Improving Race and Ethnicity Data in Health Care

RAISE Community Workshop 1

Thursday, January 26, 2023

SUMMARY

Overview of RAISE Community Workshop I

A two-hour workshop on January 26, 2023, launched the RAISE workshop series. Opened by Susan Winckler, CEO of the Reagan-Udall Foundation. RDML Richardae Araojo, FDA Associate Commissioner for Minority Health and Director of the Office of Minority Health and Health Equity provided opening remarks. During the session we heard five presentations on the importance of and barriers to reporting, collecting, and curating race and ethnicity data in health care. First, Dr. Georges Benjamin of the American Public Health Association and Reagan-Udall Foundation for the FDA Board of Directors provided an overview of why race and ethnicity data matters. Next, Ms. Nicole Evans and Dr. Hemi Tewarson of the National Academy for State Health Policy presented on how states can improve health and ethnicity data. Dr. Gary Puckrein of the National Minority Quality Forum then presented on the need for data standards and solutions, followed by Dr. Oscar Benavidez of Massachusetts General Hospital who presented a mixed-methods approach to advance health equity. Ms. Deanna Fulp of Blue Cross Blue Shield of Massachusetts presented on the payor side of collecting race and ethnicity data. Claudia Withers, JD, of Civil Rights Corp then joined the previous presenters for a profound discussion moderated by Dr. Michelle McMurry-Heath. The session closed with an overview of RAISE presented by Dr. Carla Rodriguez-Watson of Reagan-Udall Foundation for the FDA.

Keynote: What Gets Measured - Gets Done: Why Data by Race & Ethnicity Matters
Georges C. Benjamin, MD
Executive Director, American Public Health Association
Board Vice Chair and Chair, Food & Nutrition Committee, Reagan-Udall Foundation for the FDA

The on-point quote by Peter Drucker: "What gets measured gets managed" was used to underscore that measuring by race doesn't just matter for historically underserved populations - everyone can benefit as part of that process.

- **Example 1:** The overall infant mortality rate in Maryland from 2011-2020 looked to be stable (if not slightly declining). But when disaggregated, the rate was noticeably increased in Black women. This also happened several years ago as well. When public health researchers then looked further into the trend, they identified that the increase was driven by older, insured, Black women who initiated fertility treatment (fertility treatment is a risk factor for infant mortality); and that a similar, smaller, shorter increase also occurred in White women – and was missed.

- **Example 2:** Social determinants drove COVID-19 inequities as those working outside the home tend to be disproportionately racial and ethnic minorities with higher prevalence of chronic disease and

lower socio-economic status. Understanding social determinants of disease, signaled by race, helps address the inequities that exist across all people.

Health inequities arise due to differences in access to acute and chronic care, differences in the quality of care received within the health care setting, individual behavioral differences in our approach to health or health care and societal things such as political, economic, or environmental exposures (includes racism and discrimination). The examples cited above illustrate why race and ethnicity are important descriptors for public health; and signal the need for additional inquiry. The arguments for and against collecting race and ethnicity data were illustrated in two slides:

Arguments Against Reporting, Curating, Collecting Data By Race / Ethnicity

- Race is a social construct and has little to no biological basis, so they don't matter in a therapeutic sense
  – Race / ethnicity do serve as a rough proxy for risk of certain diseases (Sickle cell anemia, Thalassemia, Cystic fibrosis, etc.) & response to therapeutics (antihypertensives)
  – The social determinates of health are influenced by race
  – Access & quality of care are influenced by race & ethnicity
- Data collected by race / ethnicity can be misused to discriminate or stigmatize
  – Agreed as can any dataset- We need safeguards to minimize this

APHA

Arguments Against Reporting, Curating, Collecting Data By Race / Ethnicity

- Misclassification, grouping of like individuals into broad racial/ethnic categories, makes analysis inaccurate
  – For example: Groupings of Hispanics and Asian individuals into singular groups. Cultural and language differences abound that impact clinical outcomes
- Privacy concerns especially with small disaggregated numbers
  – Disaggregation can be a problem and must be managed
- Promotes racial / ethnic conflict and racial bias
  – A form of misuse of the information that needs to be addressed

APHA

Overall, accelerating improvements in the standard of collection and curation of race and ethnicity data in healthcare is complicated and does require a complex set of actions. This is a call to action.

Opportunities to Improve Race and Ethnicity Data in Health Care
Nicole Evans, MPH and Hemi Tewarson, JD, MPH
National Academy for State Health Policy

The National Academy for State Health Policy (NASHP) is a nonprofit organization committed to working closely with states to identify new ideas, new approaches and new steps to advance health policy.
- The COVID-19 pandemic increased the interest in race and ethnicity data as states needed quality data to:
  o identify unmet needs from disparities that existed before the pandemic that have been exacerbated;
  o understand potential policy solutions to target disparities and unmet needs to improve the health of all residents; and
  o to prioritize and take next steps when progress is being made.
- State challenges to collecting race and ethnicity data included inconsistent standards, reliance on outdated systems, missing and incomplete data, lack of public trust and data sharing, inability to disaggregate data.
- Overall, the pandemic advanced state strategies to strengthen race and ethnicity data:
  o Tennessee and South Carolina leveraged their Medicaid managed care organization contracts to collect these data.

- o In Utah, a uniform data collection standard was developed. New dashboards are being developed statewide that are transparent with the communities they serve to encourage race and ethnicity data collection as well as race and ethnicity data sharing.
  - o North Carolina created an interactive map that uses race and ethnicity data along within social drivers of health indicators to illustrate the social determinants of health impact to populations and communities during the pandemic. By layering this dashboard with vaccine uptake, they can see where the vaccine is needed most.
- Leadership buy-in at the legislative level is key for states to be able to move forward with health policy that addresses collection of race and ethnicity data.
  - o Example: Oregon's legislature requires providers to collect race, ethnicity, language, and disability data to be shared with the Oregon health authority to inform state policy.
- Where state legislatures are not inclined to collect race and ethnicity data, states have leveraged innovative ideas around collecting race and ethnicity data transparently and sensitively with cultural humility and communicated where this information is going along with its purpose.
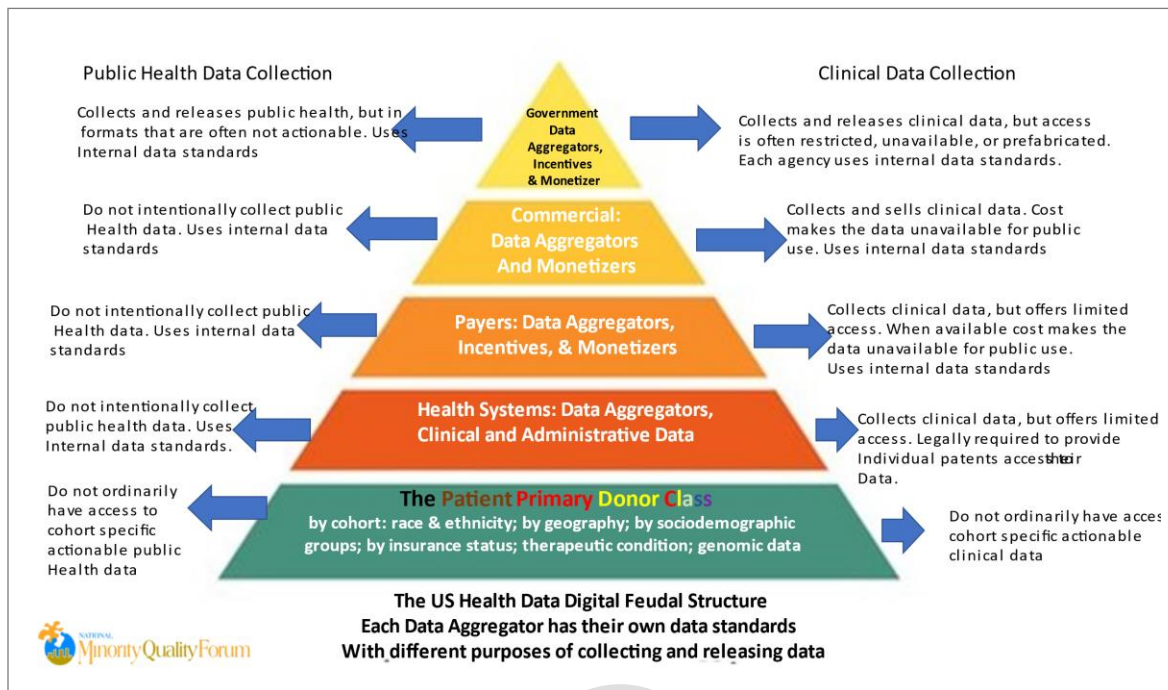
Health Data Standards in the Age of Digital Feudalism and the Health Equity Movement
Gary A. Puckrein, PhD
National Minority Quality Forum

Over the last 20 years, the National Minority Quality Forum has collected heath data and now has a database of more than 5 billion patient records across just about 100,000 different clinical conditions.

- Dr. Puckrein described the ways the current structure of health care data is misaligned with the goals of public health. Where the source of the data (i.e. the people and clinicians who collect and use the data to care for patients) is the foundation of the pyramid--but access to these data–and the information that can be gleaned from those data - are limited, primarily by the layers on top.
- Different sets of rules, different data, and different standards create challenges for those trying to collect data that can be used in the community.

Public Health Data Collection

Collects and releases public health, but in formats that are often not actionable. Uses Internal data standards

Do not intentionally collect public Health data. Uses internal data standards

Do not intentionally collect public Health data. Uses internal data standards

Do not intentionally collect public health data. Uses Internal data standards.

Do not ordinarily have access to cohort specific actionable public Health data

Clinical Data Collection

Collects and releases clinical data, but access is often restricted, unavailable, or prefabricated. Each agency uses internal data standards.

Collects and sells clinical data. Cost makes the data unavailable for public use. Uses internal data standards

Collects clinical data, but offers limited access. When available cost makes the data unavailable for public use. Uses internal data standards

Collects clinical data, but offers limited access. Legally required to provide Individual patents access their Data.

Do not ordinarily have access cohort specific actionable clinical data

Pyramid levels (top to bottom):
- Government Data Aggregators, Incentives & Monetizer
- Commercial: Data Aggregators And Monetizers
- Payers: Data Aggregators, Incentives, & Monetizers
- Health Systems: Data Aggregators, Clinical and Administrative Data
- The Patient Primary Donor Class by cohort: race & ethnicity; by geography; by sociodemographic groups; by insurance status; therapeutic condition; genomic data

The US Health Data Digital Feudal Structure
Each Data Aggregator has their own data standards
With different purposes of collecting and releasing data

NATIONAL Minority Quality Forum

- The US Health Data Digital Feudal Structure is a pyramid. Each data aggregator within a layer of the pyramid has its own data standards that are aligned with that layer's specific purpose of collecting and releasing data. Although the bottom layer is closest to the patient (for whom the data are intended to impact), access to the data is really concentrated towards the top among commercial data aggregators who monetize the data. At this point, cost limits access to the data for public health purposes.
- This paradigm creates challenges because when data standards and tools are built to support the needs of businesses at their respective level (siloed), then data must be integrated across layers to construct a complete picture of the patient and the community. Data integration requires a complex set of activities to manage data privacy and incompatible standards to produce datasets that can inform patient and community health.
- Data collection, infrastructure and standards need to be re-imagined to provide a complete view of the patient to support population health (Flashback: For those who attended the Aug. 20, 2020 meeting of the Diagnostics Evidence Accelerator you will recall how "Heidi's journey" for her COVID symptoms illustrated the data patchwork needed to understand her COVID-19 diagnostic tests). It is this view that should inform what the data collection, infrastructure & standards should look like to support patient care and population health.

Improving Race, Ethnicity, and Language Data Collection through Innovation: Mixed Methods Approach to Advance Health Equity
Oscar Benavidez, MD, MBA, MPP
Massachusetts General Hospital

Dr. Benavidez described an exciting mixed methods approach to identify, categorize, and prioritize opportunities to improve the collection of race, ethnicity, and language (REaL) data in a large hospital system; the taxonomy of barriers they identified in REaL data, and then described the development of a pilot study and presented preliminary findings.
- Missing REaL data is a ubiquitous problem that limits the ability to achieve equity and dismantle racism in health care. In 2021, the Health Policy Commission in Massachusetts identified REaL data as a critical element to dismounting racism and other longstanding and inequities. Unfortunately, REaL data are missing for the majority of patients.
- The three aims of the project: (1) to engage across functional groups of stakeholders to understand barriers to the collection of REaL data; (2) to identify patient level risk factors for missing data; and (3) to co-design a pilot to learn from and improve REaL data collection.
- The qualitative portion of the project engaged subject matter expects to conduct failure mode analysis to identify and categorize barriers to data collection – eventually creating a taxonomy. They also surveyed administrative staff for their perceptions of the relevance and feasibility of REaL data collection within their workflow.
- The quantitative portion of the project examined the administrative database from a large Academic Medical Center (AMC) to identify clinical areas with high rates of missing REaL data and develop a multivariate model to identify patient-level variables associated with missing REaL Data.

QUALITATIVE & QUANTITATIVE RESULTS
- Low engagement with the health system was associated with missing data. The results suggest that If REaL data are to be collected, different strategies are needed for different populations.

- Outpatient areas with the biggest gaps in data collection were the pediatric groups, which were used for the pilot. The pilot included a cost-estimate to scale system-wide.
- Despite acknowledging the importance of health equity, the patient-facing stakeholders (front desk, scheduling, registration) felt the collection of race, ethnicity, and language was either not relevant to their job or it belonged to somebody else to do. These stakeholders were engaged to assist in developing the race, ethnicity, and language data collection pilot.

PILOT & RESULTS
- Development of the pilot included an examination of process and workflow. Results were used to redesign workflow, user interface, and develop an education module. The pilot was implemented in outpatient pediatric cardiology and compared to all other pediatric outpatient areas.
- Interim pilot results showed a 60% decrease in the rate of missing data from baseline and 35% decrease compared to control group after two months.

CONCLUSIONS
- A problem-focused approach with diversified subject matter experts and scalable solutions is important.
- Data analytics are needed to understand the populations, feasibility of aspirational targets, the number of discrete solutions, and costs for implementation. Ongoing sensing mechanisms are needed to identify areas in need of improvement if data collection rates fall.
- A new and active REaL data collection infrastructure is needed to examine the existing categorizes and relationship to important outcomes and consider changing categories as data suggests.
- The efficacy and cost of these efforts should be assessed and compared to other efforts to ensure investment in effective strategies.

Advancing Health Equity Data Collection: Race and Ethnicity
Deanna Fulp, MBA
*Blue Cross Blue Shield of Massachusetts*

The social and political climate around COVID-19, coupled with the leadership at Blue Cross Blue Shield of Massachusetts (BCBSMA), created a moment in time to act swiftly on health equity.
- BCBSMA CEO and board of directors decided to elevate equity within their strategic framework. They placed it on equal footing with affordability and quality of care.
- BCBSMA has not encountered legal barriers to collecting race and ethnicity data for their purposes.
- Collected race and ethnicity data for only two reasons: to measure inequities and care, and to minimize inequities in care.
- BCBSMA produced equity audits for provider organizations that are part of the alternative quality contracts (value-based care payment models). Though blinded, providers can see their equity performance compared to other providers.
- To be accountable and transparent to their members, BCBSMA publishes an annual Health Equity Report on their website.
- Additionally, BCBSMA has added equity to the Alternative Quality Contract (AQC) triad by adding Pay for Equity (P4E) incentives (2023), the Equity Action Community Institute for health care Improvement (2021) and Health Equity Grants to contracted provider organizations that participated in the Equity Action Community.

- BCBSMA has several sources of data on race and ethnicity, including those from member self-report (gold standard), provider and government sources, imputed data and is working towards collecting race and ethnicity on enrollment or employment files.
- The gold standard member self-report is done through an online survey which was launched two years ago. 18.5% of members have provided their race and ethnicity data through this channel with only three member complaints. The survey was designed using behavioral science tools and the two purposes of the data use are shown. The survey is optional, and members can choose to bypass completing it.

MEASURING INEQUITIES IN CARE
BCBSMA has produced equity audits for provider organizations and for publication on our website
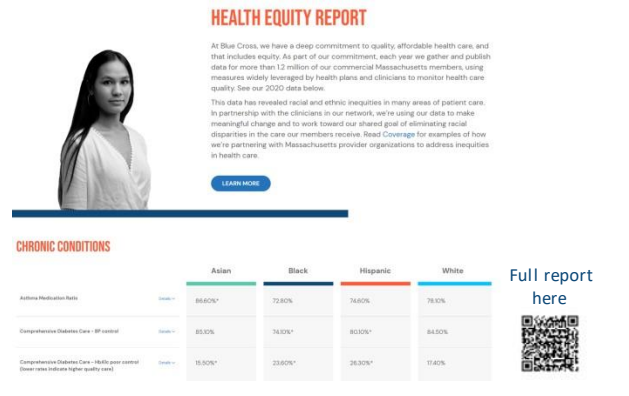
- BCBSMA also launched a paper survey to reach those not responsive online by doing a multi-arm, randomized, controlled test of messages informed by behavioral science and insights. The survey included a writing implement and that increased response rates but was also expensive.
- BCBMA chose the Fast Healthcare Interoperability Resources (FHIR) standard to enable alignment because they realize "we're not in this alone." As part of their promise to their members and to anyone who gives them data, they commit to using data only in ways that are consistent with their framework (to measure inequity and reduce inequity). And as part of their pay for equity contracts, BCBSMA strongly encouraged those that are thinking about, or working on a data collection of race and ethnicity to share it and to be part of a community.
- BCBSMA's work to improve equities is a work in progress, and all barriers are simply issues to overcome.

Moderated Discussion

Moderator: Michelle McMurry-Heath, MD, PhD
Discussants:
Citizen Voice: Claudia A. Withers, JD, Civil Rights Corps
Oscar Benavidez, Georges Benjamin, Nicole Evans, Deanna Fulp, Gary Puckrein

The moderated discussion took questions from the question-and-answer chat as well as those posed by the moderator to further expand on the workshop's presentations, and to probe about community trust and building cultural competence. The moderated discussion also reinforced that the collection of race

and ethnicity data is crucial and that we should not let "perfect be the enemy of the good" (we're never going to be satisfied with the "categories"). We must act now to improve the reporting, collection, and curation of race and ethnicity data. Hot takes from the discussion:

- As we think about community–make sure that we have ongoing conversations with the folks that we are trying to reach.

- Make sure there is a built-in constituency–may not talk about race/ethnicity or issue of division, but we in the community.

- Other countries have no problem having national data collection. We have been most successful when we acknowledge differences and bring those differences together.

- Think about how we can shift the focus to wellness when we talk about health equity. Racial and ethnic minority populations often only access health care in moments of crisis.

- There are so many misperceptions about who's driving social spending and who's on Medicaid. We need to collect race and ethnicity data, analyze data by race and ethnicity, and report results back to the public to check our assumptions.

- We have our current race and ethnicity categories, but it's unclear that they are the correct categories. [Note: OMB released on the very same day the Initial Proposals to Revise the Federal Race and Ethnicity Standards.]

- We are never going to stop talking about the categories, but we know there are inequities. It is a fraught, difficult environment, but we have to start somewhere. If this was any other industry, we probably would have done much more by now.

Closing: RAISE: What it is and where we're going!
Carla Rodriguez-Watson, PhD, MPH
Director of Research, Reagan-Udall Foundation for the FDA
Principal Investigator, RAISE

The Reagan-Udall Foundation for the FDA, in collaboration with FDA's Office of Minority Health and Health Equity, established the Real-world Accelerator to Improve the Standard of collection and curation of race and Ethnicity data in health care (RAISE) was developed in response to the large volume of missing race and ethnicity in real-world data (RWD). This missingness limits the real-world assessment of utilization and performance of medical products and care interventions, as well as the ability to utilize RWD to understand populations at risk.

RAISE has 3 aims:
Aim 1. Gather to create opportunities for Collective Learning and Community Capacity Building
**Aim 2. Create a Tool** to prioritize solutions
Aim 3. Evaluate the impact of the project to initiate a change in practice

Aim 1 is where we are today–gathering to share ideas, frameworks, solutions, implementation challenges and successes–because we know they exist!

Good Solutions Exist
...but require identification and investment.

We are bringing together **leaders** in the health care space because we believe the infrastructure investments needed to better capture and curate race data for evidence generation will be incentivized if systems see that it improves patient health, delivery of care and brings value, not just lower cost, to their patients and members, and ultimately value to the system.
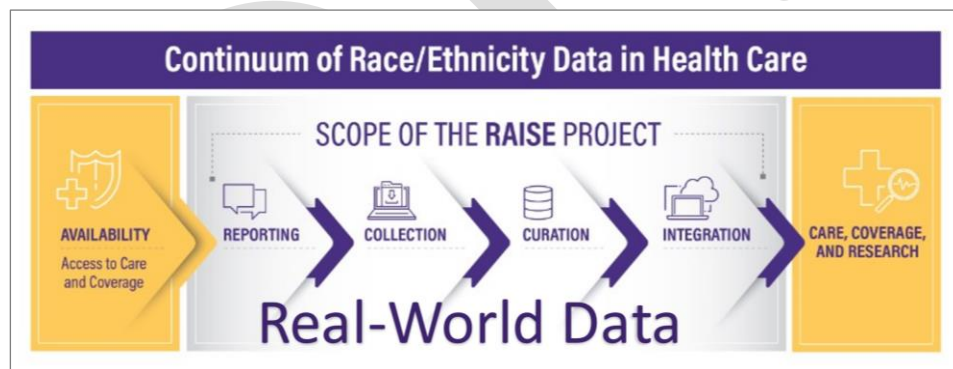
RAISE will host 11 community workshops on the 1st and 3rd Thursday of every month from February through June 2023, 2-3 PM ET. We will explore topics along the health care data continuum that may represent challenges in the reporting, collection, curation, and sharing of race and ethnicity data.

Aim 2: As we go through the workshop series, we'll identify barriers and solutions and develop a tool to prioritize short, medium and long-term solutions:
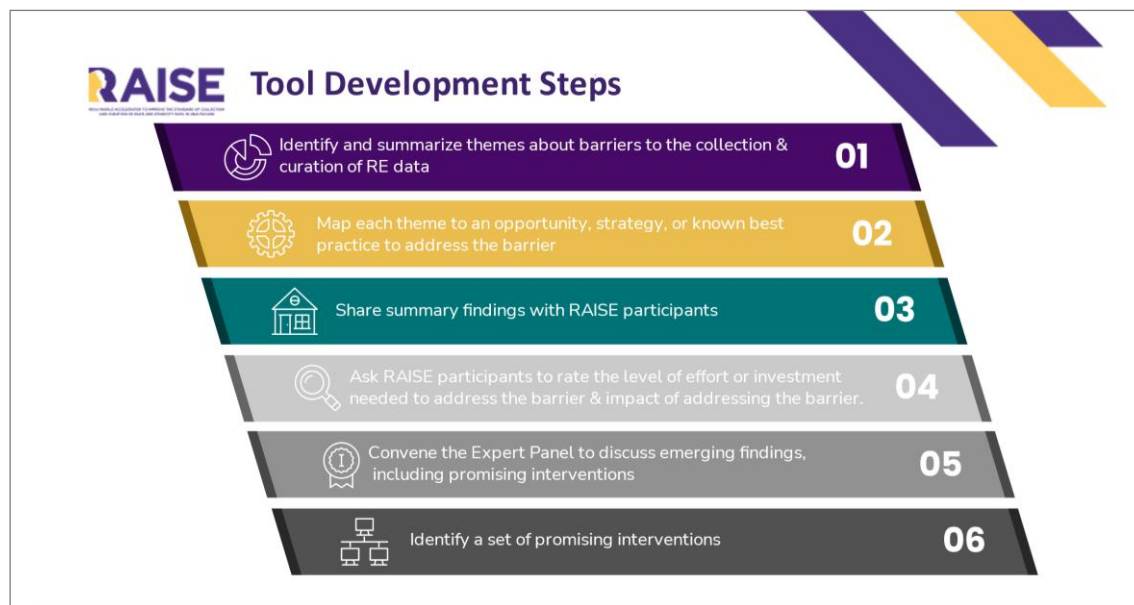
Aim 3: We will evaluate the impact of the project to initiate a change in practice using a pre and post qualitative design. Some of you may have participated in our baseline interviews.
Join us for the next conversation…and the one after that!

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