani, Cambodian, Hmong, etc.
☐ **AMERICAN INDIAN OR ALASKA NATIVE** – *Enter, for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat Tribal Government, Tlingit, etc.*
☐ **MIDDLE EASTERN OR NORTH AFRICAN** – *Provide details below.*
☐ Lebanese ☐ Iranian ☐ Egyptian
☐ Syrian ☐ Moroccan ☐ Israeli
Enter, for example, Algerian, Iraqi, Kurdish, etc.
☐ **NATIVE HAWAIIAN OR PACIFIC ISLANDER** – *Provide details below.*
☐ Native Hawaiian ☐ Samoan ☐ Chamorro
☐ Tongan ☐ Fijian ☐ Marshallese
Enter, for example, Palauan, Tahitian, Chuukese, etc.

  - Initial Proposal 4: Update Terminology in SPD 15. The updating of the terminology is to remove of some outdated terms that are holdovers from the past and update the terms so that they are understandable and respectful of individuals. Examples of terms to be

    removed include the terms ‘majority’ and ‘minority’ as well as the terms ‘negro’ and ‘far east.’
  - Initial Proposal 5: Guidance is necessary to implement SPD 15 revisions on Federal information collections. If SPD 15 is revised, what kind of implementation guidance is going to be needed for Federal agencies and others in terms of bridging between the old standard and the new standard in terms of collecting more detailed R&E data and approaches for reporting on data where individuals select more than one race or ethnicity category?
- The HHS has been collecting information related to these proposals via listening sessions, surveys, and through a formal request for feedback from each Division. Each HHS division (CMS, FDA, CDC, etc.) has an opportunity to put forward their opinion about the proposals, and that will be sent to the ITWG and inform the final proposal.
- Based on the HHS listening sessions in October 2022 there is broad support for the four content related proposals described. Another theme is the need to consult with community representatives and stakeholders with lived experience and who represent the different racial and ethnic groups, about the proposal. The final theme is that there is need for plain language in whatever the standard ultimately is.
- Besides a bridge between the current and new standards, there are several other areas where clarification is needed, including:
  - Standardization, making sure that there are clear definitions,
  - Training, and educational materials for staff contractors and grantees,
  - Immediate expectations for the application of the new standard. What is the timeline? How is this going to work and a variety of other things?
  - Guidance for treatment of collected data, such as aggregation, data of indigenous people,
  - And provision of in-depth statistical information including suppression information, guidance for conducting trend analyses across previous and updated standard.
- Details on the initial proposals were released through the Federal register notice. The public can provide input via tribal consultations (*invitations sent to Tribal leaders and shared by the National Indian Health Board and the National Congress of American Indians*), town halls, bimonthly ITWG listen sessions (*to schedule a listening session, please send a brief email expressing interest to Statistical_Directives@omb.eop.gov*), and direct comments submissions.

Moderated Discussion

Moderator: *Sarah Greene, MPH*
Senior Advisor
National Academy of Medicine

Discussants:

- *Kathy Cronin, PhD (Deputy Associate Director, Surveillance Research Program National Cancer Institute)*
- *Susan Jenkins, PhD*
- *Ronnie Tepp*
- *Citizen Voice: Warren Whyte, PhD (VP, ERACE Lead, ConcertAI)*

The moderated discussion included questions from the chat as well as those posed by our moderator to further expand on the workshop’s presentations. Our discussion emphasized that, from a health perspective, we cannot understand differences in access to services received, or the outcomes of the services across race if we do not have good, consistent, high quality, valid measures of race. R&E

categories need to be reflective of the population, and consistent and aligned among data sources. This will ensure that the community is seen, heard, and counted. Highlights from the discussion:

- There are already 6,300 comments in the Federal Register. There is this massive effort in place to update the R&E standard, and it could not be more timely or important.
- The intended purpose for the updated SPD 15 categories is consistency across Federal data collections.
- Race is not a biological or anthropological construct. It is a socio-political construct. It is understood that it changes and morphs as do other social political constructs. It is an important construct. We cannot understand differences in access to services received or outcomes of those services across racial groups if we do not have good, consistent, high quality, and valid measures of race.
- Cancer registries are very dependent on the population estimates from the census but are equally dependent on what exists in the medical records. We need standards for how data is collected across medical practices, so that all registrars can access the same data and same race categories could be combined across different datasets (i.e., medical records and the census)
- There is a continued interest in getting more granular classifications that are more meaningful to individuals. With that comes issues with identifiability. There is a need to balance what can be put out while maintaining patient confidentiality. The needs and purposes of registries are different than those for research.
- Consistency and alignment across all R&E data points will help to inform the type of care that patients are receiving, and how inadequate care or lack of access to high quality care is contributing to health inequities.
- Combining R&E into a single category can have some utility as there can be some confusion among about the distinction between R&E. In these cases of confusion, the result is often misclassification.
- Aggregation and roll up versus keeping that individual person centered can be very challenging. When the numbers get too small, data is suppressed. There are groups that get upset because they feel like their data is always suppressed. If you roll data up to larger groups, it is not as meaningful.
- “Multiple race” could mean lots of different things which can be difficult to interpret in a meaningful way. It is critical to reflect the multiple race categories appropriately and respectfully for people in those categories to feel seen.
- There are lots of people and communities that are not going to respond to a federal register. There is a need for the Government entities to go to the community as opposed to expecting the community to come to the Government. It takes an investment and time, but it is important conversation to have to get those insights.
- When we talk about engaging diverse communities and hoping to capture R&E as well as nationality and ancestry in a way that is truly reflective of the individual that we are speaking to, we also have to be mindful of the fact that there is still some mistrust about how this information is being collected and utilized by health care providers.
- Just as there were several steps involved in the Pyxis case study that ultimately led to research participation, we need to think of collecting medical information step wise. If any one of those steps is missing with regards to proper education, proper trust relationship building, then what oftentimes happens is a reluctance on behalf of certain groups, particularly among people of color and other marginalized communities, to disclose information. We need to work collaboratively with community organizations to ensure that education, transparency and trust are built into each step of the data continuum to create a safe space to disclose R&E data.
- It is becoming more difficult to fully understand R&E as the granularity increases. Because of this, individuals are looking more towards some other variables such as education and financial status as additional data points. Not depending solely on R&E is probably an important step forward. Not that

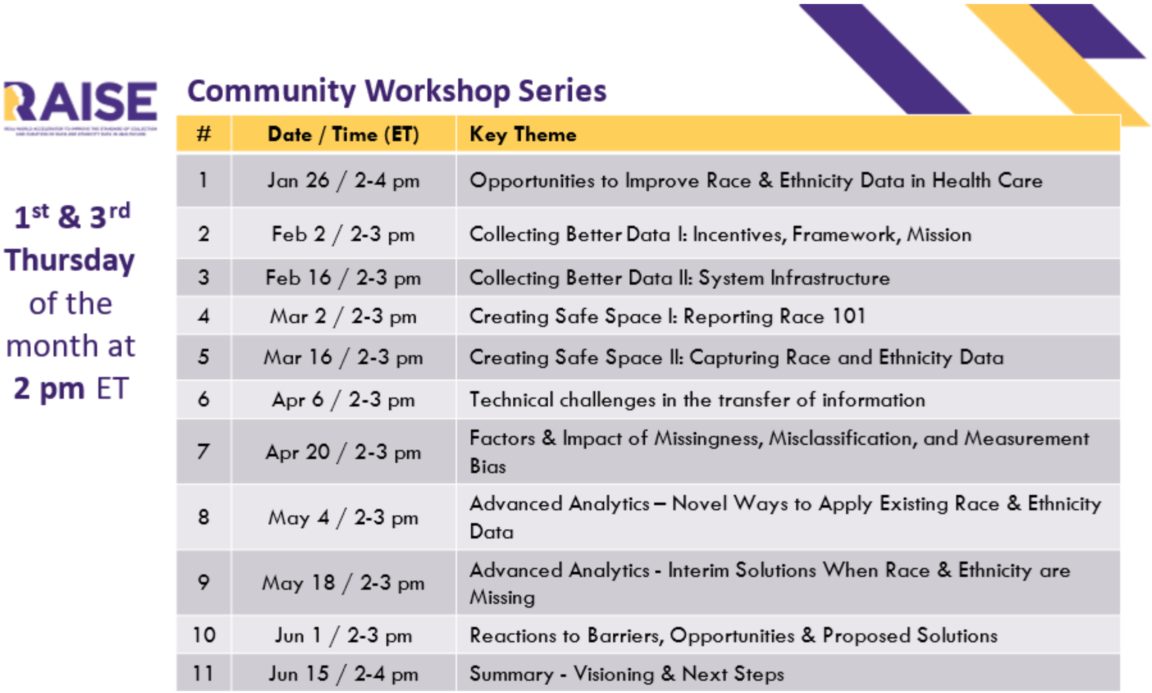
one replaces the other but trying to understand how they how they react together. The constellation of data points that makes the person unique.

- Although collecting additional data points (e.g., R&E, employment status, zip code, etc.) may improve care, there is also a risk of data privacy breach.
- When we think about health disparities across communities, it is multi-factorial in terms of the contributors. Having social determinants of health and other social economic construct can help in deciding how we address disparities and creating solutions that are tailored to the individual.
- There are a lot of differences between Japan, China, Korea compared to countries in Southeast Asia, for example. The differences in healthcare needs can vary for example, between Black and African Americans versus (relatively) newer immigrants from Africa. The proposed detailed R&E categories should help us understand the wide range of needs access and outcomes across and between racial groups.

Hot Takes and links from the Chat

- There was a question regarding prioritization of comments (e.g., whether comments from professional organizations will carry more weight than those from individuals). RESPONSE: All comments will be reviewed and considered. The ITWG is also reviewing research and conducting testing regarding approaches to data collection.
    - https://spd15revision.gov/content/spd15revision/en/proposals.html
    - https://www.federalregister.gov/documents/2023/01/27/2023-01635/initial-proposals-for-updating-ombs-race-and-ethnicity-statistical-standards: "How the 1997 Standards Define Race and Ethnicity: The categories developed represent a sociopolitical construct designed to be used in the self-reported or observed collection of data on the race and ethnicity of major broad population groups in this country and are not biologically or genetically based."
- https://www.nationalacademies.org/news/2023/03/researchers-need-to-rethink-and-justify-how-and-why-race-ethnicity-and-ancestry-labels-are-used-in-genetics-and-genomics-research-says-new-report
- Race is a socio-political construct, and its implications are long standing. The collection of detailed R&E data should help us to understand the impact of race on health outcomes.
- Intersectionality is important. R&E, along with history, politics and policy influence who has access to social determinants of health (including education and income) and health disparities.

Please join us for future RAISE Workshops:



RAISE

Community Workshop Series

1st & 3rd Thursday of the month at **2 pm** ET

#	Date / Time (ET)	Key Theme
1	Jan 26 / 2-4 pm	Opportunities to Improve Race & Ethnicity Data in Health Care
2	Feb 2 / 2-3 pm	Collecting Better Data I: Incentives, Framework, Mission
3	Feb 16 / 2-3 pm	Collecting Better Data II: System Infrastructure
4	Mar 2 / 2-3 pm	Creating Safe Space I: Reporting Race 101
5	Mar 16 / 2-3 pm	Creating Safe Space II: Capturing Race and Ethnicity Data
6	Apr 6 / 2-3 pm	Technical challenges in the transfer of information
7	Apr 20 / 2-3 pm	Factors & Impact of Missingness, Misclassification, and Measurement Bias
8	May 4 / 2-3 pm	Advanced Analytics – Novel Ways to Apply Existing Race & Ethnicity Data
9	May 18 / 2-3 pm	Advanced Analytics - Interim Solutions When Race & Ethnicity are Missing
10	Jun 1 / 2-3 pm	Reactions to Barriers, Opportunities & Proposed Solutions
11	Jun 15 / 2-4 pm	Summary - Visioning & Next Steps