Counting a Diverse Nation:
Disaggregating Data on Race and
Ethnicity to Advance a Culture of Health

PolicyLink

Acknowledgments

Cover photo: Courtesy of Southeast Asia Resource Action Center (SEARAC)

PolicyLink is a national research and action institute advancing racial and economic equity by **Lifting Up What Works®**.
www.policylink.org

Counting a Diverse Nation:
Disaggregating Data on Race and
Ethnicity to Advance a Culture of Health

Victor Rubin
Danielle Ngo
Ángel Ross
Dalila Butler
Nisha Balaram

"Haitian Flag Raising Ceremony" in Paterson, New Jersey.
(*B.C. Lorio* *is licensed under* CC 2.0)

Contents

Preface from PolicyLink

As the U.S. population becomes ever more diverse, our country is strengthened in many ways by its remarkable array of races, ethnicities, cultures, and languages. For us to fully benefit from that diversity, to fully unlock our promise, we need to be able to document, measure, and appreciate the economic conditions and the nuances of life experience of people of all backgrounds. The inequities in health outcomes that persist in American society do not show up just at the level of broad racial categories but as disparities experienced by more specific groups. When Americans of Southeast Asian and Pacific Islander backgrounds make the case to be recognized, not obscured within a much larger, undifferentiated Asian American category, they are seeking information about their own community and also to be a more visible part of the American fabric. When members of individual American Indian nations, or refugees from a Middle Eastern country, or immigrants from nations in Central America or Africa, seek to be counted and have their life circumstances documented, theirs is a call for visibility and full inclusion as well.

Creating that visibility is the power of disaggregated data when it is meant to advance health equity. It is the basis for systemic change and the empowerment of groups that have often not been heard. That is why, at PolicyLink, we have been honored to work with the Robert Wood Johnson Foundation to bring clarity to the issues, ideas, and best practices in data disaggregation and to carry those insights to new audiences.

The researchers, advocates, and policymakers who participated in this project have proposed and are ready to move forward with practical actions that can improve the quality, availability, and utilization of disaggregated data. Federal policies about the collection and use of data, from the census to the array of health surveillance surveys, will need to be more responsive to the need for disaggregation by race and ethnicity. Well thought out, scientifically sound proposals for such changes are now available. States, several of which have already taken important steps, will need to reconfigure key data sources about health, education, and other services to reflect their growing diversity. And for all levels of policymaking, good ideas and the drive for positive change will continue to come from local, grassroots leaders in health equity, immigrants' rights, racial justice, and other struggles. This report is intended to support all those who are seeking to bring about those changes.

Angela Glover Blackwell
CEO

Preface from the Robert Wood Johnson Foundation

The Robert Wood Johnson Foundation has a bold vision for a nation that strives together to build a Culture of Health, enabling all in our diverse society to lead healthy lives, now and for generations to come. The Foundation believes that this requires making health a shared value such that everyone deserves the opportunity for good health, supporting cross-sector collaboration so that everyone sees the interrelationship between health outcomes and social determinants of health as well as health care, creating healthier and more equitable societies where everyone has equitable opportunities to make healthy choices, and improving the integration of prevention and health-care services. And in order to achieve a healthy nation, we need to know where there are health disparities so that resources and efforts can be targeted appropriately. I have often heard the phrase, "what gets measured is what gets changed." However, what mediates that relationship, what really drives that relationship, is the fact that what gets measured is what gets the resources so that change can happen.

In terms of health disparities, research often focuses on major demographic characteristics like race and ethnicity. The United States comprises a diverse ethnic/racial population, which is typically categorized into five distinct groups: African-Americans/Blacks, Latinos/Hispanics, American Indians/Alaska Natives, Asian Americans/Native Hawaiians/Pacific Islanders, and Whites. Despite these broad categories, however, significant variation exists in the histories and experiences within each of those ethnic/racial groups, all of which can influence health outcomes as well as risk factors for poor health and well-being.

- There are 573 federally recognized Indian nations in the United States. In addition to members of these tribes differing ethnically, culturally, and linguistically, they can live on or off reservation, which influences their access to health services and other major resources.

- Asian Americans and Pacific Islanders accounted for more than 17 million individuals and nearly half of all refugees who arrived in the United States between 2000 and 2010. Those coming to this country as refugees have vastly different life experiences from those who voluntarily immigrated.

- Forty-two million people in the United States self-identify as Black or African American. While most of them have lived in the United States for generations, more than three million are immigrants, mostly from Africa or the Caribbean.

- The Hispanic/Latino population makes up about 16 percent of the U.S. population, and about three-quarters self-identify as Mexican, Puerto Rican, or Cuban, which represent strikingly different cultures and histories.

- Individuals self-identifying as White represent more than two-thirds of the U.S. population but the cultural diversity of three continents (North America, Europe, and Africa).

These demographic statistics were drawn from either the 2010 U.S. Census or the American Community Survey, two of the most comprehensive surveys available in terms of ethnicity/race data collected. That level of detail, however, is uncommon. Even when more detail might be collected, data are often aggregated into the five broad ethnic/racial categories when they are analyzed or reported due to limitations in sample size. Aggregating data to those higher levels ignores the significant variation that exists within those broad categories and limits the field's ability to target its resources where they are most needed, to the communities experiencing disparities. While some of the groundwork has already been laid for this research, the Robert Wood Johnson Foundation commissioned a multiphase process that first aimed to level-set across ethnic/racial populations in terms of what is known and then sought to identify needs, gaps, and next steps for the field. The project brought together a diverse set of experts, demographers, practitioners, decision makers, and advocates to encourage leaders across sectors and ethnic/racial groups to talk to and work with each other. This report represents the culmination of those activities and the Foundation's first step toward identifying solutions for improving ethnic/racial data disaggregation with the aim of promoting health equity in our nation.

Tina J. Kauh, PhD, MS
Senior Program Officer
Research-Evaluation-Learning Unit
Robert Wood Johnson Foundation

U.S. Department of Commerce
Economics and Statistics Administrati
U.S. Census Bureau
1201 East 10th Street
Jeffersonville IN   47190-4444

OFFICIAL BUSINESS
Penalty for Private Use $300

D-6A(IN) (5-2009)

AN EQUAL OPPORTUNITY EMPLOYER

U.S. Census Form Enclosed
YOUR RESPONSE
IS REQUIRED BY LAW

2010 Census forms were mailed to reach everyone living in the U.S. *(Quinn Dombrowski)*

Executive Summary

Racial and ethnic health disparities and inequities can only be eliminated if high-quality information is available by which to track immediate problems and the underlying social determinants of health. Such information can guide the design and application of culturally specific approaches to medicine and public health. Often, health outcomes are disaggregated only by broad racial and ethnic categories such as White, Black, or Hispanic. However, the great, and growing, diversity of the American population means that people's actual experiences are much more specific.

The U.S. has numerous compelling reasons to build and maintain a much more robust practice of disaggregating data below the level of major racial groups and to link these data to the factors that influence health. Improving how we create, understand, and handle disaggregated data about race and ethnicity is central to the pursuit of health equity and a deeper appreciation of American society overall, as noted in the examples below.

- If public health practitioners and policymakers cannot track differences in chronic conditions or levels of treatment at a level more detailed than broad racial groups, they will lack the knowledge to act on the conditions that are leading to disparities. For example, obesity rates for U.S.-born African Americans are twice as high as for Black immigrants from Africa,[1] and Korean Americans have rates of being uninsured that are nearly twice as high as for Asian Americans overall.[2]

- If government-generated data collection about American Indians remains inconsistent in policies and practices, even on the basic question of who is a "Native American," the chances for improving population-level health will stay low for both urban Indians and those living on reservations.

- If the economic and social circumstances of refugees and immigrants are not understood because we do not have information about their specific cultural communities, we are less likely to design and target the best public health strategies, programs, or policies for these communities.

This recognition of the value and potential of disaggregated data motivates this report, which reflects two years of exploration and interaction among advocates, policymakers, researchers, government agents, philanthropists, and others with relevant experience. This project represents the first time such a wide array of leaders and interests in the U.S. have shared their work and compared their ideas on these issues. The participants delved deeply into how these issues are experienced by populations in each major racial category in the U.S. population, for each group has a distinct history and set of contemporary challenges. From that diversity, and from the dialogue among the participants in our project, have emerged common themes and a common agenda for change in health surveillance surveys, other data collection and research, medical and public health practice, government policy, and privately managed information systems.

This project began in 2015, in a period of expansion of numerous strategies for equity in health access and in the creation of healthier communities, many of them being implemented or supported by the federal government. It was also a period of innovation and exploration for the U.S. Bureau of the Census and other federal agencies, and many of their state counterparts, with respect to improving ways to measure race and ethnicity. The past year has seen a very different environment for nearly every dimension of this exploration. The future of many federal health equity initiatives is in question, the 2020 census has been underfunded, a climate more unfriendly to immigrants has created fear in the security of data collection, and broader racial animus has been well documented and more openly displayed. Although the environment has changed, the underlying trends remain and the need for a well-conceived map of actionable steps for achieving health equity is as great as ever.

Across all the dimensions of health runs the need to more accurately, precisely, and insightfully understand conditions and to measure progress, and that is where better information about racial and ethnic subgroups will be especially critical. Neither diagnoses of problems nor plans for change can be fully successful without "meeting people where they are"—understanding and documenting basic facts about their medical conditions and economic circumstances as well as the nuances of their culture and life experiences. That, ultimately, is the potential power that disaggregated data bring to bear to advance health equity.

This report analyzes the key focus areas in data disaggregation to advance a culture of health and recommends changes and improvements to the conduct of research and data collection and to the government and corporate policies that define priorities and allocate resources.

The Arab Complete Count Committee encouraged Arab Americans to not check "White" and to write in their detailed response under "Some other race" to improve data on their community in the 2010 census. (Arab Complete Count Committee)
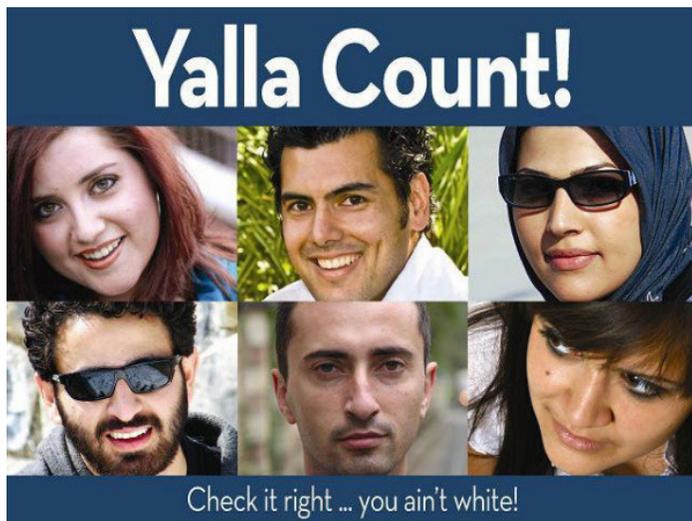
Key Focus Areas and Recommendations

Opportunities abound to improve how disaggregated data about race and ethnicity are created, understood, and managed. Our findings and recommendations are grouped into two broad areas:

- methods for collecting and analyzing data about race and ethnicity at more detailed levels, and
- government policies that can enable and enhance data disaggregation.

The report provides background on various areas of focus and why each is important, and how leaders in the field view the challenges and opportunities. Examples of recent innovations, campaigns, and resources are profiled throughout the key focus areas.

By improving data disaggregation, researchers, policymakers, and practitioners will be better able to influence the health environments, behaviors, and outcomes of communities across the U.S. The recommendations address conceptual, technical, and practical challenges to leveraging data disaggregation. Success will require significant commitments and investments from many institutions and leaders. The scientific case for making these improvements may be powerful, but as with most equity-driven improvements to government systems or the practice of research, such changes will not come into being purely through the force of logic. Efforts will require sustained, well-crafted advocacy, not only on the part of experts but also by groups that represent the interests of the communities in question.

For each of the two broad areas of change, several guiding principles capture the purposes, values, and priorities in the conduct of activity, followed by a set of more specific recommendations. These guiding principles and recommendations are provided in abbreviated form in this executive summary and with more detail—including the key actors for each recommendation—in the full report.

Methods for Collecting and Analyzing Data about Race and Identity

The research community should delineate the changes in approach and increases in resources that will be necessary to generate reliable information about a wide range of racial and ethnic groups, and to reliably connect this information to important health variables. The application of disaggregated data would be greatly improved through a series of investments, commitments, and innovations that would broaden, deepen, and improve the practice of research.

Guiding Principles

- **Methods of measurement of race and ethnicity should be selected to best serve the purpose of a given research or clinical activity.** Racial and ethnic identity is multidimensional, so researchers can and should deploy different ways of naming and measuring it. Most health surveys rely on self-reported race and ethnic identity, while health-care providers, including hospitals, often use third-party reports of race and ethnicity.

- **Health surveillance surveys should be supported well enough to be able to explore multiple dimensions of the intersection of race and ethnicity with other aspects of identity through oversampling, linguistic diversity, new questions, and other resources and techniques.** State and local health surveys, as well as the large federally supported endeavors, are in the best position to advance the collection of high-quality disaggregated data if they can get the resources and the political support to venture into these new areas.

- **Community voices should be integrated into the research design process, providing ideas, consultation, and feedback on the ultimate design, purpose, and intent of the research.** This authentic inclusion of community representatives will help ensure that the assets, strengths, and resilience of racial and ethnic identities are analyzed, as well as disparities and deficits. It will also improve the quality and representativeness of information.

- **A new source for high-quality, widely available survey data that represents key dimensions of the intersection of race, ethnicity, and health should be established.** Existing surveys should be strengthened, but ideally, a new, nationally representative, longitudinal-panel data set could function as an essential clearinghouse for the purposes of interdisciplinary population health disparities research. The data would serve to build the theoretical and practical foundation of knowledge around racial and ethnic subgroups and health.

Recommendations

1. **The research community should invest in research methods that address small "N" populations and small sample sizes.** Exploring, advocating, and refining research methods specifically targeted to these communities will generate more accurate measures of communities that are often underrepresented in the research and historically undercounted. These methods include field research, in-language surveys, ethno-racial sampling frames, accultured data tools, and pooled data sets.

2. **The research community should capture racial and ethnic self-identity with more standardized measurements, providing more nuanced and contemporary subgroup categories for respondents to choose from.** These measurements will better uncover factors throughout the life course that influence identity, such as the demographic environment surrounding multiracial individuals; the racial and ethnic context of an immigrant's home country; and the evolving norms, culture, and institutions of society in the U.S. The funders of social research should set a high bar and provide meaningful incentives for the improvement of these techniques.

3. **Survey managers and data users should determine the relative costs and benefits of sample size expansion and adding and modifying survey questions.** The effort to capture high-quality data about communities can be enhanced by careful evaluation of possible improvements to questions, sampling, outreach, and other steps in the survey process. This needs to be done in an environment of tightening budgets and during a trend of declining response rates, so it will be challenging and will often require trade-offs among the possible innovations. For example, what is the higher priority for achieving better disaggregated data: more detailed questions or broader, deeper outreach?

On July 29, 2016, over 400 people lined up around the block at 7 A.M. for the grand opening of the Filipino fast-food franchise, Jollibee in Chicago. (*Darold Higa*)

4. **The health research community should collect clinical data on biochemical factors associated with sociocultural and environmental stress and, where possible, align it with survey data on socioeconomic factors, race, and ethnicity.** Data on biomarkers, such as allostatic load or other biochemical processes, can reveal associations among heterogeneous racial and ethnic subgroups for health disparities research with an emphasis on mental health.

5. **Data providers and survey managers should provide more opportunities for researchers and data users to link health surveillance and administrative data sets to allow for more robust and accurate analyses.** Data from health surveillance surveys can benefit from linkages with data on education, employment, and other social determinants of health. When data are linked, more information is available to describe health behaviors and outcomes. It also reduces the need for each survey to collect the full scope of demographic information, reducing respondent burden and preserving valuable survey "real estate." Linking data sets creates opportunities for larger, integrated, and interdisciplinary analyses, which apply to government agencies, research institutions, and the public sector. In the process of linking data, data providers and survey managers will need to strengthen existing protocols for privacy and security of respondents' identifiable information.

6. **Data providers and survey managers should connect users to the abundance of existing public and private data applicable to population health disparities research.** There are many challenges to accessing health-related data, due to eligibility requirements, patient privacy rules, credentials, clearance, and bureaucratic processes, and often these get in the way of making high-quality, updated data on race and ethnicity accessible for important research and policymaking purposes. In some cases, the need is to streamline access to multiple survey data sets, and in other cases, the challenge is around using the information to improve community outreach to targeted groups that may potentially benefit from the data. The process of making more data available without compromising key privacy concerns will be complex, but is well worth the effort. Private-sector leaders in health-care provision and insurance, such as Kaiser Permanente and Anthem Health, can be an asset in this effort.

7. **The funders and managers of health surveillance surveys should add additional variables to surveys that more effectively reveal within-group differences.** These surveys about health and social determinants of health can reveal associations with health outcomes among racial and ethnic groups that may have more significance than associations between groups, such as skin tone, internal migration, or parents' country of birth.

8. **Advocates and the research community should thoroughly communicate the potential benefits, costs, and risks of increased data disaggregation among communities.** Disaggregation needs local champions and well-connected leaders to make the case for smart innovations. Particularly for racial and ethnic subgroups who are linguistically, geographically, and/or socio-economically isolated, the use of culturally relevant outreach will foster buy-in, build understanding, and recognize risks of greatest relevance to their communities. This can be done in partnership with culturally sensitive and appropriate communications and media channels, and the graduate and post-graduate fellowship and support programs that aim to increase diversity and attention to equity in health research.

Government Policies that Enable and Enhance Data Disaggregation

Public policy change is at the center of any systematic effort to expand data disaggregation for health-related research and practice, and those policies should include not only more resources directed to smart innovations, but also a strong commitment to racial equity and a political environment that celebrates the authentic racial and ethnic identities that comprise society.

Guiding Principles

- **Agencies supporting or conducting health surveillance surveys and other population surveys should make the collection and analysis of more finely disaggregated data about race and ethnicity a high priority.** Throughout the federal and state governments are numerous departments with the opportunity and resources to advance this commitment, and the experiences described in this report show that that commitment will pay off with a deeper understanding of health behaviors, outcomes, and social determinants.

- **Health-care and social-service institutions should receive adequate funding and technical assistance to build their data capacity.** These investments need to be made in patient intake systems, administrative and professional staff training, and IT systems to allow full compliance with data disaggregation requirements and updates.

- **Individually identifiable information about survey respondents, especially of underrepresented and marginalized groups, must be protected to ensure data security and respondent privacy.** These issues are especially important for communities threatened by civil rights and oversight abuses, such as the LGBTQ community and undocumented immigrants. Many existing policy measures require full compliance and enforcement, as well as continued strengthening. These measures for data security and respondent privacy defend against growing threats in cybersecurity and data misuse. It is also important to communicate the efforts and gravity of these measures to build public confidence and trust in survey management to ensure high response rates.

- **The relationships between the U.S. government and the American Indian/Alaska Native tribes should protect and build the integrity of research activity and data collection about tribal areas and tribal citizens.** The Census Bureau actively consults with tribal leaders and representatives, and research ethics principles protect tribal areas and citizens. These practices should continue to grow and value the tribal sovereignty of data and information about tribal communities.

- **The next generation of health researchers and practitioners should be actively recruited and well-trained with diversity, equity, and inclusion as values alongside the field of data disaggregation.** Programs with targeted funding and training will help transform the field of health equity through greater awareness and representation of racial and ethnic subgroups, among front-line staff, researchers, academics, practitioners, and advocates.

- **The consequences for policy and funding of the different ways of reporting race and ethnicity should be clearly defined and broadly understood.** The complexity of categorizing individuals' race and ethnicity, including single-race data, multiracial data, and tribal identification, affects how accurately population sizes are measured. For example, the Census Bureau reports "single-race alone" and "single-race alone or in combination with other races and ethnicities" to offer varying degrees of detail. Using "single-race alone or in combination with other races and ethnicities" would lead to larger population sizes that account for multiracial respondents. The implications of these choices are increasing in significance as the multiracial population grows, yet they are largely not understood by the public or many decision makers.

- **Investments from survey funders, philanthropies, and government agencies should be used to support data disaggregation efforts across sectors.** From recruiting the next generation of researchers and practitioners to evaluating the use of newly generated data, substantial and sustainable investments will optimize the data's utility to the field of population health disparities.

Recommendations

1. **Congress should fund the U.S. Census Bureau's budget to adequately maintain and improve operations for the 2020 census and all other surveys.** The success of the decennial census requires adequate funding to cover the extensive costs in the final years leading up to the decennial count, to ensure that the new technologies are deployed correctly, and to ensure outreach is sufficient. A well-funded census will ensure a full, fair, and accurate count of everyone living in the U.S. Funding for the American Community Survey and other surveys is very important as well. Congress should respond to the need to fully fund the census, as it has been expressed by the professional research community, leaders of many sectors within private enterprise, and a host of constituents concerned with civil rights and health equity.

2. **The U.S. Census Bureau and the Office of Management and Budget should improve the documentation of race and ethnicity in federal data collection.** The Federal Interagency Working Group's revisions to the Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity address question formatting and nonresponse, classification of people of Middle Eastern or North African race/ethnicity, additional minimum reporting categories, and relevance of terminology. All of these are important steps for data disaggregation for health equity. OMB should respond to the Working Group's recommendations and the Census Bureau and OMB should continue to consult with community representatives and leaders, researchers, policymakers, and advocates.

3. **The Office of Management and Budget and the U.S. Census Bureau should develop protocols for using data disaggregation consistently throughout the collection, analysis, and reporting of racial and ethnic subgroup data.** These protocols will encourage thorough execution of the federal data standards on race and ethnicity for greatest comparability and knowledge. These protocols can address challenges across sectors, such as how to enumerate free text responses, how to analyze "only Hispanic" and "Hispanic multiracial" subgroups apart from ethnic categories, and how to report a "roll up" of racial and ethnic subgroups when adequate data are unavailable to report more detailed information. The OMB and Census Bureau can generate these protocols with consultation of community members, data users, and researchers, and they can be applicable to data policy leaders at every agency.

4. **The research community, including Institutional Review Boards, should standardize enforcement of existing policies that facilitate research processes over activities on tribal lands and with tribal citizens.** There are policies pertaining to research ethics designed to respect the government-to-government relationship between American Indian and Alaska Native tribes and the U.S. government. Ensuring that these policies are enforced will protect tribal lands and tribal citizens from unethical research practices. Tribal elders, tribal leaders, and tribal epidemiology centers are all important decision makers and thought leaders to include in the conversation.

5. **Advocates, policy leaders, community members, and influencers should articulate strong arguments in support of data disaggregation as a tool for advancing health equity.** A solid foundation of evidence exists about the advances in data disaggregation, its positive benefits, and the innovations needed to further enhance this field. These arguments should convey the urgency for action and the costs of inaction. The benefits to society of enhancing health equity are economic as well as moral and political, and the case should be made on all fronts. Advocates in the realms of civil rights, criminal justice, disability, immigration, reproductive justice, youth organizing, and other areas can make the case and spread the word of how health equity can be enhanced with better data. The case should be made not only nationally but in the states, as shown by the recent examples of leadership in California, Minnesota, Rhode Island, and several other states.

The dialogues and explorations undertaken for this project show clearly that the nation is eager to move ahead to improve how disaggregated data about race and ethnicity is created, managed, interpreted, and applied. Enhancing this field of research and practice will be a major asset not only for advancing a culture of health, but for achieving greater mutual understanding of race and ethnicity in our diverse society.

The importance of achieving a higher level of data disaggregation is clear to people who have worked with—and come from—the communities that have been overlooked, and to researchers and policymakers who have sought the information necessary for addressing health inequities. It is a call for compelling information, evidence, and stories of many kinds. But, this is still a relatively new issue for most people. Making changes in systems, policy, and practice will require broader and deeper understanding among many new constituencies and groups of experts. It will require building the public will to bring new priorities to the fore in government, and for intelligent, progressive uses of new data sources managed by the private sector.

A Boston Cape Verdean delegation taking part in a 2016 multicultural Marian procession that concludes at the Basilica of Our Lady of Perpetual Help. (*Roman Catholic Archdiocese of Boston* / *George Martell is licensed under* CC 2.0*)*

Introduction

Does a fifth generation Chinese American have much in common with a Hmong refugee from Laos when it comes to the factors that determine their health and well-being? What are the reasons for wide variations in health outcomes between immigrants and U.S.-born people of the same race? What steps does a clinic in Los Angeles managed by Muslim medical students take to reassure its Latinx* clients that their records are safeguarded against misuse by immigration authorities? Why are there so many different—and inconsistent—ways of determining who is Native American, and what are the consequences of this inconsistency for Indian health? What are the implications on health policy of the rapidly increasing number of people in the U.S. who identify as being of more than one race?

"Disaggregation of data about health into racial and ethnic subgroups" may sound at first to be a relatively dry or technical subject, but it provides a window opening onto a broad range of important questions about the meaning of culture and identity in American society. These issues have great consequences for the prospects of advancing a culture of health, and they need closer attention. They were always important factors in the U.S., but the increasing demographic diversity of the country has brought them into sharper focus and has made them more frequent topics of debate.

Racial and ethnic health disparities and inequities can only be eliminated if high-quality information is available through which we can track immediate problems and underlying social determinants. This information is needed to guide the design and application of culturally specific approaches to medicine and public health. Often, health outcomes are disaggregated only by broad racial and ethnic categories, such as Black,

Hispanic, Asian, White, or Native American. However, the diversity of the American population means that people's actual experiences are much more specific.

Admittedly, the U.S. disaggregates information by race more than it once did, with the era of denoting people as simply White, Black, or Other being largely gone. But that gradual progress in recognizing and documenting the country's diversity has not meant that we are well-prepared for the present or the future. If, for example, we miss out on understanding the economic and social circumstances of the Hmong refugee because information is only available about an extremely diverse category called Asian Americans, we are less likely to design and target the right public health strategies, programs, or policies. If we do not appreciate the meaning of multiracial identities, we will not understand either the lives of this growing part of the population or the contemporary context of race and culture. If undocumented immigrants stay away from needed services because they are afraid that the medical record systems present a risk to their remaining in the country, then their health and that of the community will suffer. If federally run data collection about the status of American Indians and Alaska Natives remains inconsistent and often at odds with tribal sovereignty, the chances for improving population-level health will stay very low. In these and many other ways, improving how we create, understand, and handle data disaggregated by race and ethnicity is central to the pursuit of health equity.

That recognition has motivated and informed this report, which reflects two years of research and interaction among advocates, policymakers, researchers, administrators of health and statistical agencies in federal and state governments, leaders in philanthropy, and others with relevant experience. The exploration has allowed us to delve deeply into how these issues are experienced by the populations in each major racial category in the U.S., and to better understand each group's distinct history and set of contemporary challenges. But, from that diversity and the dialogue among the participants in our project, common themes and a common agenda have emerged

---

* We use the term "Latinx" throughout this report as a gender-neutral alternative to "Latino." In Spanish, the masculine version of a noun is gender neutral, but some advocates, mostly based in the United States, have called into question the neutrality of masculine nouns. Language evolves and more inclusive alternatives to "Latino" have included Latina/o and Latin@. We use Latinx to intentionally include gender nonconforming and nonbinary people of Hispanic or Latin American origin who might not identify with the masculine or the feminine. Though a relatively new term believed to have originated in the Latino LGBTQ community, and not without dissent, Latinx is being increasingly adopted by advocates, universities, and even government agencies. Stanford, for example, has a Chicanx/Latinx Studies program and the CDC names October 15 as "National Latinx AIDS Awareness Day."

around changes needed in health surveillance surveys, other data collection and research efforts, public health practices, and in the policies for the government and the purveyors of increasingly influential privately managed information systems.

This project represents the first time such a wide array of leaders from a variety of interests in the U.S. have shared their work and compared their ideas on these issues. It was important to build the deliberations upon a solid foundation and, toward that end, the Robert Wood Johnson Foundation (RWJF) commissioned research reviews of the state of data disaggregation for health,[3] one for each of five major population groups in the U.S. Also commissioned was a seven-country comparative study which showed the diverse ways in which those nations are addressing the implications for health of the heterogeneity and granularity of their populations.

Amid the wealth of distinct histories of, and contemporary issues pertaining to, each part of the American population, some important common themes emerged. Across the board, disaggregated data were limited for each racial group, and the constraints of the existing data were a serious barrier to useful basic research and policy analysis. There was a lack of clarity about the process of "rolling up," or combining, racial subgroups in analysis and reporting. Samples that were too small to be reported were a common problem, and research across all subgroups was too sparse and too limited. America's complex, conflicting, and always evolving values about race, ethnicity, and culture are reflected in the inconsistent official systems by which people are—or are not—categorized, and the ways in which such information is—or is not—deployed.

Although these shortcomings are significant, each study also found promising practices upon which to build a better system—ideas that are discussed throughout this report. The research collectively reaffirmed the intentions of the project, to explore common themes among all racial and ethnic subgroups that could support data disaggregation as an effective tool for health equity.

The issues the project participants raised, and the topics for which disaggregation is important, touched upon all the elements of the Culture of Health Action Framework[4] put forward by RWJF, from building health as a shared value to restructuring health-care delivery systems and creating healthier communities. Across all these goals runs the need to more accurately, precisely, and insightfully understand history

This young Iu Mien girl is participating in SEARAC's photo campaign to support AB 1726, with over 300 individuals identifying as a category not currently represented in the data. *(Southeast Asia Resource Action Center (SEARAC))*

and current conditions and to measure progress, and that is where better information about racial and ethnic subgroups will be especially critical. From basic facts about medical conditions and economic circumstances to the nuances of culture and life experience, neither the diagnoses nor the plans for change can be fully successful without "meeting people where they are"—understanding and documenting them in terms that genuinely pertain to them. That, ultimately, is the potential power of disaggregated data brought to bear to advance health equity.

This project began in 2015, in a period of expansion of numerous strategies for equity in health access and in the creation of healthier communities, many of them being implemented or supported by the federal government. It was also a period of innovation and exploration for the U.S. Census Bureau, other federal agencies, and many of their state counterparts with respect to improving ways to measure race and ethnicity. However, 2017 has seen a very different environment for nearly every dimension of our exploration. The future of many federal health equity initiatives is in question, the 2020 census is underfunded compared to recent censuses,

a climate that is more unfriendly to immigrants has created fear and lack of confidence in the security of data collection, and broader racial animus has been well-documented and more openly displayed.

The environment has changed, but the underlying trends remain and the need for a well-conceived map of actionable steps is as great as ever. The scholars, advocates, and policymakers exemplified by the participants in this project are ready, even impatient, to move forward with practical actions that can improve the quality, availability, and utilization of disaggregated data. Those actions range from improving survey questions about race to increasing sample sizes to better capture smaller population groups. Such steps require political will and resources as well as technical proficiency, of course. But the proposals raised through this exploration also reflect the importance of committing to a more participatory vision of research when working in and with marginalized communities, and to a more nuanced understanding of race, culture, health, and their intersectionality than basic statistics alone can provide.

The guiding principles and recommendations in the report recognize the crucial role of "making the case"—of broadening and deepening public and professional awareness of these opportunities and challenges and of building support for positive change. Taken together, the recommendations provide a starting point for what we hope will be a persistent, creative, and growing movement to understand and take full advantage of the nation's diversity, and thereby, advance health equity.

Project Activities and Research Methods

This project was a multifaceted investigation of the leading issues and opportunities for disaggregating data by race and ethnicity for use in furthering health equity. We employed a combination of methods and sources to take advantage of the knowledge and experience of a wide range of experts.

One project goal was to bring together specialists who may not have connected previously. With that in mind, the network of project participants was a multidisciplinary set of academic researchers, policy analysts, community-based organizers, health practitioners, data scientists, managers of health surveillance surveys, government officials, funders, and other data users. The wide range of professional roles and experiences was complemented by the diverse racial and ethnic backgrounds of the participants, all of which led to a well-informed dialogue about the most salient aspects of identity that influence health.

As the starting point for the project, six research reviews were commissioned by RWJF and conducted by leading health equity researchers (see Appendix A). Five of these were based in the U.S. and were concerned with major racial categories: non-Hispanic White; Black and African American; Asian American, Native Hawaiian, and Pacific Islander; Latinx; and American Indian/Alaska Native. The reviews of literature were supplemented by key informant interviews, legislative timelines, and theoretical frameworks. More than 460 databases or published sources were cited in the U.S. reviews. All the reviews concluded with recommendations for the field to improve the quality of disaggregated data and its use for achieving health equity.

The issues presented in this report have a U.S. orientation, although in their most general form they are applicable across the globe as well. Many of the key focus areas are translatable across countries and contexts. The conceptualization and realization of race and ethnicity varies and it is helpful to understand and compare those variations outside of the U.S. context. The sixth review examined seven nations, each with a distinct approach to population definitions and policies:, Aotearoa New Zealand, Canada, Denmark, Great Britain, Hungary, Malaysia, and the Plurinational State of Bolivia, as well as the European Union.[5] The ways in which these countries relate to indigenous populations, define and document ethnicity, handle

information about immigration and citizenship, and conduct research on health disparities are illuminating and remind us that many things we may assume to be natural or inevitable are actually highly specific to U.S. culture and government.

A series of three convenings followed the research reviews to provide space for agenda setting, issue generation and refinement, and creative thought about data disaggregation for all racial and ethnic subgroups. Attendees were invited to the sessions held in Los Angeles, Atlanta, and Washington, DC, during the spring of 2017 (see Appendix B). Round tables and interactive workshops served as a meeting ground for cross-pollination, mutual education, and dialogue about the issues most salient to advancing data disaggregation.

- The first convening focused on **making the case for data disaggregation** to a variety of audiences, for example, within major racial groups and subgroups, research disciplines, data managers and users, policymakers, funders, and others.

- The second convening focused on **multiple definitions of race and the multiracial community,** advancing our understanding of how race and ethnicity play a role in health and other interacting factors.

- The third convening focused on **immigrants and migrants** and how their pathways to the U.S. and varied circumstances need to be better analyzed and understood in relation to their impact on health.

The convenings brought together close to 100 individuals. Participants addressed policy issues from multiple perspectives, listened to advocates and on-the-ground organizers about their campaigns, explored the boundaries of theoretical and practical health equity issues, and learned about how race and ethnicity resonated differently across communities. In all, the participants proposed 400 suggestions for action.

A webinar held in October 2017 was open to the full project network, which included those who had attended a convening as well as other influential leaders who had been invited or referred. During the webinar, our project team proposed a set of recommendations curated from the suggestions in the convenings. More than a dozen participants then reviewed the first draft of this report.

While the project activities were underway in 2016 and 2017, a working group of staff from numerous agencies, commissioned by the Office of Management and Budget (OMB), was revising the 1997 Standards for Federal Data Collection on Race and Ethnicity. This process was separate from our project, but many of the leaders on these topics from the Census Bureau, U.S. Department of Health and Human Services (HHS), OMB and other participating agencies shared their plans and joined in our deliberations, as did members of their advisory committees. The OMB's Federal Interagency Working Group proposals to revise data collection standards directly influence guidelines for all data collected or funded by the federal government. As such, they also indirectly influence the standards for state and local data collection.

With this set of activities, we sought to be comprehensive, interdisciplinary, and attentive to the needs of the field in order to build a common agenda for data disaggregation. This work provides a solid foundation of research and relationships upon which that agenda can be advanced. This project network is an important ongoing component of the project, providing expert opinion and community perspectives, making connections across issues, and bringing to bear members' understanding of the nuances of data disaggregation.

❝ ❞

It is rare to be in fellowship with other data geeks and not have to prove that data disaggregation is important. [PolicyLink] has tapped over 100 organizations, all of whom are a part of this fellowship. We've never been able to participate with such a wide array of organizations across the country. We hope that this kind of network can continue, across all racial and ethnic categories, to help us in our work.

Kathy Ko Chin,
President and Chief Executive Officer of Asian & Pacific Islander American Health Forum

The USDA Interfaith Iftar Celebration recognizes and celebrates the "breaking of the fast" for Muslims in Washington, DC, and around the world, in June 2017. *(U.S. Department of Agriculture / Lance Cheung)*

Data Disaggregation and a Culture
of Health: Key Focus Areas

Data presented at the level of broad racial and ethnic categories tend to mask differences within those categories and render invisible distinct subpopulations who are doing better than or worse than the average. Health disparities have been disproportionately faced by certain racial and ethnic subgroups for a long time, but it is only recently that a critical amount of attention has been paid to these differences.[6] As demographic diversity in the U.S. grows at an accelerating rate, the value and necessity of tracking and analyzing these patterns will only increase.[7]

Information on racial and ethnic identity can illuminate how communities' risks, behaviors, assets, and challenges relate to health outcomes. By analyzing data at this level, health disparities that would otherwise be undetected, exacerbated, or ignored can be identified and addressed. Understanding specific cultures allows for the design of strategies to increase resilience, protective factors, and the drivers of positive change in health outcomes for those subgroups.

Race and ethnicity are not the only dimensions of identity that matter to health outcomes. Other variables, such as gender expression, disability, country of origin, ancestry, or skin tone, each have consequences for health. Well-executed data collection and research about all these factors can be a powerful tool to help achieve health equity.

The issues presented here arose throughout the project as pressing concerns, common themes, shared targets, and goals for public health research and practice. They are necessarily snapshots of complex topics, but, taken as a set, they represent a strong picture of the insights and commitments of our diverse participants to moving the field forward.

Key focus areas are organized into two main sections. First, the section on **methods** examines data collection and analysis to see where a strong commitment to greater disaggregation with respect to race and ethnicity would take the field. It describes the challenges to obtaining statistically representative samples of small communities, and the ways in which innovative researchers are overcoming those odds. It explores the intersection of race and ethnicity with other aspects of identity in the study of population health disparities. The subsections convey why some populations have been largely invisible with respect to health-related data, and what the opportunities are for increasing their visibility.

The section on **policies and practices** details the federal, state, and local activities that can provide the foundation for better publicly sponsored data collection. The upcoming 2020 census and the proposed new federal standards for measuring race and ethnicity are the most prominent, current, politically charged arenas, and the debates about the standards and the census have implications for a dozen federal departments and scores of surveys and other data sources. At the state and local levels, a number of initiatives in the health and education sectors have enacted disaggregation beyond the federal minimum requirements.

Following each of these two narratives on key focus areas are guiding principles and recommendations to define the conduct of work in these areas. The guiding principles are broadly applicable to the field and convey priorities for how all groups are counted and seen in the data. The recommendations are more specific and directly actionable by a range of organizations.

Please report babies as age 0 whe

Print nu

Age on April 1, 2010   Month

→ **NOTE: Please answer BOTH Quest**
Question 9 about race. For this ce

8. **Is Person 1 of Hispanic, Latino,**

☐ **No**, not of Hispanic, Latino, or S

☐ Yes, Mexican, Mexican Am., Ch

☐ Yes, Puerto Rican

☐ Yes, Cuban

☐ Yes, another Hispanic, Latino, o
Argentinean, Colombian, Dominican, Nicaragu

On the 2010 census, Question 8 asks, "Is Person 1 of
Hispanic, Latino, or Spanish origin?" A person of Hispanic,
Latino, or Spanish origin can be of any race.
(Quinn Dombrowski)

19

9   n 1's race? Mark X

Methods for Collecting and Analyzing Data about Race and Ethnicity

Our project examined many dimensions of personal and group identity; how these are related to health; and how they intersect with, complement, or confound the understanding of race and ethnicity. New approaches to research will be needed to raise the visibility of small and hard-to-reach populations, and to document and address their health issues.

This section covers a range of methodological challenges and opportunities raised by the commitment to disaggregate data for distinct groups within the major racial and ethnic categories. There are several reasons to take up these opportunities.

- The health of the diverse population of the U.S. cannot be assessed without understanding migration in all its aspects, from formal designations of newcomers, such as their legal status, to social factors in the immigrant or refugee experience, such as acculturation.

- As the proportion of the population identifying as multiracial grows, the ways to measure and understand race need to keep up with these changes in both numbers and social norms.

- Health inequities are also a function of factors beyond race and ethnicity, such as skin tone or gender identity, and those factors need to be better measured and understood. They are important in their own right and as components of an intersectional understanding of the determinants of health.

- The ethical conduct of research in and with marginalized communities is as essential as any methodological innovations. Issues of health and race were traditionally examined in ways that were often highly inappropriate, and the current environment calls for data collection that is guided by principles of respect for community and the authentic engagement of residents in the research.

Visibility of Small and Hard-to-Reach Communities

One of the most compelling reasons to pursue disaggregated data is to understand the circumstances of small ethnic groups and other populations whose numbers are modest. Among small populations, low response rates may limit the representativeness of results. Data collection on small and hard-to-reach populations is a chronic challenge for surveys because of logistical, financial, and outreach constraints. Hard-to-reach populations include proportionally more people without cell phones or physical addresses as well as linguistically isolated people and migrants. As a result, these communities are often undercounted by conventional sampling methods; these shortcomings, in some cases, exacerbate the vulnerabilities and disconnectedness of these populations.

Researchers and survey managers have developed strategies to better capture data on these populations, including providing surveys in-language and drawing on nontraditional sampling techniques. One of the most commonly used nonprobability methods is snowball sampling, which includes using a respondent's networks to recruit additional subjects who fit the study inclusion criteria. To use this approach, researchers generate "referral chains" from an initial sample of "seeds." [8] Though findings derived from this method may not always be generalizable to the broader group, it is especially fruitful when working with stigmatized or hard-to-reach groups. As inequities within subpopulations become more apparent and pronounced, researchers are developing innovative methods to ensure that these communities are represented in health surveys.

We have a little bit of everybody, but not enough of anybody to really focus in on.

A.J. Scheitler, National Coordinator of the National Network of State and Local Health Surveys

When data are collected on subgroups, a "roll up" of these small populations into larger population groups is inevitable in certain circumstances to maintain statistical reliability or to protect the identities of members of small subgroups. For instance, the American Indian/Alaska Native community is often noted as the "Asterisk Nation," a nation too small and "difficult" to enumerate in data reporting, and too costly to be sampled appropriately. Instead, in reporting, an asterisk (*) is used to note that the data are there, but are insufficient to report because of small sample sizes.[9] Oftentimes, American Indians and Alaska Natives are lumped together with "Other" or left off reports entirely. This illustrates chronic challenges of being a small and hard-to-reach population, and the need for data disaggregation and advanced research methods to shed light on this community. In many cases, this is preferable to having no data at all.

A group of Hmong Americans pose during the 10th annual Hmong New Year Celebration in downtown Chico, California. (*Carol M. Highsmith*)

Multiple Definitions of "Indian"

In addition to being a relatively small population, American Indian and Alaska Native (AI/AN) populations must also contend with the multiple federal definitions of "Indian" currently in use.[10] For example, the Census Bureau and the U.S. Department of Education defer to the OMB, which defines an American Indian or Alaska Native as "a person having origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment." On the other hand, the U.S. Department of Housing and Urban Development only counts people who are a member of a federally recognized tribe. In addition, the Census Bureau publishes population numbers by race "alone," "in combination" with other races and ethnicities, and "alone and in combination" with other races and ethnicities. These numbers produce radically different estimates for the size of the American Indian and Alaska Native population. To complicate it further, the Department of Education decided in 2007 to only report AI/AN-specific data for non-Hispanic or non-Latino students and to classify as Hispanic or Latino those AI/AN students who selected both Hispanic/Latino and AI/AN.[11] These cross-agency differences impact the standardization and thus utility of disaggregated data on AI/AN peoples, limiting the ability of tribes to effectively plan for and govern their nations.

> **❝❞**
>
> **The AI/AN population is missing in a lot of research and resources on migration, but migration is very important for AI/AN health and outcomes. A lot of times this population seeks service outside of the reservation because of the lack of services on the reservation or maybe the disease they have has a stigma on the reservation.**
>
> Adrian Dominguez, Scientific Director at Urban Indian Health Institute Measurement Tools for the Multiracial Population

Because of the status of American Indian and Alaska Native people as a racial group and as a group with unique political status, AI/AN researchers, including those at the National Congress for American Indians, warn of data gathered based purely on self-identification.[12] The value of self-identified AI/AN race (as collected by the U.S. Census and many other health surveillance surveys) is contested because of the sovereign authority of tribes to confer membership. Whether through blood quantum or other means, each tribe, including tribes that are not federally recognized, has its own requirements for membership.

Many American Indian and Alaska Native peoples have left reservations and other areas, by choice or by federal relocation policy, which has led to a large population of urban American Indians and Alaska Native peoples. By some counts, nearly 67 percent of the nation's self-identified AI/AN people are urban Indians, complicating definitions that consider "tribal affiliation or community attachment." [13] Urban AI/ANs are approximately twice as likely as the general population to be poor, to be unemployed, and to not have a college degree. [14] There are also people who identify as pan-tribal, and people who migrate on and off tribal lands throughout their life. Still, each federal agency uses its own definition of AI/AN—from self-identification to formal tribal affiliation—in its funding formulas. This has significant implications for how federal dollars are allocated and distributed. Any solution to standardize across agencies must honor the sovereign authority of tribes to confer tribal membership.

Measurement Tools for the Multiracial Population

The multiracial population is one of the fastest growing segments of the U.S. population. According to the Pew Research Center, in 1970, just 1 percent of babies were multiracial. But by 2013, one in 10 babies were multiracial.[15] The 2000 census was the first decennial census that allowed respondents to select more than one race, making the 2000 and 2010 censuses the first nationally representative comparisons of self-identified multiracial populations.[16] Between 2000 to 2010, the multiracial population grew 32 percent from 6.8 million to 9 million people. It is understood that a portion of this growth is due to measurement improvement, or more and more people choosing to self-identify as multiracial, in addition to natural growth of the multiracial population. Some researchers, however, still ask multiracial people to select the racial/ethnic identity they identify with most strongly and recode them into single-race groups.[17] Other studies and surveys opt to bunch all multiracial people together as one group distinct from single-race people, regardless of their backgrounds.

As the multiracial population grows, the understanding of race as a social construct rather than a fixed biological category is also becoming more mainstream. There is no such thing as "pure" racial/ethnic groups, and most of the population is of multiracial ancestry.[18] Existing research on the multiracial population recognizes the "multidimensionality of multiracial status" and how the measurement of the population varies depending on the purpose of the research.[19]

A growing body of literature points to the specific consequences of different racial/ethnic combinations when it comes to health. For example, one study on multiracial health outcomes used data from the Behavioral Risk Factor Surveillance System (BRFSS), a national telephone survey about risk behaviors, chronic health conditions, and prevention, and found that self-rated health varies considerably across multiracial subgroups, with White-Black respondents reporting self-rated health most closely to Whites, White-Asians reporting better self-rated health than either group, and White-AI/ANs and Black-AI/ANs reporting poorer self-rated health than either group.[20] Better and more frequent measurement of multiracial identities should improve the prospects for understanding the reasons for these differences and what else matters for health for the growing multiracial population.
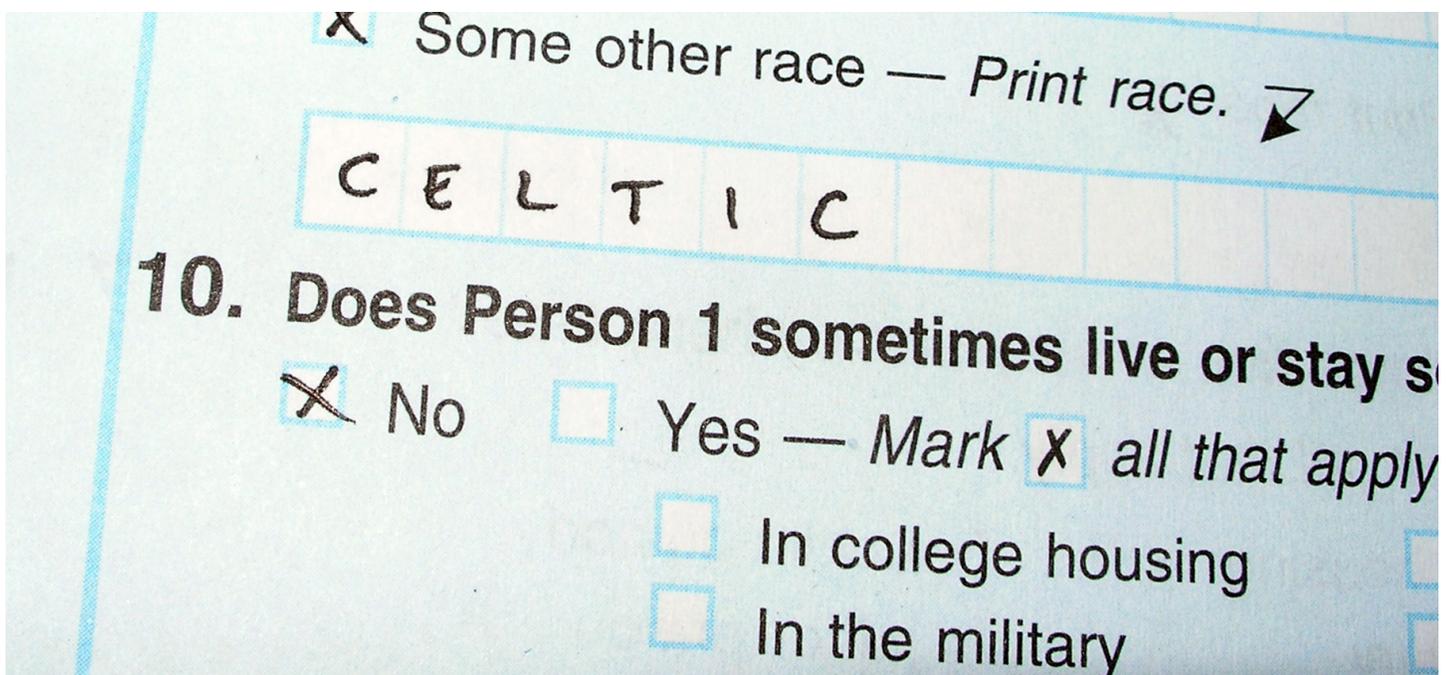
In the era of readily available DNA testing for ancestry, many people in the U.S. are reexamining their racial and ethnic backgrounds, leading perhaps to evolving attitudes about having "mixed" heritage, not to mention the creation of potentially huge information bases. But despite that trend and the presence of country-of-ancestry information in the American Community Survey, among other sources, there is very little health outcomes research in the U.S. that compares ethnic groups within the non-Hispanic White population, and White is generally used as a monolithic reference group for comparison with other races. This, of course, is different from many countries where ethnic differences within the White population are the basis for some of the most basic distinctions, if not divisions, in society.

Aspects of Racial Misclassification

Misclassification of individuals' background is unavoidable when some surveys rely on self-reported race and ethnicity data and others are based on observed race and ethnicity data from a third party, as is common in hospital admissions, for example. One reason for misclassification includes phenotypical diversity within multiracial and single-race groups. For the multiracial population, this often means classification into a single-race group, which may or may not be part of their ancestry. One way that researchers have tried to account for and predict these discrepancies is by asking about "reflected race" on self-reported surveys.[22] In addition to capturing self-reported race and ethnicity data, researchers also collect data on the race that a respondent believes other people would assign to them, which is known as reflected race. In some cases, this measure is an even better predictor of health outcomes for the multiracial population than self-reported race because multiracial adults may identify with a single-race group whose health experience they do not share.[23] These inaccuracies exist throughout the research process from data collection to analysis and reporting, challenging the understanding of the health of multiracial populations.

On the 2010 census, in response to "What is Person 1's race?," an individual can select "Some other race" and write in their detailed answer. (*Mitch Barrie*)

"There has been documentation in the literature that when nurses, doctors, funeral directors, child welfare workers, and school staff are required to report race for AI/AN people, they often misclassify race.[24, 25]

from the research review of American Indian and Alaska Native populations, by National Congress of American Indians

In addition, many surveys rely on self-reported data, but designations connected to health and human services are often based on how a third party interprets one's race or ethnic identity. This generates misclassification issues for the multiracial population, and misclassification is very common in the American Indian and Alaska Native population. The Native American Cancer Research Corporation identifies several reasons for misclassification: the use of Spanish surnames to determine race and ethnicity, the fact that AI/AN is not a ubiquitous response category in medical records (including hospitals and other health clinics), inconsistent definitions of AI/AN people, and changing self-identification over the life course.

Racial misclassification also results in significant underestimations of death rates and cancer incidences, as uncovered by research that linked Indian Health Service data with National Death Index records and the National Vital Statistics System.[26] The undercount of AI/AN people caused by misclassification makes effective planning of public health programs more difficult. Racial misclassification of AI/AN people was more likely in regions where they made up a smaller share of the population. An undercount of the AI/AN population in vital statistics, cancer, and other health data limits our understanding of AI/AN health and leads to underfunding of programs serving the AI/AN population.

Community Ownership and Participation in Health Research

The technical aspects of data disaggregation cannot be separated from the reality that this effort is centered on, and will take place among, populations of color who historically have often been treated poorly in the process of research. Therefore, the endeavor to improve this data collection and the use of the information should be guided and informed by progressive principles that are attuned to those populations. Community ownership of data, and community-based participatory research, can go a long way to improve the quality of information and to advance health equity goals in ways that are respectful of cultures, local contexts, and people.

Community ownership of data and community-based participatory research place value on the perspectives of those most affected by health disparities to own, create, and use data in ways that have greater accountability, transparency, and significance to those directly affected by health disparities. The implementation of data disaggregation on the ground can be realized in steps such as the linguistic and cultural competency of interviewers, the contributions of local leaders to the design of empirical fieldwork, and the local dissemination and discussion of findings.

Nothing about us, without us! It matters who is asking the questions and how the disaggregated data will be used.

Richard Chase, Senior Research Manager at Wilder Research, Amherst H. Wilder Foundation

These equitable practices involve all parties affected and view both community members and researchers as producers of knowledge, with the chance to obtain better data and more insights than would otherwise be the case.[27] This commitment to local ownership can extend to questioning the traditional interpretation of statistics when that approach perpetuates false stereotypes and narratives of communities that are overly focused on identifying disparities, rather than assets and strengths. Community-based participatory research is adaptable and can work well with small sample sizes, can align with active policymaking and planning, and can guide knowledge and asset building among communities.

Indigenous Data Sovereignty in Practice

Indigenous data sovereignty is the right of a nation to govern the collection, ownership, and application of their own data. This sovereignty originates from Native American tribes' right to govern their own land and people, in accordance with treaties, international covenants, and declarations where the U.S. is a participant, such as the United National Declaration on the Rights of Indigenous Peoples (UNDRIP). With the framework of data sovereignty, Native scholars are seeking to protect the fundamental rights and interests of indigenous people.[30]

Tribal governance strategies, like the tribes themselves, are diverse, but there may be some common principles that can guide tribes from data dependency toward data sovereignty. Community members and researchers have co-created approaches to gathering disaggregated data in a way that ensures that the community owns the data from collection to dissemination. Wilder Research conducts a study of homelessness and near-homelessness on six American Indian reservations in Minnesota every three years.[31] This study is testament to the practice of data sovereignty, where each tribe is a sole owner of their own data. The study is paid for by the Minnesota Department of Human Services and involves a partnership between Wilder Research and tribal housing representatives to address issues around data for policy and planning. Key recommendations concern not only housing availability, but also access to adequate transportation and employment opportunities. Wilder Research only reports aggregate results for these tribes to ensure that descriptive data can inform statewide policy while also respecting tribal privacy.

Another area of growth has been through the U.S. Indigenous Data Sovereignty Network (USIDSN), which helps ensure that data for and about indigenous nations and peoples in the United States (American Indians, Alaska Natives, and Native Hawaiians) are used to advance indigenous aspirations for collective and individual well-being. Hosted by the Native Nations Institute, a unit of the University of Arizona Udall Center for Studies in Public Policy, the USIDSN's primary function, according to their website, is "to provide research and policy advocacy to safeguard the rights and

promote the interests of indigenous nations and peoples in relation to data."[32] The USIDSN maintains that when data are collected from or about the people of an indigenous nation, it comes under the control of that nation, even if it is collected by federal, state, or local governments.

Don't come to Native communities because of high disparities, come because we have solutions.

Michele Suina (Cochiti Pueblo), Program Director for the CDC Good Health and Wellness in Indian Country Program at the Albuquerque Area Southwest Tribal Epidemiology Center

In 2016, Navajo Nation President Russell Begaye participates in a listening session hosted by the U.S. Department of the Interior. *(U.S. Department of the Interior)*

The ethics and legal environment for community-driven research are distinct for American Indian and Alaska Native people as members of sovereign tribal nations. There are 573 federally recognized tribes, each with their own customs and laws.[28] Tribal scholars emphasize the importance of indigenous data sovereignty, which asserts the authority of tribal governments to govern the collection, ownership, and application of data about their people in alignment with their values and needs. In a 1989 case with the Havasupai Tribe, researchers from Arizona State University misused blood samples originally intended for type II diabetes research. After completing their research, and being unsuccessful at finding a genetic tie between type II diabetes and the Havasupai, the researchers continued to use the blood samples for unrelated studies on schizophrenia, migration, and inbreeding—all taboo topics for the Havasupai tribe. Settled in *Havasupai Tribe of the Havasupai Reservation v. Arizona Board of Regents and Therese Ann Markow 2004,* this case highlights lack of informed consent, violation of civil rights, unapproved use of data, and violation of medical confidentiality. Tribal leaders are working to translate indigenous data sovereignty into actions in research practices, such as informed consent, and other safeguards to preventing similar problems.[29]

❝ ❞

Egyptian immigrants and Americans of Egyptian descent comprise a unique population whose cultural and religious beliefs impact on decision making and behaviors related to cancer diagnosis and treatment. This population is overwhelmingly Muslim, although a sizeable minority are members of Eastern Christian sects. Dietary restrictions, social conduct, and religious observance are among the areas that require understanding by health providers.[34]

Jen'nan Read, Duke University, in the research review on the non-Hispanic White population done for this project

Intersection of Other Factors with Race and Ethnicity

Efforts to understand racial health disparities need to be informed by many dimensions of identity and life circumstances. There are many other aspects of one's identity, such as being an immigrant, that intersect with race and ethnicity to have an influence on health. The field of population health research is investigating aspects such as international and internal migration, acculturation, citizenship and legal status, gender expression, and sexual identity to discern nuanced patterns in racial health disparities. The theoretical and methodological advances are growing and ultimately expanding the intersectional lens of health disparities.

Health of Immigrants

People who move to the U.S. as immigrants or refugees encounter influences on their health that are distinct from those faced by the native-born population, and the disaggregation of immigrant and refugee populations is essential to understanding those factors. There may be something protective and unique about being an immigrant with respect to health outcomes, such as socioeconomic status in their home countries or environments there that promoted physical activity, healthy diets, and, in some situations, possibly less overall inequality. The health status of immigrants is often better than that of U.S.-born people despite immigrants generally having lower socioeconomic status than their U.S.-born counterparts, and the health of immigrants tends to decline the longer they live in the U.S.[33] The so-called "immigrant health paradox" is clearly not a consequence of membership in broad categories such as Latinx or Asian, but rather of being in specific cultural groups from particular places—exactly the kind of detail that more disaggregated data are intended to pick up.

Often, the differences in health outcomes between immigrants and U.S.-born populations of the same racial and ethnic subgroup are very substantial.[35] The rates, risks, and protective factors that shape the mental health in the Latinx community, for example, vary by acculturation factors, such as nativity status and age of migration. Alegría and colleagues found that being an immigrant functioned as a protective factor, or strength, for most psychiatric disorders among Latinx people, but to varying degrees depending on the subgroup and disorder.[36] Other researchers found that Mexican American immigrants reported significantly lower rates of any depressive,

anxiety, and substance-use disorder than U.S.-born Mexican Americans, and lower rates than U.S.-born non-Latinx Whites.[37] Alegría and colleagues also found that Puerto Rican American and Cuban American immigrants reported significantly lower rates only for substance-use disorders.[38]

Poverty is highly correlated with health disparities, and it interacts with nationality and nativity in complex ways. In 2015, Latinx poverty was 25 percent compared with 11 percent among Whites. U.S.-born Latinx people have a slightly higher poverty rate than Latinx immigrants. Within the U.S.-born Latinx population, people of Central American ancestry had the highest poverty rate at 29 percent. But even among U.S.-born Central Americans, the rates varied greatly by country of origin. People of Honduran and Guatemalan ancestry were more than twice as likely as people of Costa Rican ancestry to be in poverty (roughly 36 percent versus 14 percent, respectively).[39] These differences by nativity and ancestry translate into health disparities across subpopulations. The disaggregation calls attention to disparities that might otherwise not have been noticed, let alone explained or addressed, and knowing them deepens our understanding of which disparities to further assess and address.

Acculturation is the process of individuals and groups of one culture adopting, modifying, and borrowing traits, practices, and values from another culture.[40] This happens as immigrants move to new environments, leading to both positive and negative outcomes as the two cultures interact. Acculturation includes immigrant integration and assimilation, and broadly includes all factors of culture—language, religion, diet, social norms, and more. The experiences faced by newcomers fall along a spectrum, and standardizing the measurements of how different subgroups acculturate can help us better understand how experiences mediated by race, ethnicity, and culture affect health. Some aspects of native culture may be protective and may facilitate positive health outcomes, or some interactions between the native and introduced culture may lead to positive outcomes over time; these aspects may not be seen at the level of broad racial and ethnic categories, but at the level of cultures and nationalities.

❝❞

Some studies on immigrant health point to increased stress exposure the longer immigrants live in the U.S. which over time increases their likelihood to engage in health risk behaviors. There also may be specific cultural health factors prevalent in the beginning or changes that take place as a result of contact with culturally dissimilar people and groups related to integration that reduce or heighten their health risk.

Carmela Alcántara, Associate Professor of Social Work, Columbia University

Acculturating communities can be linguistically isolated, making them difficult to count, and therefore, they are underrepresented in the data. Stress can come from linguistic isolation. Jamil et al, in a piece about post-traumatic stress disorder among Iraqi refugees, pointed out that beyond surveys of the Census Bureau, there is no national, federally sponsored data collection that is consistently administered in languages other than English and Spanish.[41] In Census Bureau data on languages spoken at home, results showed that 49 percent of Haitian speakers lived in Florida, 48 percent of Punjabi speakers lived in California, and 39 percent of Bengali speakers lived in New York.[42] Linguistic isolation, and limited English proficiency, is associated with lower socioeconomic status and burden of health inequities. Most national surveys may undercount and underrepresent these health inequities among these communities.[43]

To improve the measurement of acculturative stress in population health research, some researchers are looking to biomarkers, which are pieces of biological information that are collected in social surveys, such as cortisol for stress, interleukin 6 for immune functioning, and heart rate for cardiovascular health. These biomarkers can help measure allostatic load, which in turn can help identify acculturative stress, which is the stress experienced from the process of acculturation.[44] Notably, there is low participation of communities of color in clinical trials that collect biomarker data because of negative historical experiences in research exploitation.

Citizenship and Legal Status

The ways that people formally experience living in the U.S. determine many aspects of their health, directly and indirectly, and indicators of their legal status are important for addressing health equity issues.[45] Undocumented communities are some of the most vulnerable to health disparities because they are ineligible for health insurance and preventive and supportive programs.

Citizenship and legal status are sensitive information, especially for communities fearful of surveillance and discrimination. Revealing undocumented status may lead to losing a job, being evicted from housing, or even deportation. The recent, currently unresolved proposal by the U. S. Department of Justice to add a question about citizenship to the 2020 census has provoked a negative response from many quarters that it would depress response rates and lead to an even larger undercount of the undocumented population, with all its implications for redistricting and allocation of federal funds. The undocumented population often has no, or subpar, access to quality health-care services, secure employment, and decent housing. For example, undocumented immigrants are prevented from enrolling in health insurance programs through the federal and state exchanges, and they are not eligible for Medicaid under federal guidelines.[46]

In this context, a population health survey that can safely and successfully elicit reliable responses from undocumented residents is a very important asset. The California Health Interview Survey (CHIS) has a few questions to measure citizenship and immigration. CHIS is the largest state health survey, collecting responses over the phone about health status, health conditions, mental health, oral health, health behaviors, access to and use of health care, health insurance, employment, and respondent characteristics. When asking, "Are you a permanent resident with a green card?" CHIS immediately reassures the respondent that, "Your answers are confidential and will not be reported to Immigration Services."[47] CHIS has asked this question since 2001, with response options of "yes," "no," "application pending," "refused," and "don't know." Additionally, the CHIS asks, "In what year did you become naturalized?" and, if the respondent is currently on a visa, "Is this visa or document still valid or has it expired?" Through these responses, CHIS can determine data on respondents without legal permanent residency status and without a green card to use as a proxy for the undocumented population.

> **When [health professionals] take a medical history, often the social and family history gets short. What neighborhood do they live in? When did they arrive in the U.S.? What social supports do they have? It's important to understand how that connects to community health. I'm going to have way more of an impact if this person is going to receive culturally competent care.**
>
> Yousef Turshani, Pediatrician and Former Chief Medical Officer of UMMA Community Clinic, Los Angeles

Although it can be difficult to collect citizenship and legal status data, the potential is immense for this part of individuals' background to inform population health disparities research. When it is combined with detailed information about race and nationality such as the CHIS is often also able to collect, a much more complete picture of these communities begins to emerge.

A migrant worker from Oaxaca living in a small town outside of Sacramento who has built a peaceful life without being bothered by I.C.E. He stays there because, "the money keeps me here." (*Eneas De Troya is licensed under* CC 2.0)

Methods of Estimating the Undocumented Population for Health

Rates of immigration and the status of immigrants are complex topics to measure, but that hasn't stopped groups such as the Migration Policy Institute from doing what they can. Categories of immigrants vary in terms of legal status, including asylum seekers, legal permanent residents, temporary student and work visas, and the undocumented. While the undocumented population is of growing social, economic, and political importance, this population is inherently difficult to study because of their limited contact with and representation in official data collection systems. This lack of measurement has led to a lot of informal or indirect knowledge and questions around the growth of the undocumented population.

The undocumented population is uniquely vulnerable to compounding social problems, such as lack of access to stable employment, safe housing, and health care, which can adversely impact the health of these individuals. Without documentation, these individuals are ineligible for health insurance and many other social safety-net programs, such as the Supplemental Nutrition Assistance Program (SNAP). By being uninsured, this population faces severely reduced options for receiving health care. This restricts their contact with health-care professionals and health-care systems, which leads to worsening health disparities because they are unable to address symptoms and illnesses as needed. If a good estimate of the undocumented population was available through data collection systems, researchers could better identify health needs of and disparities among this population.

The Migration Policy Institute, based in Washington, DC, is a leader in estimating the size of the undocumented population for public policy. Their Data Hub imputes unauthorized status using the Census Bureau 2010–2014 American Community Survey and 2008 Survey of Income and Program Participation (SIPP) data.[48] The Institute has used the Data Hub to research a variety of aspects of the undocumented population, such as an unauthorized wage penalty, effects of local enforcement efforts, and legal status and immigrant integration.

Other methods are available for estimating the undocumented population, such as the residual method, which subtracts the number of legal non-immigrants and legal permanent residents from the total number of noncitizens. Logical imputation methods identify likely documented noncitizens using legal or other predetermined criteria, again leaving a remaining pool of likely undocumented noncitizens. Small sample surveys in studies such as the Los Angeles Family and Neighborhood Survey, the National Agricultural Workers Survey, and the SIPP can help to provide this information about groups such as welfare recipients, government workers, veterans, and students.

Isela Garcia, 23, of Little Village, Chicago, came to the United States with her parents when she was nine years old, and spent Wednesday, August 15, 2012 at Chicago's Navy Pier, preparing her application for deferred deportation. (WBEZ/Peter Holderness *is licensed under* CC 2.0*)*

Internal Migration Patterns

Even within the country, a household's migration patterns over generations influences their accrual of wealth, as well as their health. The Great Migration of African Americans from the rural South to urban areas in Northern states in the first half of the 20th century was a pursuit of economic opportunity and an escape from Jim Crow segregation.[49] There are also secondary and tertiary resettlements of refugees and other immigrants to the U.S., as immigrants concentrate and disperse to share resources and build communities and networks.[50] A sizable portion of the residents of Appalachia, mostly of Scots-Irish ancestry, moved to Midwestern and Eastern cities and industrial towns from the 1960s through the present day.[51] Cyclical migration is ongoing, on and off tribal lands, by American Indians and Alaska Natives because access to services varies between the U.S. and tribal government health and social service entities.[52] All of these internal migration patterns have influenced health behaviors, risks, and outcomes. Tracking internal migration is an important component of building a complete portrait of a racial or ethnic subgroup in the U.S. Much of what we know is based on historical research, community studies, and first-hand accounts, for these internal migration factors are difficult to represent, standardize, and analyze on a large-scale statistical basis. Migration patterns are, however, compelling assets for health policy research when they can be obtained.

❝❞

With respect to the Black population, data from the 2001–2003 National Survey of American Life (NSAL) suggest that 4 percent of the Black population is very dark, 24 percent dark, 47 percent medium, 16 percent light, and 8 percent very light.

Jackson et al, in the research review on the Black population conducted for this project

Skin Tone and Within-Group Differences in Health

One of the ways that researchers have captured phenotypical differences, particularly but not exclusively within the multiracial population, is by measuring skin tone. This was also a central dimension discussed in the research review about the Black and African American population. Skin tone is an important measure by which to assess the role of discrimination in health because it measures how other people may view and therefore treat the respondent.[53] While health inequities between Blacks and Whites are widely known, skin tone variation within the Black population is associated with even greater health disparities than those across Black and White groups.

While there are some methodological challenges in capturing data on skin color, nearly everyone responds to questions of skin color when data are collected alongside other phenotypical markers, like eye color.[54] Researchers measure skin tone in multiple ways, including "objective" data from a reflectance meter and self- and observer-reported data on a continuum (e.g., very light, light, medium, dark, and very dark). The researchers on the Black and African American review recommend collecting both measures and emphasize the importance of matching the races of the observer and respondent.

Much of the existing research on skin color is focused on the Black population, but work has also been done on attitudes about skin tone within the Asian population, for which the multibillion-dollar skin bleaching industry is an indicator.[55] In addition, research on skin tone is evolving legal interpretations of discrimination and civil rights beyond race and ethnicity alone.[56]

The Role of Skin Tone in Discrimination and Health Inequities

Skin color and skin tone shape the lived experiences and health outcomes of Black people in important ways, and scholarly interest in this phenomenon has grown over the last several decades. There are wide, documented, intra-racial disparities in the social determinants of health, including income and education. Even when looking across similar occupations, for example, the earnings of the lightest-skinned Black men were comparable to those of White men and significantly higher than those of darker-skinned Black men. Lighter-skinned Black people also achieve higher levels of educational attainment than their darker-skinned counterparts, and research has found greater variability in education within the Black population than across Black and White populations. After adjusting for factors that may influence earnings, lighter-skinned Black people have higher adjusted earnings at the same level of education relative to darker-skinned Black people. According to the research review on the Black population produced for this project, skin tone is one of the key understudied analytic domains within the Black population, and has important consequences for health outcomes.

There are benefits of using multiple measures of skin color to understand the relationships between color, discrimination, and health. Objective measures of skin tone rely on readings from spectrophotometers, which are used to measure melanin levels in cancer research. Subjective measures are also available, notably interviewer-rated skin tone and self-reported skin tone. Ellis Monk argued that self-rated skin tone is especially well-suited to measure perceived discrimination and health because it acts as a form of "subjective social status" (e.g., a self-assessment of one's place in the social order).[57] Both subjective measures of skin tone are significantly associated with perceptions of discrimination: darker-skinned Black people are more likely to perceive discrimination than their lighter-skinned counterparts.

❝❞

In a growing multiracial society, color is still used as a marker of difference.

Trina Jones, Professor of Law at Duke University

Consistent with other research, Monk found that higher rates of perceived discrimination within the Black population were associated with poorer physical and mental health outcomes, as was the darkness of self-reported skin tone.[58] Importantly, Monk's findings challenge the race paradox in mental health. This paradox is supported by studies that found Black Americans have relatively similar or even better mental health than Whites.[59,60] But Monk found that both the darkest-skinned and the lightest-skinned African Americans have similar or worse self-rated mental health than Whites and worse self-rated mental health than medium-tone African Americans, who make up a majority of the Black population.[61] Importantly, discrimination by people in other groups as well as skin color discrimination within the Black population are significant predictors of health outcomes.

Gender, Sexuality, and Social Class

The demographic factors most commonly collected in health surveys, in addition to race and ethnicity, include gender and socioeconomic status. Life expectancy varies considerably by race and ethnicity; for example, it is higher for women than men across race and ethnicity. And one of the most established determinants of health inequities is socioeconomic status (SES), which also varies by race and ethnicity. In fact, racial health differences in heart disease mortality persist across SES categories.[62] Attention to the confluence of these factors, rather than studying one in isolation, is where the health research field should continue to move.

Most health surveys rely on a binary gender system where respondents select either male or female, but these categories mask the diverse experiences of people who do not fit neatly within them. Research suggests that transgender and gender nonconforming people, especially people of color, face health inequities related to social stigma, discrimination, and limited access to necessary health services and procedures. In the absence of nationally representative health surveillance surveys that capture data on gender identity, state and local surveys are filling the void.[63]

The California Health Interview Survey has collected data on sexual orientation since its inception, and in 2015, it started collecting data on transgender people. The 2015–2016 CHIS was the first population-based survey with a large enough sample size in the U.S. to use the standard two-step approach for measuring gender identity.[64] A CHIS policy brief concluded that compared to cisgender adults, "transgender adults are more than three times more likely to have ever thought about suicide, nearly six times more likely to have ever attempted suicide, nearly four times more likely to have experienced serious psychological distress, and more than three times more likely to have emotions that interfere with their relationships, social life, ability to do chores, and work performance."[65] The report emphasizes that transgender and cisgender adults are similar in many ways, such as demographic characteristics, but experience health disparities in mental health, disability status, and health-care access.

Small sample sizes hinder the ability to analyze the experiences of LGBTQ people by gender, race, and SES, but these intersections are ripe for future research and practice to reduce health inequities. The risk of HIV among Black gay and bisexual men poignantly illustrates the interconnectedness of these demographic factors. According to data from the Centers for Disease Control and Prevention (CDC) as cited in a June 2017 *New York Times Magazine* report, if current rates persist, the lifetime risk of HIV is one in 99 for all Americans and one in 11 for White gay and bisexual men. For Black gay and bisexual men, the number is one in two. In Jackson, Mississippi, the capital of the poorest state in the country, 40 percent of gay and bisexual men are living with HIV. For comparison, the world's highest rate of HIV is 29 percent in Swaziland.[66] Looking at these numbers by race or by sexual orientation or by gender in isolation masks their cumulative impacts on the health and well-being of diverse communities.

Without better representative data on the LGBTQ population that is collected alongside other demographic information, including race and ethnicity (at the level of subgroups where possible), language, and SES, it will be difficult to develop effective and comprehensive strategies to build an inclusive culture of health.

The Potential of Private and Public Data Assets

Not all efforts to improve and disaggregate racial and ethnic data call for collecting new information or conducting bigger surveys. Between government-generated data, administrative records, and privately owned data, a wealth of existing information exists about the health of communities that could be put to better use. A lot of potentially powerful data collected or authorized by the federal government goes unused because it is restricted to a small number of users. Effective access to these large databases is extremely limited outside of academia and certain government agencies, making it challenging for community-led efforts to use empirical evidence in their advocacy and practice. The federal data are often restricted when the confidentiality of respondents cannot be assured. While it is critical to ensure the data security and safety of respondents, it may be possible to do so while also relaxing the criteria for data access. At the very least, a clearly written process for accessing data and meta-data on certain racial and ethnic subgroups, or at small geographies, is necessary to advance the field.

The California Cancer Registry (CCR), a statewide population-based cancer surveillance system, is a good example of integrating and making effective use of existing resources. CCR links state and national databases and routinely updates vital statistics and causes-of-death information. To determine and refine data on Hispanic ethnicity, CCR employs many tactics such as (1) medical records on Hispanic ethnicity; (2) inference from the race and ethnicity of parents, maiden name, surname, birthplace, or death records; and (3) the National American Association of Central Cancer Registries' Hispanic Identification Algorithm. These methods help improve the classification of the Latinx ethnicity and overall knowledge of cancer incidence among the community, and provide an exemplar in linking existing data sources to fill in gaps.[67]

Another compelling case resulted from the disaggregation of enrollment in Affordable Care Act-sponsored health insurance, to identify and respond to disparities among Asian American groups. The Asian and Pacific Islander American Health Forum found that Korean Americans had a high uninsured rate of 23 percent, compared to just 12 percent for all Asian Americans. That high level for Koreans was a consequence of their high rate of self-employment in small businesses, so the Forum built

The All Dulles Area Muslim Society met with Former Secretary of Homeland Security Jeh Johnson to discuss ongoing efforts to safeguard local communities through partnerships between community organizations, state officials, and local officials on civil rights and civil liberties. *(U.S. Department of Homeland Security/Jetta Disco)*

a network to do specific outreach, education, and enrollment for them. Overall, such targeted efforts assisted 1,000,000 people and resulted in a 56 percent decrease in the uninsured among the Asian and Pacific Islander population.[68]

Although most of the attention of this project has been on government-managed information systems, a massive new trend worth tracking is the proliferation of health-related data collected, integrated, and controlled by private sources. In 2017, *The Economist* ran an article titled "The world's most valuable resource is no longer oil, but data."[69] With the growth of machine learning, algorithms can predict when a consumer is ready to make a purchase and, increasingly, the degree to which a person is at risk of disease.

Technology giants like Google, Facebook, and Apple benefit from user-generated data. The more users, the more data, and thus the greater the power to predict and influence decisions. But, private-sector data users do not operate in a vacuum apart from government; these worlds share information and build off each other, with different implications for community health and well-being. Many private firms have, in fact, built their businesses upon reuse and reformulation of government statistics.

The widespread adoption of electronic health records has generated massive databases. Murdoch and Detsky described the potential applications of "big data" in health care, which includes linking traditional health-related data such as family history or medication lists to personal data found on other sites, such as income, education, and neighborhood qualities.[70] Administrative data from hospital and health clinic systems can be linked to other data sets to fill gaps, draw correlations, and increase information.

With the proliferation of "big data," the private and public sectors face expanding opportunities to link personalized health and genetic screening data with information on the social determinants of health; however, ethical challenges accompany such possibilities. The place of race and ethnicity in these new configurations of information needs to be sorted out carefully, not only with respect to the broad categories, but also for the disaggregated subgroups. The potential is there for gathering valuable insights and increasing access. However, there is also a risk that highly specific background information can be used for targeting of dubious marketing or for exclusion from care or insurance options.

Our collection of next steps for improved disaggregation in each of the two domains of change—Methods and Government Policy—takes two forms. First, we offer several Guiding Principles, which are broad statements of values and priorities for action. These are followed by Recommendations, which are more specific or specialized and which identify a range of actors and direct opportunities which can be undertaken this year. In the case of Methods of Data Collection and Analysis, the recommendations are mainly about survey practices and even though the concepts could be similar across the country and across fields, the action would be relatively decentralized, with groups responsible for each survey usually having the authority and the need to act.

Guiding Principles for Methods of Data Collection and Analysis for Disaggregation

- **Methods of measurement of race and ethnicity should be selected to best serve the purpose of a given research or clinical activity.** Racial and ethnic identity is multidimensional, so researchers can and should deploy different ways of naming and measuring it. Most health surveys rely on self-reported race and ethnic identity, while health-care providers, including hospitals, often use third-party reports of race and ethnicity.

- **Health surveillance surveys should be well supported to be able to explore multiple dimensions of the intersection of race and ethnicity with other aspects of identity through oversampling, linguistic diversity, new questions, and other resources and techniques.** State and local health surveys, as well as the large federally supported endeavors, are in the best position to advance the collection of high-quality disaggregated data if they can get the resources and the political support to venture into these new areas.

- **Community voices should be integrated into the research design process, providing ideas, consultation, and feedback on the ultimate design, purpose, and intent of the research.** This authentic inclusion of community representatives will help ensure that the assets, strengths, and resilience of racial and ethnic identities are analyzed, as well as disparities and deficits. It will also improve the quality and representativeness of information.

- **A new source for high-quality, widely available survey data that represents key dimensions of the intersection of race, ethnicity, and health should be established.** Existing surveys should be strengthened, but ideally, a new, nationally representative, longitudinal-panel data set could function as an essential clearinghouse for the purposes of interdisciplinary population health disparities research. The data would serve to build the theoretical and practical foundation of knowledge around racial and ethnic subgroups and health.

Recommendations for Methods

1. **The research community should invest in research methods that address small "N" populations and small sample sizes.** Exploring, advocating, and refining research methods specifically targeted to these communities will generate more accurate measures of communities that are often underrepresented in the research and historically undercounted. These methods include field research, in-language surveys, ethno-racial sampling frames, accultured data tools, and pooled data sets. This pertains to the Population Centers at major universities and their "national" association, the National Network of State and Local Health Surveys, survey managers at the National Center for Health Statistics and Tribal Epidemiology Centers, and the philanthropic and government funders who support surveillance survey research.

2. **The research community should capture racial and ethnic self-identity with more standardized measurements, providing more nuanced and contemporary subgroup categories for respondents to choose from.** These measurements will better uncover factors throughout the life course that influence identity, such as the demographic environment surrounding multiracial individuals; the racial and ethnic context of an immigrant's home country; and the evolving norms, culture, and institutions of society in the U.S. This type of research may pertain to the members of groups such as the Critical Mixed Race Studies Association, the American Sociological Association, and numerous professional associations for ethnic studies research, as well as the National Center for Health Statistics and the Census Bureau. Again, the funders of social research should set a high bar and provide meaningful incentives for the improvement of these techniques.

3. **Survey managers and data users should determine the relative costs and benefits of sample size expansion and adding and modifying survey questions.** The effort to capture high-quality data about communities can be enhanced by careful evaluation of possible improvements to questions, sampling, outreach, and other steps in the survey process. This needs to be done in an environment of tightening budgets and during a trend of declining response rates, so it will be challenging and will often require trade-offs among the possible innovations: what is the higher priority for achieving better disaggregated data: more detailed questions versus broader or deeper outreach, for example. This process of estimating and then advocating for intelligent improvements to survey practice can be done in partnership with the funders, the survey managers, and key data users, such as the National Network of State and Local Health Surveys, the National Center for Health Statistics, the Association of State and Territorial Health Officials, and the National Association of County and City Health Officials, as well as data policy leaders at every federal agency.

4. **The health research community should collect clinical data on biochemical factors associated with sociocultural and environmental stress and, where possible, align it with survey data on socioeconomic factors, race, and ethnicity.** Data on biomarkers, such as allostatic load or other biochemical processes, can reveal associations among heterogeneous racial and ethnic subgroups for population health disparities research with an emphasis on mental health. This may pertain to the Collaborative Psychiatric Epidemiology Survey and surveys administered by the National Center for Health Statistics.

5. **Data providers and survey managers should provide more opportunities for researchers and data users to link health surveillance and administrative data sets to allow for more robust and accurate analyses.** Data from health surveillance surveys can benefit from linkages with data on education, employment, and other social determinants of health. When data are linked, more information is available to describe health behaviors and outcomes. It also reduces the need for each survey to collect the full scope of demographic information, reducing respondent burden and preserving valuable survey "real estate." Linking data sets creates opportunities for larger, integrated, and interdisciplinary analyses, which apply to government agencies, research institutions, and the public sector. For example, there is potential to link birth and death rate data with the American Community Survey to allow a better understanding of racial and ethnic misclassification, by contrasting third-party observation with self-reported identification. In the process of linking data, data providers and survey managers will need to strengthen existing protocols for privacy and security of respondents' identifiable information.

6. **Data providers and survey managers should connect users to the abundance of existing public and private data applicable to population health disparities research.** There are many challenges to accessing health-related data, due to eligibility requirements, patient privacy rules, credentials, clearance, and bureaucratic processes, and often these get in the way of making high-quality, updated data on race and ethnicity accessible for important research and policymaking purposes. In some cases, the need is to streamline access to multiple survey data sets, and in other cases, the challenge is around using the information to improve community outreach to targeted groups that may potentially benefit from the data. The process of making more data available without compromising key privacy concerns will be complex, but is well worth the effort. Private-sector leaders in health-care provision and insurance, such as Kaiser Permanente and Anthem Health, can be an asset in this effort. The organizations who are part of the National Network of State and Local Health Surveys, the National Center for Health Statistics, the Census Bureau, and the constellation of advocates who concentrate on the regulatory environment for health data can also contribute to making these sources available in an appropriate manner.

7. **The funders and managers of health surveillance surveys should add additional variables to surveys that more effectively reveal within-group differences.** These surveys about health and social determinants of health can reveal associations with health outcomes among racial and ethnic groups that may have more significance than associations between groups, such as skin tone, internal migration, or parents' country of birth. This can be done in consultation with communities, and in partnership with the National Survey of American Life, the National Latino and Asian American Survey, and other surveys that focus on heterogeneous groups.

8. **Advocates and the research community should thoroughly communicate the potential benefits, costs, and risks of increased data disaggregation among communities.** Disaggregation needs local champions and well-connected leaders to make the case for smart innovations. Particularly for racial and ethnic subgroups that are linguistically, geographically, and/or socio-economically isolated, the use of culturally relevant outreach will foster buy-in, build understanding, and recognize risks of greatest relevance to their communities. This can be done in partnership with culturally sensitive and appropriate communications and media avenues, such as the Black Immigrant Network, the National Latino and Asian American Survey, and the National Network of State and Local Health Surveys. The fellowship and support programs that aim to increase diversity and attention to equity in health research, such as those supported by the Robert Wood Johnson Foundation, the W. K. Kellogg Foundation, and others, should train and engage graduate students and younger scholars to improve and increase data disaggregation relevant to various communities of color.

Representative Ilhan Omar (MN-60B), the first Somali-American elected to office in the United States, in 2017.
(Wikimedia Commons / Leopaltik1242)

Government Policies that Enable and Enhance Data Disaggregation

A complex ecosystem of public policies and practices shapes the defining, collecting, analyzing, and reporting of racial and ethnic identity. These policies and practices affect all types of data, including health surveillance surveys, administrative data, clinical records, and other population surveys.

If stronger guidance and support for the disaggregation of data on race and ethnicity are to come about, they will emerge from actions taken at every level of this federated system. The federal government establishes the principal foundation by setting the terms and guidelines that affect the data it generates or supervises. Those minimum expectations provide a foundation upon which states, localities, and tribal governments can go further as they choose. State and local health departments may collect a wealth of data that is comparable and consistent with the terms of national statistics. Some states, localities, and tribes are already taking the lead, disaggregating further than the minimum expectations and pioneering innovative practices that address smaller and concentrated communities of racial and ethnic subgroups.

The states that are moving forward with these new progressive laws and practices for disaggregation are doing so at a time when the federal government's stance on updating the measurement of race and ethnicity is unresolved. In the absence of a response from OMB about the Interagency Working Group's proposed changes to the federal data standards, those ideas for modernization and improvement have not been adopted. Some states will move ahead on their own, and such efforts would reflect the broader national political environment in which a number of states are challenging the direction of Congress or the administration on various issues.

❝ ❞

Sometimes even when the data is there, there's hesitation to focus resources on the AAPI [Asian American/Pacific Islander] populations; cost and time of oversampling small populations is a factor but buy-in and political will is what is really holding us back.

Akil Vohra, Director of Strategic Initiatives, White House Initiative on Asian Americans and Pacific Islanders

Federal Data Collection on Race and Ethnicity

As noted earlier, the Office of Management and Budget standardizes the definitions to be used throughout all federal agencies and all data collection efforts receiving funding from the federal government. The OMB also generates guidelines for the U.S. Federal Statistical System. The Census Bureau is the primary statistical agency that collects, analyzes, and reports data about society and the economy, including racial and ethnic demographic data. The Census Bureau conducts these operations primarily through the decennial census and the American Community Survey.

Data collected on race and ethnicity helps guide the allocation of funding and resources to communities across the U.S., so the accuracy of this information has many implications for quality of life, access to services, and the enforcement of civil rights. The definitions of race and ethnicity have evolved in OMB policy and Census Bureau practice over time, in response to demographic shifts, social and historical events, and changing societal perspectives.[71]

Significant investments have been made in accurately counting and recording the race and ethnicity of the U.S. population. These definitions and policies are only the minimum requirements for federal activity, and entities are encouraged to be more detailed than these requirements. When that greater detail is obtained, the entity doing so is encouraged to maintain comparability by following guidelines on the "roll up" of racial and ethnic subgroups into larger categories. The U.S. Department of Health and Human Services manages several surveys that go beyond these minimum federal standards, and there are also examples at the state and local level.

1997 Standards for Data Collection on Race and Ethnicity and their Revision

The definitions in OMB's 1997 Standards for Data Collection on Race and Ethnicity[72] are intended to govern all federally funded data collection, and state, local, and tribal entities often use those definitions to maximize comparability with federal sources. While the OMB details the major categories, such as "Black or African American," it is the Census Bureau that details the subgroups, such as "Haitian," "Oromo," or "Afro-Caribbean."[73] When data on a specific subgroup is reported because the sample size is large enough, the end user should still be able to roll up to the OMB categories. This allows for more detailed race categories while also maintaining consistency in the major racial groups over time. This process as outlined in the 1997 Standards needs further guidance and standardization to maintain the integrity of subgroups when reporting data at multiple levels.

During the last revisions of OMB standards in 1997, the Native Hawaiian and Pacific Islander community was disaggregated from the Asian American community. The revision also clarified the definition for Hispanic, Latino, or Spanish ethnicity, and deferred taking a position on adding a Middle Eastern/North African race or ethnicity until further research was conducted.[74] Since then, a great deal of local research has been done, and the Arab American community in particular has been advocating for a Middle Eastern or North African racial/ethnic category in the decennial census.[75, 76]

❝❞

The fastest-growing Asian American subgroups based on the 2000–2010 U.S. census include the Bhutanese population with an 8,255 percent increase (15,290 Bhutanese reporting single race in U.S. in 2010), along with Nepalese (561 percent increase; 51,907 in U.S. in 2010), and Burmese (500 percent increase; 100,200 in U.S. in 2010) subgroups.[77]

N. Ponce, A. J. Scheitler and R. Shimkhada, in the research review on Asian American, Native Hawaiian, and Pacific Islander populations conducted for this project

In 2017, the OMB convened the Federal Interagency Working Group for Research on Race and Ethnicity,[78] which was tasked with updating the standards for the first time in 20 years. In those last two decades, the demographics of the U.S. changed dramatically, and so updating the standards would expedite an improved understanding of how the country is diversifying and improve the prospects for statistically valid disaggregation. (See Appendix D for more information on the Working Group's focus areas.)

The Federal Interagency Working Group has members from every agency in the Federal Statistical System. Their work is driven by research conducted by the Census Bureau, as well as a public comment process and consultation with community leaders. The working group has been receptive to grassroots advocacy from organizations and coalitions such as the Census Project, the Census Task Force convened by the Leadership Conference on Civil and Human Rights, and the Asian Pacific Islander American Health Forum. The Working Group is also tasked with updating terminology and developing guidance for implementation of the revisions. At the time this report is being completed, OMB has not responded publicly to the Working Group's recommendations, a response that had been scheduled to take place by December 1, 2017. This delay has meant that the 2020 census will, for the most part, need to revert to its 2010 questions and procedures regarding race and ethnicity, steps which the Bureau announced in early 2018. The eventual disposition of the Working Group's recommendations could have influence on other federal data systems and surveys, and the study of population and race, but it is not possible to predict how that will play out.

Decennial Census and American Community Survey

As required by the U.S. Constitution and as carried out every 10 years since 1790, counting all people living in the U.S. is the basis for allocating political representation. It is the only survey in the U.S. that requires everyone's participation. The population count is, of course, not just a tally of the number of people living in the U.S., but also collects demographic and economic information such as age and size of household. Accurate and detailed racial and ethnic identity is essential to help ensure fair representation in legislative districts, enforcement of civil rights, and the equitable distribution of resources. An inaccurate decennial count—one that systematically undercounts low-income communities, communities of color, and young children, among others—adversely impacts a well-functioning democracy and society.

❝❞

The Leadership Conference commends OMB for recognizing that continued racial and ethnic change in the United States requires an evolution in the statistical policy governing how we measure the composition of our population. That policy must strike a balance between the compelling individual interest in identifying oneself and society's interest in ensuring compliance with laws that uphold the civil and constitutional rights of all people.[82]

The Leadership Conference on Civil and Human Rights, in a letter dated April 28, 2017 to the Office of the U.S. Chief Statistician, Office of Management and Budget

The Census Bureau is facing severe budget cuts during the preparations for the 2020 census, which will if continued affect its ability to ensure a full, fair, and accurate count.[79] Typically, in the 10 years of preparation leading up to each census, the eighth, ninth, and tenth years receive amplified funding to meet the needs of field-testing, community outreach, and final operations (see Appendix C).

However, Congress is thus far allocating a flat line of funding in this eighth year of preparation. Already, Census Bureau operations have experienced adverse impacts: reduced capacity for outreach to small and hard-to-reach populations, limited end-to-end testing of the full questionnaire, and constrained testing of the new IT system.[80] Census budgets for FY 2018 and 2019 were both in play as this report was being completed. In 2017, the U.S. Government Accountability Office included the 2020 Census on their high-risk list, which labels federal programs that are "especially vulnerable to waste, fraud, abuse, and management, or need transformative change."[81]

The American Community Survey (ACS)—the sample survey that is the successor to the "long form" of the decennial census—gathers a wide breadth of data about individuals and households, and provides some of the best opportunities to disaggregate results by, among other things, countries of origin. The ACS is a nationally representative random sample survey conducted every month. ACS data often guide public programs and private sector investment decisions, because the data are generated more frequently than the decennial census and describe smaller geographic areas. The ACS replaced the decennial census long form in 2010 and publishes its data in 1-, 3-, and 5-year estimates. The data describe a much greater wealth of information than the decennial census, including data on education, language, and disability, all of which are sortable by the standard racial and ethnic categories; data on countries of origins for those born outside the U.S. were added recently. The richness of ACS data is immensely critical for researchers studying health disparities, as the data set includes much greater information about the social determinants of health that can be analyzed by race and ethnicity.

Surveys of the U.S. Department of Health and Human Services

Within HHS, the National Center for Health Statistics (NCHS) is the primary agency that generates the statistical information to guide HHS actions and policies to improve the health of the American people.[83] A few important surveys that NCHS administers are the National Health Interview Survey, the National Health and Nutrition Examination Survey, and the National Survey on Drug Use and Health. The CDC also generates data, as part of HHS, to save lives and protect people from health, safety, and security threats. The CDC administers BRFSS, a telephone survey about risk behaviors, chronic health conditions, and disease prevention. HHS data are also helpful to researchers, practitioners, and policymakers outside of the federal government.

HHS has taken the initiative in many ways for the disaggregation of data on racial and ethnic subgroups. Recently, NCHS conducted an oversample of the Native Hawaiian and Pacific Islander community in the National Health Interview Survey.[84] The data were released in 2014, and it was the first time the Native Hawaiian and Pacific Islander community could access and use disaggregated information that could be compared to national data. Similarly, CDC conducted an oversample of the American Indian/Alaska Native community in BRFSS in 2017, through additional interviews in 11 states with high proportions of American Indians and Alaska Natives.[85] These oversamples better represent these racial and ethnic subgroups that would otherwise be masked in the larger data sets.

66 99

Since 9/11, the [MENA] community has been savagely stereotyped and felt targeted and under siege, and this relates to how they open up and share info with the outside world. Still, there is wide support across the community for the [MENA] category because the benefits outweigh the costs.

Hassan Jaber, Executive Director and Chief Executive Officer of ACCESS (Arab Community Center for Economic and Social Services), and member of the National Advisory Committee of the U.S. Census Bureau

Imperative for Respondent Confidentiality and Data Security

The appeal of more disaggregated data has run up against an environment in which the people whom the better data would most assist may be less willing to share their information with authorities. The current political context has created heightened concerns over the safety and confidentiality of data, fear of targeted surveillance, and prospects of discrimination based on race and ethnicity.

In 2017, an operational test from the Census Bureau observed an increase in respondent confidentiality concerns, despite survey questions having nothing to do with personally identifiable information.[86] When respondents were asked about their willingness to participate in the 2020 census, respondents would reply, " Does it make a difference if I'm not a citizen?" This was seen particularly among immigrants or those who live with immigrants, communities of color, and marginalized communities.

Apart from the census, undocumented people may be wary that information they provide even in a benign, service-based setting, might be obtained by immigration enforcement.[87] This wariness has consequences at a larger scale—when communities are concerned about surveillance and their safety, they can retreat from public life and reduce the utility of necessary services, such as health care. This retreat can worsen health outcomes by reducing the efficacy of preventive health measures, closing off access to help, leaving symptoms and ailments untreated, and diminishing windows of opportunity to address health concerns.[88]

The fears and stress experienced by some immigrants and other groups reinforce the need for respondent confidentiality and data security. Communities can be reluctant to answer surveys, even from sources that have good reputations and intentions.[89] Unfortunately, there is a precedent of misuse of confidential data—the Census Bureau shared information with the Department of War to identify Japanese Americans for internment during World War II.[90] Since then, data confidentiality laws have been strengthened, but the historical memory and skepticism persist in some communities. With recent references to a national "Muslim registry," although it is not close to being enacted, and with a proposal having been made by the U.S. Department of Justice and approved by the Secretary of Commerce to add a citizenship question to the 2020 census, many groups fear for their security when they provide answers to surveys or otherwise make personal data available to authorities.[91, 92]

Data are more secure and confidential when personal identifiers are anonymized, processes for data access are rigorous and transparent, and errors and breaches in protocol carry significant consequences. The data should not reveal personal identifiers of individuals without consent. The data may be publicly available and helpful for research, but with strict privacy of respondents maintained.

Sensitivity of Counting the Middle Eastern or North African Community

Leaders in the Middle Eastern or North African (MENA) community have for decades advocated to include MENA as a new racial or ethnic category in the Census Bureau. Various organizations representing the MENA American populations, including the Arab American Institute, have been working with the Census Bureau to bring awareness to this issue since the 1980s and to address what is believed to be a significant undercount of these individuals. When the federal standards to measure race and ethnicity were reviewed in 1997, the OMB stated that additional research was necessary to explore the MENA category. On the 2010 census, many from the MENA population wrote in "Middle Eastern" or "some other race" as an ethnicity, but the census still classifies them as White. According to the Institute, adding a MENA category would allow for a more accurate count, leading to better outcomes for federal spending in many areas. With respect to health, "...because researchers are unable to disaggregate the current data of most persons with MENA origin, it is impossible to address diseases that are ethnic-specific, such as lactose intolerance or the prevalence of diabetes among Arab Americans. Inclusion on the census will foster greater access to health information and services for Arab Americans, as well as funding for services for the elderly and disabled."[93]

However, along with this interest in the MENA category has come heightened concern and sensitivity about increasing the visibility of the population in this political environment.[94] Post-9/11 sentiment and the current increase in Islamophobia threaten the security and integrity of collecting disaggregated MENA data. Personal data has on occasion been shared and used to profile and conduct surveillance. An example of this surveillance occurred in 2002, when the New York Police Department's Demographics Unit tracked individuals' records to identify suspected terrorist "hot spots."[95] This unit extensively mapped out communities, including searches based on license plates found within mosque parking lots, community gathering spots for specific ethnic subgroups, and the personal information of participants in Muslim student groups. The program also extended to Muslim communities within 100 miles of New York, in New Jersey, Connecticut, and Pennsylvania. A contentious case of census data on ethnicity being released occurred when the Census Bureau was found in 2004 to have shared data on where Arab Americans lived by city and zip code with the U.S. Department of Homeland Security, even though the information about specific individuals was not made available.[96]

To address concerns such as these, the Arab American Institute notes on its website that it will emphasize the importance of the laws and protections put in place to keep this data confidential. By law (Title 13, U.S. Code), the Census Bureau keeps personal data confidential, such that individual responses cannot be shared with anyone, including law enforcement agencies or other federal agencies. The Census Bureau keeps the personally identifiable information of decennial census respondents private for 72 years.[97] Furthermore, the Census Bureau does not share statistical totals for geographic regions that are so small that the information could identify individuals.

A man in prayer at the 2017 USDA Interfaith Iftar Celebration. *(U.S. Department of Agriculture / Lance Cheung)*

State and Local Data Collection on Race and Ethnicity

States and localities have a range of responsibilities for which they need information on the specific demographics of their jurisdictions, and they are increasingly taking steps to disaggregate the information more thoroughly. Four states are majority people of color: California, Hawaii, New Mexico, and Texas.[98] Many communities of racial and ethnic subgroups are regionally concentrated across the U.S. and many ethnic communities are both locally concentrated and widely dispersed across the U.S., such as the Cambodian American communities in Lowell, Massachusetts, and Long Beach, California.[99] The Twin Cities in Minnesota are home to the largest populations of Somali Americans and Hmong Americans.[100, 101] Southeast Michigan is home to the largest concentration of Arab Americans, notably Lebanese Americans and Chaldeans in Dearborn and other suburbs of Detroit.[102] California has a large and internally diverse population of Pacific Islanders, most prominently in San Diego and Los Angeles counties. Such concentrations and the resulting advocacy of those groups have led to greater attention to disaggregation in these states.

Many of these recent efforts have been concentrated in the health and education sectors for the Asian American, Native Hawaiian, and Pacific Islander populations, although efforts are not limited to these sectors or populations. Whether it be based in health, education, labor, or other issues, there is a growing network of policies and practices to identify the nuances of the experiences of racial and ethnic subgroups at smaller scales.

❝❞

It's hard to find the numbers on Arab Americans because they are included in the White population. We looked nationally and found only 34 reliable studies on Arab Americans, and 26 of them were conducted in Michigan.

Hassan Jaber, Executive Director and Chief Executive Officer of ACCESS (Arab Community Center for Economic and Social Services), and member of the National Advisory Committee of the U.S. Census Bureau

Initiatives in Health and Education

The Asian American, Native Hawaiian, and Pacific Islander populations recently received policy wins for data disaggregation in a number of states.

- In California, the Accounting for Health and Education in API Demographics Act (AHEAD Act) was signed into law in September 2016.[103] It requires the California Department of Public Health to collect more detailed data on Asian American, Native Hawaiian, and Pacific Islander subgroups, such as Bangladeshi, Malaysian, Taiwanese, Fijian, and Tongan.[104] These data will include rates for major diseases, leading causes of death per demographic, and pregnancy rates.

- In Rhode Island, the All Students Count Act was signed into law in June 2017.[105] It requires the State Department of Elementary and Secondary Education to disaggregate Southeast Asian subgroups whenever they collect demographic data on students. These additional subgroups include Cambodians, Filipinos, and Laotians.

- In Minnesota, the 2016 All Kids Count Act passed with bipartisan support; it enhances student achievement data. In 2017, the advocates and policymakers strengthened the legislation to ensure that key student groups are included and clarify the implementation requirements and timeline.[106] The data includes five new ethnic categories that have over 1,000 students and data on students' home language, immigrant or refugee status, and the students' history with foster care.[107]

These three states are recent highlights from a national trend to pass data disaggregation legislation in education and health. But two states with a longer history in addressing these issues are Michigan and Massachusetts. In 2001, the Michigan Department of Community Health designed a survey in conjunction with the Arab Community Center for Economic and Social Services (ACCESS). The Health Survey of the Arab, Muslim, and Chaldean American Communities in Michigan was the first of its kind to quantify the needs and health conditions of these subgroups considered "White" in federal statistics. The survey was translated into Arabic, and participants were recruited by a convenience sampling through 34 different community centers across Michigan, such as ACCESS. While this survey has not been regularly administered since 2001, it detailed poorer health for the Arab American population than that of the general population, and discerned trends between Arab and Chaldean communities.[108]

In 2000, the Massachusetts Department of Public Health developed an alternative data collection tool for collecting the race, ethnicity, and language preference of patients.[109] The Boston Public Health Commission, the Division of Health Care Finance and Policy, and the Massachusetts Department of Public Health were early adopters of using these forms, training personnel, and developing online training tools. The data collection form encouraged self-reported data and offered additional racial and ethnic subgroups as options, such as Cape Verdean, Haitian, and Puerto Rican, and languages such as Portuguese and Albanian. The disaggregated data helped uncover differences within all the broad race categories for health outcomes, such as the incidence of cesarean delivery being three times higher in Asian Indian women in the U.S. than among Cambodians.[110]

University students have also been organizing. At the University of California-Los Angeles, the Count Me In! Campaign advocated for and won data disaggregation in enrollment and admissions data. Students led a University of California system-wide campaign for nine months to add 23 different Asian and Pacific Islander ethnic groups to application, admissions, and enrollment forms, in addition to the nine groups already offered within this category. The data helped evaluate the flaw in a "model minority myth" of uniformly high academic achievement across the Asian and Pacific Islander population.[111] In 2017, there was traction in the University of Michigan system to add the MENA identity category to forms and applications. Several students and faculty members have organized for over a decade, and in a meeting with the Board of Regents, they presented their case under the banner #WeExist, after receiving support from the Literature, Science, and the Arts Student Government earlier in the fall.[112] With MENA data, the students hope to support student recruitment and retention efforts and identify trends in the hiring of faculty, staff, and administration and in bias incident reporting.

All of these advocacy campaigns, legislation, and administrative changes show the high energy and broad diversity of the movement to generate better disaggregated data. Those efforts could be multiplied many times over with the implementation of the recommendations that have arisen from the deliberations of this project—recommendations that are presented in the following section.

In 2016, Southeast Asian American students from UC Davis came to Sacramento for the All Californians Count Community Rally to support AB 1726. *(Southeast Asia Resource Action Center (SEARAC))*

Disaggregation of Asian American, Native Hawaiian, and Pacific Islander (AANHPI) Communities in California: Organizing and Opposition

The Accounting for Health and Education in API Demographics (AHEAD) Act (AB 1726) was signed in 2016 by California Governor Jerry Brown. The act mandates the collection and reporting of detailed demographic data of Asian Americans and Pacific Islanders by the Department of Public Health.[113]

The passage of the AHEAD Act was a great success for the AANHPI community in California, and has built momentum for similar policy proposals in other states. This legislation goes beyond the minimum expectations set by OMB for collecting tnd reporting AANHPI data. After 1997, when OMB disaggregated Native Hawaiians and Pacific Islanders from the larger Asian American community, the Department of Health and Human Services estimated that it took eight years for all federal agencies to comply. Advocates of AB 1726 and similar pieces of legislation across the U.S. are now encouraging state and local-level advances in data disaggregation, rather than just relying on the slow pace of change at the federal level.

Assemblyman Rob Bonta sponsored AB 1726 to collect disaggregated data on 10 AANHPI subgroups with a population of 439,809. In the original bill's text, it would have applied to both public health and education systems to address both health disparities and the achievement gap. AB 1726 was co-sponsored by the Southeast Asia Resource Action Center (SEARAC), Asian & Pacific Islander American Health Forum (APIAHF), Empowering Pacific Islander Communities (EPIC), and California Pan-Ethnic Health Network (CPEHN). The coalition collected signatures in support from over 100 organizations, and 500 individuals from 25 states, Guam, and even New Zealand.

However, the advocacy campaign for AB 1726 revealed a divide within the AANHPI community. A vocal contingent, mainly composed of recent Chinese American immigrants, opposed data disaggregation. Opponents protested in front of the Capitol, placed ads in local newspapers, and published articles advocating against data disaggregation. The coalition of AB 1726 co-sponsors was surprised by the intensity and number of these protestors, for its emergence as a wedge issue within the AANHPI community was unexpected.

The opposition used WeChat to communicate and organize. WeChat, a Chinese-language social media platform, is a hybrid of text messaging, media sharing, and group chats. On WeChat, users had conversations conflating data disaggregation with segregation, insinuating that disaggregated data would be used to target communities as happened to Jews and other minorities under Nazi Germany and to Japanese Americans interned by the U.S. government in the 1940s. WeChat's users also associated disaggregated data with "affirmative action policies" that occurred during the Cultural Revolution in China, which were vastly different and more prescriptive than affirmative action policies in the U.S. Although WeChat spread misinformation about data disaggregation with regard to the AB 1726 coalition, it was also a highly effective, in-language, culturally relevant, social media platform.

This is a valuable lesson about organizing and advocacy among immigrant and linguistically isolated groups, where community outreach and educational campaigns must consider a broader cultural context, anticipate adverse reactions, and distribute multilingual materials to reach a wider audience. While AB 1726 passed and will generate benefits to disaggregated data for health, the issue may continue to divide groups and weaken the case for systems-level change.

Opposition to Disaggregation

Several sources of doubt and opposition to data disaggregation have arisen as the movement to promote it has become more prominent. For some, the whole concept of being attentive to race and ethnicity in public policy is contentious. Data disaggregation has thus been criticized in the context of attacking so-called "identity politics," with the argument that making such distinctions can lead to unfair advantages between groups and to further cultural fragmentation of U.S. society.

There is also some wariness within communities of color that disaggregation could dilute their overall size, decrease their visibility, and thereby reduce their political power and influence. The U.S. pan-ethnic term "Hispanic" uniquely aggregated many Spanish-speaking racial and ethnic subgroups together, in a way that would not have occurred otherwise and is not done similarly in other countries. Univision, and other Spanish-language media outlets, celebrities, and politicians were invested in coining the term "Hispanic" to unify once disparate groups in the 1970s and 1980s. Many Spanish-speaking subgroups (e.g., Mexican, Puerto Rican, Cuban) were dispersed unevenly throughout the U.S., and despite regional concentrations, did not show up fully in the national data. Aggregating, not disaggregating, into the pan-ethnic "Hispanic" community allowed them to become sizable enough to gain political, economic, and social power.[114]

Opposition to data disaggregation has also been expressed on a more visceral level. In 2016, the California legislature was deliberating a bill about disaggregation of public health data about the Asian and Pacific Islander community. In front of the Capitol in Sacramento, and in local newspapers and social media, opposition leaders used fear tactics, spread misinformation, and preyed upon limited-English proficient communities to build opposition to the bill. Similar sentiments were expressed the following year in Massachusetts.

With a few exceptions, these sources of opposition have not been saying directly that raising the empirical visibility of the health and social conditions faced by an overlooked ethnic group would, in a literal sense, be a bad thing to do. Rather, they are seeking to make ideological points that are more part of the broader "culture wars" and general political divisions in the country than making a critique specific to public health or demographic research.

Resource Requirements for Generating High-Quality Data

For population surveys, the detailed enumeration of racial and ethnic subgroups can be an expensive and complex operation. Asking more questions and reaching more people both require more resources. Better pretesting, larger and more stratified samples, surveys administered in more languages, additional items in limited questionnaire space, greater outreach efforts to hard-to-contact households, more diverse and culturally competent staff, and more sophisticated coding to integrate previously unrelated databases are all improvements that will take larger budgets to be realized.

State and local health surveys run on tight budgets, and expansion can usually only come incrementally. Several survey managers reported that they raise funds each year to keep and add new questions. When they are unable to secure funding for a year or two, it can compromise the ability to keep the same questions and compromise the ability to make comparisons over time. Some survey operations also fundraise to add languages in which the questionnaires are administered.

Sometimes, to get disaggregated results, survey managers need larger sample sizes in targeted communities. Oversampling is commonly used to increase the reliability of survey results about small populations, but the cost of oversampling a geographic area or an ethnic or language group can be significant.

Limited and fluctuating funding have so far prevented many state and local health surveys and ethnic health equity organizations from maintaining the high-quality and longitudinal data sets necessary to understand disparities faced by smaller groups. But those surveys that have been able to expand, including the California Health Interview Survey, are showing that with enough philanthropic and government support, a great deal of new information and insight can be obtained.

The same resource challenges pertain to federally funded health research grants, from the National Institutes of Health, CDC, and other agencies, and to government-administered surveillance surveys. Oversampling, more extensive outreach, additional languages, and new questions comprise the mechanisms for disaggregation, and all of them have costs. Budget allocations that will generate the means to reach more people will be made if the relevant communities and the experts in research and health practice all make an effective case for the resources.

"

The Census Bureau makes the most comprehensive data collection effort and has made numerous provisions to ensure representation of Asian Americans through outreach efforts, in-language interviewing, subgroup categorization of Asian Americans, and oversampling in some areas.

N. Ponce, A. J. Scheitler and R. Shimkhada, in the research review on Asian American, Native Hawaiian, and Pacific Islander populations conducted for this project

Guiding Principles for Government Policies

- **Agencies supporting or conducting health surveillance surveys and other population surveys should make the collection and analysis of more finely disaggregated data about race and ethnicity a high priority.** Throughout the federal and state governments are numerous departments with the opportunity and resources to advance this commitment, and the experiences described in this report show that that commitment will pay off with a deeper understanding of health behaviors, outcomes, and social determinants.

- **Health-care and social-service institutions should receive adequate funding and technical assistance to build their data capacity.** These investments need to be made in patient intake systems, administrative and professional staff training, and IT systems to allow full compliance with data disaggregation requirements and updates.

- **Individually identifiable information about survey respondents, especially of underrepresented and marginalized groups, must be protected to ensure data security and respondent privacy.** These issues are especially important for communities threatened by civil rights and oversight abuses, such as the LGBTQ community and undocumented immigrants. Many existing policy measures require full compliance and enforcement, as well as continued strengthening. These measures for data security and respondent privacy defend against growing threats in cybersecurity and data misuse. It is also important to communicate the efforts and gravity of these measures to build public confidence and trust in survey management to ensure high response rates.

- **The relationships between the U.S. government and the American Indian/Alaska Native tribes should protect and build the integrity of research activity and data collection about tribal areas and tribal citizens.** The Census Bureau actively consults with tribal leaders and representatives, and research ethics principles protect tribal areas and citizens. These practices should continue to grow and value the tribal sovereignty of data and information about tribal communities.

- **The next generation of health researchers and practitioners should be actively recruited and well-trained with diversity, equity, and inclusion as values alongside the field of data disaggregation.** Programs with targeted funding and training will help transform the field of health equity through greater awareness and representation of racial and ethnic subgroups, among front-line staff, researchers, academics, practitioners, and advocates.

- **The consequences for policy and funding of the different ways of reporting race and ethnicity should be clearly defined and broadly understood.** The complexity of categorizing individuals' race and ethnicity, including single-race data, multiracial data, and tribal identification, affects how accurately population sizes are measured. For example, the Census Bureau reports "single-race alone" and "single-race alone or in combination with other races and ethnicities" to offer varying degrees of detail. Using "single-race alone or in combination with other races and ethnicities" would lead to larger population sizes that account for multiracial respondents. The implications of these choices are increasing in significance as the multiracial population grows, yet they are largely not understood by the public or many decision makers.

- **Investments from survey funders, philanthropies, and government agencies should be used to support data disaggregation efforts across sectors.** From recruiting the next generation of researchers and practitioners to evaluating the use of newly generated data, substantial and sustainable investments will optimize the data's utility to the field of population health disparities.

The National Network of State and Local Health Surveys

The National Network of State and Local Health Surveys is a group of survey leaders, data users, policymakers, and advocates who understand the importance of high-quality, local-level, population health data to all areas of public health. The Network organizes programming that supports quality state and local population health surveys and the use of data. Network members include leaders of state-based and local health surveys, such as the California Health Interview Survey or the New York City Community Health Survey, as well as national representatives from the Behavioral Risk Factors Surveillance Survey and the National Health Interview Survey. Managers of these surveys identified several challenges to data disaggregation, including issues of stakeholder interest, funding, the scarcity of questionnaire "real estate," sample size, and survey methodologies. Though the collection of race and ethnicity is roughly similar across the different surveys, collection is not completely standardized, making comparisons across surveys all but impossible. Importantly, many of these survey leaders are innovating new ways of capturing more detailed racial/ethnic data like ancestry, birthplace, and parent's country of birth.

Recommendations for Government Policies

1. **Congress should fund the U.S. Census Bureau's budget to adequately maintain and improve operations for the 2020 Census and all other surveys.** The success of the decennial census requires adequate funding to cover the extensive costs in the final years leading up to the decennial count, to ensure that the new technologies are deployed correctly, and to ensure outreach is sufficient. A well-funded census will ensure a full, fair, and accurate count of everyone living in the U.S. Funding for the American Community Survey and other surveys is very important as well. Congress should respond to the need to fully fund the census, as it has been expressed by the professional research community, leaders of many sectors within private enterprise, and a host of constituents concerned with civil rights and health equity, including the Asian Pacific Islander American Health Forum, the Arab Community Center for Economic and Social Services, the Leadership Conference on Civil and Human Rights, and others.

2. **The U.S. Census Bureau and the Office of Management and Budget should improve the documentation of race and ethnicity in federal data collection.** The Federal Interagency Working Group's revisions to the Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity address question formatting and nonresponse, classification of people of Middle Eastern or North African race/ethnicity, additional minimum reporting categories, and relevance of terminology. All of these are important steps for data disaggregation for health equity. OMB should respond to the Working Group's recommendations, and the Census Bureau and OMB should continue to consult with community representatives and leaders, researchers, policymakers, and advocates.

3. **The Office of Management and Budget and the U.S. Census Bureau should develop protocols for using data disaggregation consistently throughout the collection, analysis, and reporting of racial and ethnic subgroup data.** These protocols will encourage thorough execution of the federal data standards on race and ethnicity for greatest comparability and knowledge. These protocols can address challenges across sectors, such as how to enumerate free text responses, how to analyze "only Hispanic" and "Hispanic multiracial" subgroups apart from ethnic categories, and how to report a "roll up" of racial and ethnic subgroups when adequate data are unavailable to report more detailed information. The OMB and Census Bureau can generate these protocols with consultation of community members, data users, and researchers, and they can be applicable to data policy leaders at every agency, such as the departments of Health and Human Services, Housing and Urban Development, Transportation, and Labor.

4. **The research community, including Institutional Review Boards, should standardize enforcement of existing policies that facilitate research processes over activities on tribal lands and with tribal citizens.** There are policies pertaining to research ethics designed to respect the government-to-government relationship between American Indian and Alaska Native tribes and the U.S. government. Ensuring that these policies are enforced will protect tribal lands and tribal citizens from unethical research practices. Tribal elders, tribal leaders, and tribal epidemiology centers are all important decision makers and thought leaders to include in the conversation.

5. **Advocates, policy leaders, community members, and influencers should articulate strong arguments in support of data disaggregation as a tool for advancing health equity.** A solid foundation of evidence exists about the advances in data disaggregation, its positive benefits, and the innovations needed to further enhance this field. These arguments should convey the urgency for action and the costs of inaction. The benefits to society of enhancing health equity are economic as well as moral and political, and the case should be made on all fronts. Advocates in the realms of civil rights, criminal justice, disability, immigration, reproductive justice, youth organizing, and other areas can make the case and spread the word of how health equity can be enhanced with better data. The case should be made not only nationally but in the states, as shown by the recent examples of leadership in California, Minnesota, Rhode Island, and several other states.

San Jose, California, has the largest Vietnamese population in the world outside of Vietnam, along with high concentrations of other Asian Americans, Middle Eastern Americans, and Latinx Americans.
(Richard Masoner and www.cyclelicio.us)

Conclusion

The importance of achieving a higher level of data disaggregation is clear to people who have worked with—and come from—the communities that have been overlooked, and to researchers and policymakers who have sought the information necessary for addressing health inequities. It is a call for compelling information, evidence, and stories of many kinds. But this is still a relatively new issue for most people. Making changes in systems, policy, and practice will require broader and deeper understanding among many new constituencies and groups of experts. It will require building the public will to bring new priorities to the fore in government, and for intelligent, progressive uses of new data sources managed by the private sector.

The case for data disaggregation can unite many groups around a common agenda. This agenda can build support within and among racial and ethnic subgroups, across research disciplines and data users, among policymakers and decision makers, as well as throughout the greater society. It can broaden people's consciousness and understanding of how race and ethnicity affect health.

Plenty of positive examples in research, health practice, and policy change make the case for data disaggregation, and a fair amount of momentum has been built as a result. At the same time, our exploration for this project took place during a time in which the larger context for the use of data took some darker turns. The current period is characterized by fears about how information about immigrants might be misused and new awareness of the potential for abuse of "big data" of various kinds. The best responses to this environment will be made through generating community-level support and building public and expert awareness, not only by showing "proof of concept" for innovative techniques, but also by changing narratives, increasing transparency, and building trust. Everyone cited in this report has a job in moving the data disaggregation effort forward: community-based advocates, grassroots organizers, survey managers, researchers, health systems administrators, elected officials, and leaders in philanthropy. Our hope is that the ideas, connections, and strategies detailed in this report have been set forth in a way that helps all of us move ahead.

Notes

1   Irma T. Elo, Neil Mehta, and Cheng Huang, "Health of Native-Born and Foreign-Born Black Residents in the United States: Evidence from the 2000 Census of Population and the National Health Interview Survey," (*PARC Working Paper Series, WPS 08-04*, 2008).

2   Comments of Kathy Ko Chin, President and CEO of Asian Pacific Islander American Health Forum, at a convening for this project, Washington, DC, June 29, 2017.

3   The research reviews can be found on the web page for this project on the PolicyLink website (www.policylink.org/our-work/community).

4   Robert Wood Johnson Foundation, "Culture of Health Action Framework, Evidence for Action", http://www.evidenceforaction.org/what-culture-health (accessed February 1, 2018).

5   The sixth research review was conducted by a team of researchers at the University of Edinburgh, University of Alicante, and the University of Copenhagen. The review compared how the European Union, Aotearoa New Zealand, Canada, Denmark, Great Britain, Hungary, Malaysia, and Plurinational State of Bolivia, collect ethnic classification on official population censuses or registers. The level of detail in ethnic classification was assessed using the U.S. OMB Standards for the Classification of Federal Data on Race and Ethnicity. For more information, see the international research review by Raj Bhopal, Emma Davidson, Nazmy Villarroel, Pamela Pereyra-Zamora, Allan Kransnik, and EMEHRG, *Heterogeneity/Granularity in Ethnicity Classifications outside the United States (HGEC project)* (The University of Edinburgh, Edinburgh, Scotland, 2016), http://www.policylink.org/sites/default/files/HGECproject12102016.pdf (accessed January 25, 2018).

6   Centers for Disease Control and Prevention, "CDC Health Disparities and Inequalities Report—United States, 2011," U.S. Department of Health and Human Services, January 14, 2011. https://www.cdc.gov/mmwr/pdf/other/su6001.pdf (accessed December 21, 2017).

7   Bill Chappell, "Census Finds a More Diverse America, As Whites Lag Growth," *National Public Radio*, June 22, 2017.

8   Douglas D Heckathorn, "Respondent-Driven Sampling: A New Approach to the Study of Hidden Populations," *Social Problems* 44 (May 1997): 174-199.

9   Sarah Dewees and Benjamin Marks, *Twice Invisible: Understanding Rural Native America*, (First Nations Development Institute Research Note #2, April 2017).

10 Desi Rodriguez-Lonebear, "Building a Data Revolution in Indian Country," In T. Kukutai and J. Taylor (Eds.), *Indigenous Data Sovereignty* (Australia University Press, 2016).

11 "Also consider the U.S. Department of Education 2007 Guidance on OMB Racial Classifications (https://www2.ed.gov/policy/rschstat/guid/raceethnicity/questions.html). While the USDOE will continue to collect data on American Indian and Alaska Native (AI/AN) students whether or not they report a racial/ethnic status that is in combination with other racial/ethnic groups (e.g., Hispanic/Latino, White), the USDOE will only report AI/AN specific data for students who indicate they are not Hispanic/Latino ethnically and select only American Indian and Alaska Native as their race. American Indian/Alaska Native students who indicate that they are also Hispanic/Latino ethnically will only be reported in the Hispanic/Latino category. Regardless of whether they indicate Hispanic/Latino ethnicity, American Indian/Alaska Native students selecting an additional racial category will only be reported as multiracial. The effect is major and detrimental at local, state, and national levels as AI/AN communities have historically relied on USDOE data as a quality source of information for planning and development efforts (see the 2012 NCES STATS☐DC Presentation prepared by NCAI and NIEA at https://nces.ed.gov/whatsnew/conferences/statsdc/2012/session_V.asp). There is a similar problem with the use of redistricting data used after each decennial census, and should the Census Bureau make Hispanic ethnicity equivalent to a race in the 2020 census as it appears to be planning, this could have a dramatic effect on the counts of the AI/AN population, especially in a number of large metropolitan areas." Excerpt from National Congress of American Indians Policy Research Center, *Disaggregating American Indian & Alaska Native data: A review of literature* (Robert Wood Johnson Foundation, 2016).

12 National Congress of American Indians Policy Research Center, *Disaggregating American Indian & Alaska Native data: A review of literature* (Robert Wood Johnson Foundation, 2016).

13 Urban Indian Health Commission, *Invisible Tribes: Urban Indians and Their Health in a Changing World* (Urban Indian Health Commission, 2007).

14 M. L. Castor, M. S. Smyser, M. M. Taualii, A. N. Park, S. A. Lawson, and R. A. Forquera, "A Nationwide Population-Based Study Identifying Health Disparities between American Indians/Alaska Natives and the General Population Living in Select Urban Counties," *American Journal of Public Health* 96 (2006): 1474-1484. Cited in National Congress of American Indians Policy Research Center, *Disaggregating American Indian & Alaska Native data: A review of literature*.

15 Pew Research Center, "Multiracial in America," Social & Demographic Trends, June 11, 2015. http://www.pewsocialtrends.org/2015/06/11/multiracial-in-america/.

16 Nicholas Jones, "Collecting and Analyzing Data for the Multiracial Population in the United States: Research by the U.S. Census Bureau," (presented at PolicyLink Convening on Multiple Racial Identities and What They Mean for Health, June 8, 2017).

17 L. Charmaraman, M. Woo, A. Quach, and S. Erkut, "How Have Researchers Studied Multiracial Populations: A Content and Methodological Review of 20 Years of Research," *Cultural Diversity & Ethnic Minority Psychology* 20 (2014): 336.

18 john a. powell, *Racing to Justice: Transforming Our Conceptions of Self and Other to Build an Inclusive Society* (Indiana University Press, 2012).

19 M. Woo, S. B. Austin, D. R. Williams, and G. G. Bennett, "Reconceptualizing the Measurement of Multiracial Status for Health Research in the United States," *Du Bois Review* 8 (2011): 25-36.

20 Jenifer L. Bratter and Bridget K. Gorman, "Does Multiracial Matter? A Study of Racial Disparities in Self-Rated Health," *Demography* 48 (2011): 127-152.

21 For a complete list of titles and contributors to the six research reviews on disaggregating data about racial groups, commissioned by RWJF in conjunction with this project, see Appendix A.

22 Wendy D. Roth, "The Multiple Dimensions of Race," *Ethnic and Racial Studies Review* 39 (January 6, 2016): 1310-1338.

23 Woo et al., "Reconceptualizing the Measurement of Multiracial Status for Health Research in the United States."

24 U. E. Bauer and M. Plescia, "Addressing Disparities in the Health of American Indian and Alaska Native People: The Importance of Public Health Data," *American Journal of Public Health* (2014): S255-S257.

25 K. A. Earle and A. Cross, *Child Abuse and Neglect among American Indian/Alaska Native Children: An Analysis of Existing Data* (Casey Family Program, 2001). http://muskie.usm.maine.edu/helpkids/rcpdfs/B060041.pdf (accessed December 21, 2017).

26 M. A. Jim, E. Arias, D. S. Seneca, M. J. Hoopes, C. C. Jim, N. J. Johnson, and C. L. Wiggins, "Racial Misclassification of American Indians and Alaska Natives by Indian Health Service Contract Health Service Delivery Area," *American Journal of Public Health* 104 (2014): S295-S302.

27 Nina B. Wallerstein and Bonnie Duran, "Using Community-Based Participatory Research to Address Health Disparities," *Health Promotion Practice* 7 (2006): 312-323.

28 National Congress of American Indians, "NCAI Congratulates Tribal Nations in Virginia on Federal Recognition," News, January 30, 2018 http://www.ncai.org/news/articles/2018/01/30/ncai-congratulates-tribal-nations-in-virginia-on-federal-recognition.

29 National Congress of American Indians, "Havasupai Tribe and the Lawsuit Settlement Aftermath," American Indian & Alaska Native Genetics Resource Center, n.d. http://genetics.ncai.org/case-study/havasupai-Tribe.cfm (accessed December 21, 2017).

30 C. Matthew Snip offers three features of "decolonized indigenous data": (1) the sole power of tribes to define and confer membership; (2) the data reflects indigenous interests, values, and priorities; and (3) tribal nations decide who has access to their data.

31 Sarah Gehrig, Walker Bosch, June Heineman, Nicole MartinRogers. Homelessness and Near-Homelessness on Six Minnesota American Indian Reservations: Findings from the 2015 study. Wilder Research. March 2017. http://mnhomeless.org/minnesota-homeless-study/reports-and-fact-sheets/2015/2015-homeless-reservations-3-17.pdf

32 Indigenous Data Sovereignty Network. Native Nations Institute at the University of Arizona. 2018. http://usindigenousdata.arizona.edu/about-us-0.

33 K. S. Markides and J. Coreil, "The Health of Hispanics in the Southwestern United States: An Epidemiological Paradox," *Public Health Reports* 101 (1986): 253.

34 N. S. Ali, "Providing Culturally Sensitive Care to Egyptians with Cancer," *Cancer Practice* 4 (1996): 212-215.

35 Carmela Alcántara, Leopoldo J. Cabassa, Shakira Suglia, Irene Perez Ibarra, A. Louise Falzon, Elliot McCullough, and Talhah Alvi, *Disaggregating Latina/o Surveillance Health Data across the Lifecourse: Barriers, Facilitators, and Exemplars* (Robert Wood Johnson Foundation, 2017).

36 M. Alegría, G. Canino, F. S. Stinson, and B. F. Grant, "Nativity and DSM-IV Psychiatric Disorders among Puerto Ricans, Cuban Americans, and non-Latino Whites in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions," *Journal of Clinical Psychiatry* 67 (2016): 56-65.

37 B. F. Grant, F. S. Stinson, D. S. Hasin, D. A. Dawson, S. P. Chou, and K. Anderson, "Immigration and Lifetime Prevalence of DSM-IV Psychiatric Disorders Among Mexican Americans and Non-Hispanic Whites in the United States: Results from the National Epidemiologic Survey on Alcohol and Related Conditions," *Archives of General Psychiatry* 61 (2004): 1226-1233.

38 Alegría et al., "Nativity and DSM-IV psychiatric disorders among Puerto Ricans, Cuban Americans, and non-Latino Whites in the United States: results from the National Epidemiologic Survey on Alcohol and Related Conditions."

39 PolicyLink and USC PERE, "Poverty. United States. Percent People Below Poverty by Race/Ethnicity and Ancestry: United States, Latino, U.S.-born, 100%, 2015," National Equity Atlas, http://nationalequityatlas.org/indicators/Poverty/By_ancestry:40256/United_States/false/Race~ethnicity:Latino/Nativity:U.S.-born/Poverty_Level:100/Year(s):2015/ (accessed February 14, 2018.)

40 Carmela Alcántara, Leopoldo J. Cabassa, Shakira Suglia, Irene Perez Ibarra, A. Louise Falzon, Elliot McCullough, and Talhah Alvi, *Disaggregating Latina/o Surveillance Health Data across the Lifecourse: Barriers, Facilitators, and Exemplars* (Robert Wood Johnson Foundation, 2017).

41 H. Jamil, M. Farrag, J. Hakim-Larson, T. Kafaji, H. Abdulkhaleq, and A. Hammad, "Mental Health Symptoms in Iraqi Refugees: Posttraumatic Stress Disorder, Anxiety, and Depression," *Journal of Cultural Diversity* 14 (2007) 19-25.

42 U.S. Census Bureau, "Commuting, Median Rents and Language Other Than English Use in the Home on the Rise," *Census Bureau News*, CB17-204, December 7, 2017.

43 Jamil et al., "Mental Health Symptoms in Iraqi Refugees: Posttraumatic Stress Disorder, Anxiety, and Depression."

44 In health research, allostatic load is an index measuring the physiological consequences of repeated or chronic stress. It can be measured through biomeasures, including systolic and diastolic blood pressure, body mass index, glycated hemoglobin, albumin, creatinine clearance, triglycerides, C-reactive protein, homocysteine, and total cholesterol. For more information, see the Latina/o research review by Alcántara et al., *Disaggregating Latina/o Surveillance Health Data across the Lifecourse: Barriers, Facilitators, and Exemplars.*

45 In demography, "immigration status" means whether a person is an immigrant (i.e., foreign born), their country of origin and/or birth, and when that person immigrated to the U.S. On the other hand, "legal status" means whether a person is a legal permanent resident or asylee; holds a student or business visa; has Temporary Protected Status or Deferred Action for Childhood Arrival (DACA); or has some other status, such as living in the U.S. without proper documentation. For more information, see The Leadership Conference Education Fund's Factsheet "Citizenship and Legal Status Questions on the 2020 Census | Preventing a Decennial Disaster."

46 "Eligibility for Non-Citizens in Medicaid and CHIP," Centers for Medicare & Medicaid Services, November 2014.

47 "Citizenship and Immigration," California Health Interview Survey 2016 Adult Questionnaire, March 3, 2017, p. 56.

48 Migration Policy Institute, "Migration Data Hub," Migration Policy Institute, n.d. https://www.migrationpolicy.org/programs/migration-data-hub (accessed December 21, 2017). Their analysis of the SIPP data is from James Bachmeier at Temple University and Jennifer Van Hook at The Pennsylvania State University.

49 Isabel Wilkerson, *The Warmth of Other Suns: The Epic Story of American's Great Migration* (Knopf Doubleday Publishing Group, October 4, 2011).

50 Ayumi Takenaka, "Second Migration: Who Re-Migrates and Why These Migrants Matter," Migration Policy Institute, April 26, 2007. https://www.migrationpolicy.org/article/secondary-migration-who-re-migrates-and-why-these-migrants-matter.

51 Jim Webb, *Born Fighting: How the Scots-Irish Shaped America* (Broadway Books, October 11, 2005).

52 Urban Indian Health Commission, *Invisible Tribes: Urban Indians and Their Health in a Changing World*.

53 Ellis P. Monk Jr., "The Cost of Color: Skin Color, Discrimination, and Health among African-Americans," *American Journal of Sociology* 121 (2015): 396-444."

54 James S. Jackson, Tod G. Hamilton, Mosi A. Ifatunji, Krim K. Lacey, Hedwig E. Lee, and Jane A. Rafferty, *Using Analytic Domains Within the Black Population to Understand Disparities in Population Health* (Robert Wood Johnson Foundation, 2017).

55 Trina Jones, "The Significance of Skin Color in Asian and Asian-American Communities: Initial Reflections," *University of California-Irvine Law Review* 3 (2013): 1105.

56 Trina Jones, "Shades of Brown: The Law of Skin Color," *Duke Law Journal* 49 (2000): 1487.

57 Ellis P. Monk Jr., "The Cost of Color: Skin Color, Discrimination, and Health among African-Americans."

58 Ibid.

59 Corey L. Keyes, "The Black–White Paradox in Health: Flourishing in the Face of Social Inequality and Discrimination," *Journal of Personality* 77 (2009): 1677-1706.

60 David R. Williams and Tara R. Earl, "Commentary: Race and Mental Health—More Questions than Answers," *International Journal of Epidemiology* 36 (2007): 758-760.

61 Ellis P. Monk Jr., "The Cost of Color: Skin Color, Discrimination, and Health among African-Americans."

62 David R. Williams, Naomi Priest, and Norman Anderson. "Understanding Associations between Race, Socioeconomic Status and Health: Patterns and Prospects," *Health Psychology* 35.4 (2016): 407–411.

63 Jeff Krehely, "How to Close the LGBT Health Disparities Gap: Disparities by Race and Ethnicity," Center for American Progress, 2009. https://cdn.americanprogress.org/wp-content/uploads/issues/2009/12/pdf/lgbt_health_disparities_race.pdf (accessed January 25, 2018). Gary J. Gates, "Demographics and LGBT Health," *Journal of Health and Social Behavior* 54 (2013): 72-74.

64 The GenIUSS Group. *Best Practices for Asking Questions to Identify Transgender and Other Gender Minority Respondents on Population-Based Surveys.* J.L. Herman (Ed.). (Los Angeles, CA: The Williams Institute, 2014).

65 Jody L. Herman, Bianca D.M. Wilson, and Tara Becker, *Demographic and Health Characteristics of Transgender Adults in California: Findings from the 2015-2016 California Health Interview Survey* (Williams Institute and UCLA Center for Health Policy Research, 2017).

66 Linda Villarosa, "America's Hidden H.I.V. Epidemic," *The New York Times Magazine*, June 2017. https://www.nytimes.com/2017/06/06/magazine/americas-hidden-hiv-epidemic.html (accessed February 2, 2018).

67 Alcántara et al., *Disaggregating Latina/o Surveillance Health Data across the Lifecourse: Barriers, Facilitators, and Exemplars*.

68 Comments of Kathy Ko Chin, President and CEO of Asian Pacific Islander American Health Forum, at convening for this project, Washington, DC. June 29, 2017.

69 "The World's Most Valuable Resource Is No Longer Oil, but Data," *The Economist*, May 6, 2017.

70 Travis B. Murdoch and Allan S. Detsky, "The Inevitable Application of Big Data to Health Care," *Journal of the American Medical Association* 309 (2013): 1351-1352.

71 The racial categories included in the census questionnaire generally reflect a social definition of race recognized in this country and not an attempt to define race biologically, anthropologically, or genetically. In addition, it is recognized that the categories of the race item include racial and national origin or sociocultural groups. People may choose to report more than one race to indicate their racial mixture, such as "American Indian" and "White." People who identify their origin as Hispanic, Latino, or Spanish may be of any race. For more information, see U.S. Census Bureau, "Race: About," https://www.census.gov/topics/population/race/about.html (accessed January 25, 2018).

72 Office of Management and Budget, "Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity," *Federal Register*, 62 (1997): 58782-58790. https://www.whitehouse.gov/omb/fedreg_1997standards (accessed December 21, 2017).

73 U.S. Census Bureau, "2010 Census Summary File – Technical Documentation," Code Lists, 2010. https://www.census.gov/prod/cen2010/doc/sf1.pdf (accessed December 21, 2017).

74 Office of Management and Budget, "Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity," *Federal Register*, September 30, 2016. https://www.gpo.gov/fdsys/pkg/FR-2016-09-30/pdf/2016-23672.pdf.

75 Arab American Institute, "Adding a MENA Category to the U.S. Census," Arab American Institute, May 16, 2015. http://www.aaiusa.org/2020census.

76 The U.S. Census Bureau defines persons of Arab ethnicity as persons who trace their ancestry to one of the following 17 countries: Algeria, Bahrain, Egypt, Iraq, Jordan, Kuwait, Lebanon, Libya, Morocco, Oman, Qatar, Saudi Arabia, Syria, Sudan, Tunisia, United Arab Emirates, and Yemen. The International Organization for Migration defines the Middle East and North Africa (MENA) region using the same 17 countries. For more information, see the Census Bureau's Forum on Ethnic Groups from the Middle East and North Africa.

77 E. Hoeffel, P. Rastogi, M. O. Kim, and H. Shahid, "The Asian Population: 2010," *U.S. Census Bureau* C2010BR-11 (March 2012): 1-24.

78 U.S. Census Bureau, "Working with OMB and Federal Statistical Agencies," Research to Improve Data on Race and Ethnicity, 2017. https://www.census.gov/about/our-research/race-ethnicity.html (accessed December 21, 2017).

79 Jason Jordan, "Inadequate Funding Imperils a Fair and Accurate 2020 Census," American Planning Association, 2017. https://www.planning.org/blog/blogpost/9125999/ (accessed December 21, 2017).

80 Danny Vinik, "Is Washington Bungling the Census?" *Politico*, October 11, 2017.

81 "High Risk List," U.S. Government Accountability Office, 2017. https://www.gao.gov/highrisk/overview (accessed January 28, 2018).

82 Letter from the Leadership Conference on Civil and Human Rights and 80 undersigned organizations to Jennifer Park, Office of the U.S. Chief Statistician, OMB, April 28, 2017, http://www.civilrightsdocs.info/pdf/policy/letters/2017/Leadership%20Conference%20OMB%20FRN2-RE%20Standards-Comments-04%2028%202017%20final.pdf.

83 The role of the National Center for Health Statistics (NCHS) is not only to identify health-related data, but also to compile statistical information to inform public health decisions. NCHS is a key player in documenting the U.S. population's health status, as well as the health of selected subgroups, by race and ethnicity, socioeconomic status, geographic region, and other population characteristics.

84 National Center for Health Statistics, "Native Hawaiian and Pacific Islander (NHPI) National Health Interview Survey (NHIS)," Centers for Disease Control and Prevention, 2017. https://www.cdc.gov/nchs/nhis/nhpi.html (accessed December 21, 2017).

85 Office of Minority Health, "OMH & CDC Collaborating to Improve Understanding of the Health Status of American Indians and Alaska Natives," U.S. Department of Health and Human Services, 2017. https://minorityhealth.hhs.gov/omh/browse.aspx?lvl=2&lvlid=89 (accessed December 21, 2017).

86 Mikelyn Meyers, "Respondent Confidentiality Concerns and Possible Effects on Response Rates and Data Quality for the 2020 Census," National Advisory Committee on Racial, Ethnic, and Other Populations Fall Meeting, November 2, 2017. https://www2.census.gov/cac/nac/meetings/2017-11/Meyers-NAC-Confidentiality-Presentation.pdf.

87 Annie Lowrey, "Trump's Anti-Immigrant Policies Are Scaring Eligible Families Away from the Safety Net," *The Atlantic*, March 24, 2017.

88 John T. Cacioppo and Louise C. Hawkley, "Social Isolation and Health, with an Emphasis on Underlying Mechanisms," *Perspectives in Biology and Medicine* 46 (Summer 2003): S39-S52.

89 Meyers, "Respondent Confidentiality Concerns and Possible Effects on Response Rates and Data Quality for the 2020 Census."

90 Steven A. Holmes, "Report Says Census Bureau Helped Relocate Japanese," *The New York Times*, March 17, 2000, https://nyti.ms/2l4HARe. The Second War Powers Act of 1942 temporary repealed protections and allowed this transfer of data to be legal, and to date, it is the first and only time known where the Bureau shared information for just one group (names, addresses, age, sex, citizenship, and country of birth).

91 Ben Kamisar, "Haley: No Muslim Registry under Trump Administration," *The Hill*, January 18, 2017.

92 The Leadership Conference Education Fund, *Factsheet: Citizenship and Legal Status Questions on the 2020 Census | Preventing a Decennial Disaster* (The Leadership Conference Education Fund, August 28, 2017).

93 Arab American Institute, "Adding a MENA Category to the U.S. Census."

94 Tara Bahrampour, "A U.S. Census Proposal to Add Category for People of Middle Eastern Descent Makes Some Uneasy," *The Washington Post*, October 21, 2016.

95 Matt Apuzzo and Joseph Goldstein, "New York Drops Unit That Spied on Muslims," *The New York Times,* April 15, 2014: A1.

96 Lynette Clemetson, "Homeland Security Given Data on Arab-Americans," *The New York Times*, July 30, 2004. http://www.nytimes.com/2004/07/30/us/homeland-security-given-data-on-arab-americans.html.

97 U.S. Census Bureau, "The '72-Year Rule,'" U.S. Census Bureau History, n.d. https://www.census.gov/history/www/genealogy/decennial_census_records/the_72_year_rule_1.html (accessed December 21, 2017).

98 PolicyLink and USC PERE, "Race/ethnicity in the United States – Map by State, 2014," National Equity Atlas, 2017. http://nationalequityatlas.org/indicators/Race~ethnicity/Map:41391/United_States/false/geotype:state/value1:2014/value2:7/ (accessed January 25, 2018).

99 Pew Research Center, "Top 10 U.S. Metropolitan Areas by Cambodian Population, 2015. Cambodians in the U.S. Fact Sheet," Pew Research Center, 2017. http://www.pewsocialtrends.org/fact-sheet/asian-americans-cambodians-in-the-u-s/ (accessed December 21, 2017).

100 Elizabeth Dunbar, "Survey: Nearly 1 in 3 US Somalis Live in Minnesota," *Minnesota Public Radio*, December 14, 2010.

101 Center for American Progress/AAPI Data, *Who Are Hmong Americans?* (Center for American Progress Factsheet, 2015).

102 Arab American Institute Foundation, "National Demographic Profile," Arab American Institute Foundation, 2014. http://www.aaiusa.org/demographics (accessed December 21, 2017).

103 State of California, "Accounting for Health and Education in API Demographics (AHEAD) Act," AB-1726, 2016. http://leginfo.legislature.ca.gov/faces/billNavClient.xhtml?bill_id=201520160AB1726 (accessed December 21, 2017).

104 The full list of groups covered under the AHEAD Act is Bangladeshi, Hmong, Indonesian, Malaysian, Pakistani, Sri Lankan, Taiwanese, Thai, Fijian, and Tongan.

105 State of Rhode Island, "All Students Count Act," RI S0439, 2017. https://www.billtrack50.com/BillDetail/845732 (accessed December 21, 2017).

106 EdAllies, "Ensure that All Minnesota Students Count," EdAllies, 2017. https://edalliesmn.org/issues/data-disaggregation/ (accessed December 21, 2017).

107 Kristoffer Tigue, "Why a New Education Data Law Is So Important to Minnesota's Asian and Pacific Islander Community," *MinnPost*, August 5, 2016.

108 Mustafa Aswad and Adnan Hammad, *Health Survey of the Arab, Muslim, and Chaldean American Communities in Michigan* (Michigan Department of Community Health and Arab Community Center for Economic and Social Services, 2001).

109 Summer Sherburne Hawkins, Brunilda Torres, Georgia Simpson, and Bruce B. Cohen, "Setting the Standards for Collecting Ethnicity Data in the Commonwealth of Massachusetts," *Journal of Public Health Management and Practice* 17 (2011): 550-553.

110 Asian & Pacific Islander American Health Forum, *Asian American, Native Hawaiian, & Pacific Islander Maternal Health Disparities* (APIAHF Issue Brief, May 2012).

111 Cynthia Lee, "Students: 'Count Me In!,'" *UCLA Newsroom*, December 11, 2007.

112 Martin Slagter, "Middle East, North African Students Want University of Michigan Recognition," *MLive*, December 11, 2017.

113 Southeast Asia Resource Action Center, "The AHEAD Act – AB 1725," SEARAC, 2016. http://www.searac.org/content/ahead-act-ab-1726 (accessed December 21, 2017).

114 G. Cristina Mora, *Making Hispanics: How Activists, Bureaucrats, and Media Constructed a New American* (University of Chicago Press, 2014).

Appendix A:
Research Reviews of Disaggregation of Racial and Ethnic Data for Health

Six research reviews of the state of data disaggregation for health research and policy provided a solid foundation of knowledge for this project and the start of a network of leading researchers. These reviews, completed in 2016 and 2017, were thorough accounts of the conceptual frameworks, empirical research, databases—and policies about those databases—that pertained to each population. There was a research review for each of the five major population groups in the U.S. (Black and African American; Latinx; American Indian and Alaska Native; Asian American, Native Hawaiian, and Pacific Islander; and Non-Hispanic White) and a research review of a seven-country comparative study (covering Aotearoa New Zealand, Canada, Denmark, Great Britain, Hungary, Malaysia, and the Plurinational State of Bolivia, as well as the European Union). The six reviews were commissioned by the Robert Wood Johnson Foundation under separate grants to the universities or home organizations of the authors. All of the contributors are listed below, and where they have moved since the report was completed, their new affiliation is also shown.

Using Analytic Domains Within the Black Population to Understand Disparities in Population Health (May 2, 2018).

- **James S. Jackson, University of Michigan**
  Director of the Institute for Social Research
  Daniel Katz Distinguished University Professor of Psychology

- **Tod G. Hamilton, Princeton University**
  Assistant Professor of Sociology
  Faculty Associate of the Office of Population Research

- **Mosi A. Ifatunji, University of North Carolina, Chapel Hill**
  Assistant Professor of Sociology
  Carolina Population Center

- **Krim K. Lacey, University of Michigan-Dearborn**
  Assistant Professor of Sociology and African and African American Studies

- **Hedwig E. Lee, University of Washington**
  Associate Professor of Sociology

- **Jane A. Rafferty, University of Michigan**
  Research Area Specialist Senior, Institute for Social Research

Disaggregating Latina/o Surveillance Health Data across the Lifecourse: Barriers, Facilitators, and Exemplars (April 27, 2017).

- **Carmela Alcántara, Columbia University**
  Assistant Professor of Social Work
  Faculty Affiliate, Social Intervention Group

- **Leopoldo Cabassa, Washington University in St. Louis**
  Associate Professor of Social Work

- **Shakira Suglia, Emory University**
  Associate Professor of Epidemiology

- **Irene Perez Ibarra, Columbia University**
  Associate Research Scientist, School of Social Work

- **A. Louise Falzon, Columbia University**
  Information Specialist, Center for Behavioral and Cardiovascular Health

- **Elliot McCullough, Columbia University**
  Joint MSW and MPH Student, School of Social Work and Mailman School of Public Health

- **Talhah Alvi, Columbia University**
  MS Student, Teacher's College

Disaggregating American Indian & Alaska Native Data: A review of literature (July 2016).

- **Malia Villegas**
  Former Director, National Congress of American Indians Policy Research Center

- **Amber Ebarb**
  Program Manager, NCAI Policy Research Center

- **Sarah Pytalski**
  Former Policy Research & Evaluation Manager, NCAI Policy Research Center

- **Yvette Roubideaux**
  Director, NCAI Policy Research Center

Understanding the Culture of Health for Asian American, Native Hawaiian and Pacific Islanders (AANHPIs): What do population-based health surveys across the nation tell us about the state of data disaggregation for AANHPIs? (September 2017).

- **Ninez Ponce, University of California, Los Angeles**
  Professor of Public Health
  Principal Investigator, California Health Interview Survey

- **AJ Scheitler, University of California, Los Angeles**
  Coordinator of Stakeholder Relations, Center for Health Policy Research
  Coordinator, National Network of State and Local Health Surveys

- **Riti Shimkhada, University of California, Los Angeles**
  Public Administration Analyst, Center for Health Policy Research

Challenges and Prospects for Disaggregating Health Data among Non-Hispanic Whites (March 30, 2017).

- **Jen'nan Read, Duke University**
  Associate Professor of Sociology and Global Health

Heterogeneity/Granularity in Ethnicity Classifications outside the United States (December 10, 2016).

- **Raj Bhopal, University of Edinburgh**
  Principal Investigator
  Bruce and John Usher Professor of Public Health and Honorary Consultant in Public Health Medicine

- **Emma Davidson, University of Edinburgh**
  Co-Principal Investigator
  Research Fellow

- **Nazmy Villarroel-Williams, University of Edinburgh**
  Research Fellow

- **Pamela Perayra-Zamora, University of Alicante**
  Assistant Professor of Community Health and History of Science

- **Allan Krasnik, University of Copenhagen**
  Professor of Public Health
  Investigator, Denmark

- **Dr. Peter Aspinall**
  Investigator, United Kingdom

- **Dr. Liv Stubbe**
  Investigator, Denmark

- **Dr. Inez Koller**
  Investigator, Hungary

- **Dr. Donna Cormack**
  Investigator, New Zealand

- **Dr. Tahu Kukutai**
  Investigator, New Zealand

- **Professor Hude Quan**
  Investigator, Canada

- **Dr. Kelsey Lucyk**
  Investigator, Canada

- **Dr. Karen Tang**
  Investigator, Canada

- **Dr. Shyamala Nagaraj**
  Investigator, Malaysia

- **Dr. Chiu Wan Ng**
  Investigator, Malaysia

Appendix B:
Convenings With Roster of Attendees

The series of three convenings explored different components of research, health practice, and policy changes pertaining to data disaggregation:

Making the Case for Data Disaggregation to Advance a Culture of Health: Understanding the Health of Racially and Ethnically Diverse Populations
Los Angeles, CA, May 24 & 25, 2017

Multiple Racial Identities and What They Mean for Health: Clarifying Issues and Raising the Visibility of Populations
Atlanta, GA, June 8 & 9, 2017

The Role of Data Disaggregation in Understanding the Health of Immigrants and Migrants of Diverse Backgrounds
Washington, DC, June 28 & 29, 2017

The participants in these convenings, presented in alphabetical order, not including staff of PolicyLink or the Robert Wood Johnson Foundation who also attended:

- **Dolores Acevedo-Garcia, Brandeis University**
  Professor of Human Development and Social Policy
  Director, Institute for Child, Youth, and Family Policy

- **Carmela Alcántara, Columbia University**
  Assistant Professor of Social Work
  Faculty Affiliate, Social Intervention Group

- **Tara Becker, UCLA Center for Health Policy Research**
  Statistician

- **Maya Berry, Arab American Institute**
  Executive Director

- **Dulce Bustamante-Zamora, California Department of Public Health**
  Office of Health Equity, Research Scientist II

- **Leopoldo Cabassa, Columbia University**
  Associate Professor of Social Work
  Assistant Director of the New York State Center for Excellence for Cultural Competence

- **Mary Campbell, Texas A&M University**
  Associate Professor of Sociology

- **Randy Capps, Migration Policy Institute**
  Director of Research for U.S. Programs

- **Richard Chang, Empowering Pacific Islander Communities**
  Director of Policy

- **Richard Chase, Wilder Research**
  Senior Research Manager

- **Kathy Ko Chin, Asian & Pacific Islander American Health Forum**
  President and CEO

- **Marianne Chung, Asian & Pacific Islander American Health Forum**
  Senior Director of Community Initiatives

- **Ben D'Avanzo, Asian & Pacific Islander American Health Forum**
  Senior Policy Analyst

- **Adrian Dominguez, Urban Indian Health Institute**
  Scientific Director

- **Rashida Dorsey, HHS Office of the Assistant Secretary for Planning and Evaluation**
  Director, Division of Data Policy
  Senior Advisor on Minority Health and Health Disparities

- **Amber Ebarb, National Congress of American Indians**
  Budget/Policy Analyst, Policy Research Center

- **Abigail Echo-Hawk, Urban Indian Health Institute**
  Director

- **Sameera Fazili, Federal Reserve Bank of Atlanta**
  Senior Community and Economic Development Visiting Adviser

- **Christy Gamble, Black Women's Health Imperative**
  Director, Health Policy & Legislative Affairs

- **Jeffrey Hall, Centers for Disease Control and Prevention**
  Deputy Associate Director for Science, Office of Minority Health & Health Equity

- **Tod Hamilton, Princeton University**
  Assistant Professor of Sociology

- **Sonia Harb, University of Michigan School of Social Work**
  Director, Technical Assistance Center

- **Mosi Ifatunji, University of North Carolina - Chapel Hill**
  Assistant Professor of Sociology

- **Hassan Jaber, Arab Community Center for Economic & Social Services**
  Executive Director and CEO

- **Iyanrick John, Asian Pacific Islander American Health Forum**
  Policy Director

- **Nicholas A. Jones, US Bureau of the Census**
  Director and Senior Advisor of Race and Ethnic Research and Outreach, Population Division

- **Trina Jones, Duke University School of Law**
  Professor

- **Richard Kang, UCLA Medical Group, Dignity Health Northridge Hospital Medical Center**
  Medical Director, PICU
  Multiracial Americans of Southern California

- **Krim Lacey, University of Michigan**
  Research Fellow

- **Vincent Lafronza, National Network of Public Health Institutes**
  President and CEO

- **Thomas Lopez, Multiracial Americans of Southern California**
  President

- **Mona Makki, ACCESS (Arab Community Center for Economic and Social Services)**
  Director, ACCESS Community Health and Research Center

- **Luis Maldonado, University of North Carolina – Chapel Hill**
  PhD Student, Nutritional Epidemiology

- **Rachel Marks, US Bureau of the Census**
  Senior Analyst, Ethnicity and Ancestry Branch, Population Division

- **Nicole MartinRogers, Wilder Research**
  Senior Research Manager

- **Paul Mattessich, Wilder Research**
  Executive Director

- **Meghan Maury, National LGBTQ Task Force**
  Policy Director

- **Faith Mitchell, Grantmakers In Health**
  President and CEO

- **Ellis Monk, Princeton University**
  Assistant Professor of Sociology

- **David Norris, The Kirwan Institute for the Study of Race and Ethnicity at Ohio State University**
  Senior Researcher

- **Leslie Parks, Black Women's Health Imperative**
  Senior Director, Programs & Advocacy

- **Ana Penman-Aguilar, Centers for Disease Control and Prevention**
  Associate Director for Science at the Office of Minority Health and Health Equity

- **Erica Phillips, ESRI**
  Federal Health & Human Services

- **Bev Pratt, Office of Management and Budget**
  Statistical & Science Policy Analyst, Office of Information & Regulatory Affairs

- **Ninez Ponce, University of California, Los Angeles**
  Principal Investigator, California Health Interview Survey, Professor of Public Health

- **Sarah Pytalski, National Congress of American Indians**
  Policy Research & Evaluation Manager, Policy Research Center

- **Jane Rafferty, University of Michigan**
  Research Area Specialist

- **Jen'nan Read, Duke University**
  Associate Professor of Sociology and Global Health

- **Merarys Rios, U.S. Census Bureau**
  Chief of the Ethnicity and Ancestry Branch, Population Division

- **AJ Scheitler, University of California, Los Angeles**
  Coordinator of Stakeholder Relations, Center for Health Policy Research
  Coordinator, National Network of State and Local Health Surveys

- **Sonya Schwartz, National Immigration Law Center**
  Consultant

- **Justin Scoggins, USC Program for Environmental and Regional Equity (PERE)**
  Data Manager

- **Riti Shimkhada, University of California, Los Angeles**
  Public Administration Analyst, Center for Health Policy Research

- **Hyon B. Shin, U.S. Bureau of the Census**
  Chief of Racial Statistics Branch, Population Division

- **Audrey Singer, Urban Institute**
  Senior Fellow, Metropolitan Housing and Communities
  Policy Center

- **Shakira Suglia, Emory University School of Public Health**
  Associate Professor of Epidemiology

- **Michele Suina, Albuquerque Area Southwest Tribal Epidemiology Center**
  Program Director for the CDC Good Health and Wellness in
  Indian Country Program

- **Priya Thaker, UCLA Center for Health Policy Research**
  Public Administration Analyst, California Health Interview
  Survey

- **Yousef Turshani, UMMA Community Clinic**
  Chief Medical Officer

- **Alexander Vigo, HHS Office of Minority Health**
  Hispanic/Latino Health Policy Lead, Division of Policy and
  Data

- **Nazmy Villarroel-Williams, University of Limerick**
  Post-doctoral Researcher, Graduate Entry Medical School

- **Akil Vohra, White House Initiative on Asian Americans and Pacific Islanders**
  Director of Strategic Initiatives

- **Minh Wendt, HHS Office of Minority Health**
  Data Policy Lead, Division of Policy and Data

- **Jessica West, Duke University**
  PhD Student, Sociology

- **Shondelle Wilson-Frederick, Centers for Medicare and Medicaid Services**
  Public Health Advisor, Office of Minority Health

- **Meghan Woo, Abt Associates**
  Associate

- **Stella Yi, New York University School of Medicine**
  Assistant Professor of Health Equity

- **Corrine Yu, The Leadership Conference on Civil and Human Rights**
  Senior Counsel and Managing Policy Director

Appendix C:
U.S. Census Bureau, Timeline for Revisions to 2020 Census
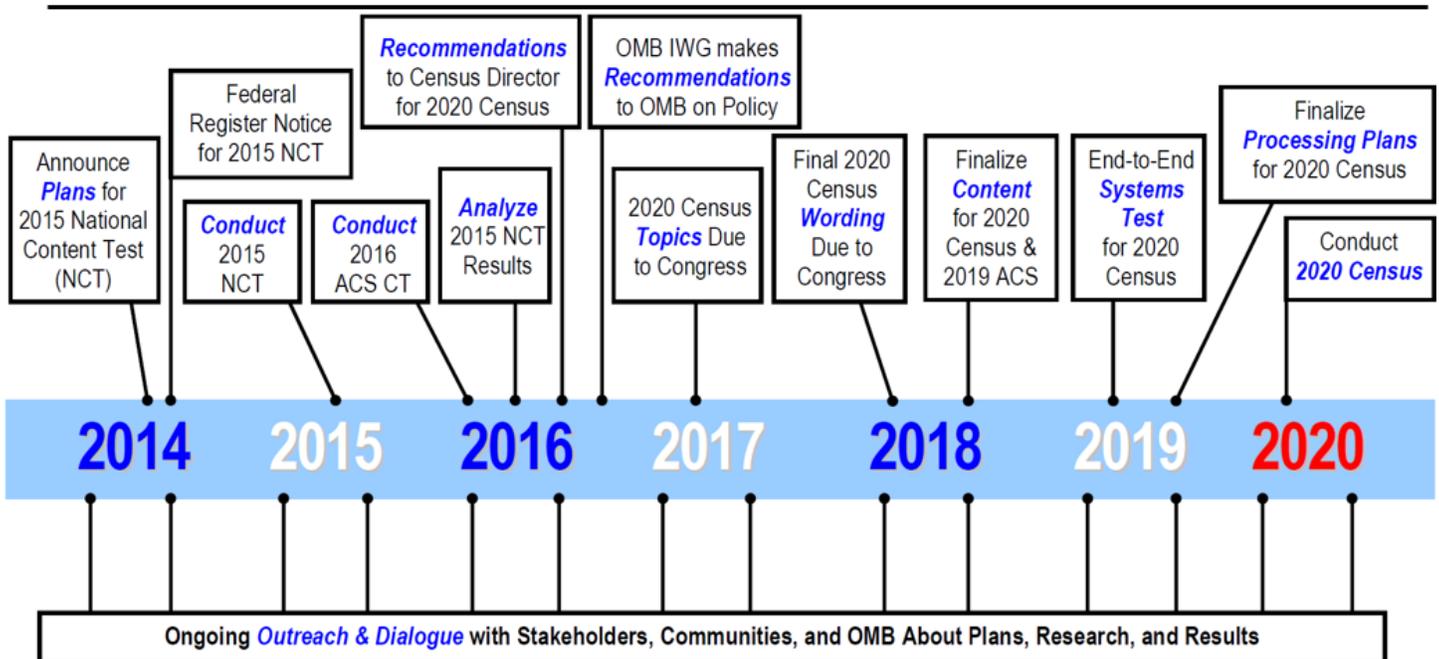
Appendix D:
Federal Interagency Working Group for Research on Race and Ethnicity, 2017 Interim Report, Excerpt

I. MANDATE AND SCOPE OF REVIEW

Under the *Paperwork Reduction Act (PRA)*, OMB is required to ensure the efficiency and effectiveness of the Federal Statistical System as well as the integrity, objectivity, impartiality, utility, and confidentiality of information collected for statistical purposes. OMB also develops and oversees the implementation of Federal-wide principles, policies, standards, and guidelines concerning the development, presentation, and dissemination of statistical information.

OMB's *Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity* (https://www.gpo.gov/fdsys/pkg/FR-1997-10-30/pdf/97-28653.pdf) were developed to provide consistent and comparable data on race and ethnicity throughout the Federal government for statistical and administrative programs. The standards were developed using

A. Background

Since the standards were last revised, the ways in which people in the U.S. self-identify their race and ethnicity may have evolved. To ensure that measures of race/ethnicity remain relevant for policy making purposes, in 2014 OMB formed an Interagency Working Group for Research on Race and Ethnicity (Working Group) to exchange research findings, identify implementation issues, and collaborate on a shared research agenda to improve Federal data on race and ethnicity. The Working Group comprises representatives from ten cabinet departments and three other agencies that collect or use race and ethnicity data.

Through its systematic review of the implementation of the 1997 revisions across Federal and state governments, and examining stakeholder feedback, the Working Group identified particular areas where revisions to the 1997 standards could potentially improve the quality of race and ethnicity information collected and presented by Federal agencies. On September 30, 2016, OMB requested public comment on the principles that should govern its work and the areas initially identified for review. (See https://www.regulations.gov/document?D=OMB-2016-0002-0001).

Specifically, the Working Group proposed further exploration in four areas. Subgroups were formed to prepare initial proposals for the Working Group's consideration.
1. The use of separate questions versus one combined question to measure race and ethnicity and question phrasing as a solution to race/ethnicity question nonresponse;
2. The classification of a Middle Eastern and North African group and distinct reporting category;
3. The description of the intended use of minimum reporting categories; and
4. The salience of terminology used for race and ethnicity classifications and other language in the standards.

Over the 30 day period, 3,750 comments were received from the public. After reviewing these public comments and continuing its analysis, the Working Group developed this Interim Report, which describes the Working Group's progress to date and requests further public comment. In this Interim Report, first, the principles guiding the Working Group's review (Section IIA) and the results of the September 30's Federal Register Notice (Section IIB) are summarized. Second, the review process, taking into account public comment, empirical analysis (Section IIC), statutory requirements and public burden (Section IID), is described. This is followed by an in-depth discussion of each area identified for review (Section III). The report concludes with initial proposals and requests for further public comment. Additional detail on the analyses is included in the attachments.

B. Summary of initial proposal

1. Separate Questions or Combined Question for Race and Ethnicity

The current standard of Separate Questions calls for **collection** of information with a question first asking ethnicity (Hispanic or non-Hispanic) and then in a separate question asking race. Self-reported information collected in Federal surveys more often follows this standard than non-Federal administrative data collections. However, Federal statistics often are not **reported** using the separate categories. The two major barriers to fully implementing the standards for collection seem be: (1) how agencies calculate race when there is no response to race or "other" is selected, and (2) the limitation of non-Federal administrative data systems in accommodating a "mark all that apply" approach that allows respondents to select multiple races. Barriers on the reporting side seem to be the size of the sampled population and the associated risk to respondent confidentiality and statistical reliability when the size of the race group is small in the sample. Federal agencies have tried to achieve the spirit of the current standards despite these barriers, but challenges remain. However, the use of a Combined Question format in certain circumstances, where race and ethnicity are not asked separately, may address nonresponse challenges if sufficient data quality and comparability can be assured across the Federal government. The subgroup will explore these approaches further before preparing a proposal for the Working Group's consideration.

2. Classification of Middle Eastern or North African (MENA) Race/Ethnicity

Since the Federal standards on race and ethnicity first were issued in 1977, the Federal government has classified White as people having origins in any of the original peoples of Europe, the Middle East, or North Africa. As part of periodic review of the standards during the mid-1990s, OMB considered several suggestions to improve the accuracy and reliability of Federal race and ethnicity statistics. Among the suggestions proposed was the inclusion of an additional, distinct minimum reporting category for respondents identifying as "Arabs or Middle Easterners." At the conclusion of that review, agreement could not be reached among public stakeholders on the intended measurement concept (i.e., whether the category should be based on language, geography, etc.) nor on a definition for this category. As a result, an additional, minimum reporting category for this group was not created. Instead, the *1997 Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity* advised that further research be done to determine the best way to improve data for "Arabs/Middle Easterners."

Based on public comment and analyses to date (see the Census Bureau's 2015 Forum on Ethnic Groups from the Middle East and North Africa and a review public comments on *Proposed Information Collection; Comment Request; 2015 National Content Test* (12/2/2014)), the Subgroup proposes that a Middle Eastern or North African classification be added to the standards and be geographically based. The MENA classification should be defined as: A person having origins in any of the original peoples of the Middle East and North Africa. This includes, for example, Lebanese, Iranian, Egyptian, Syrian, Moroccan, Israeli, Iraqi, Algerian, and Kurdish[1]. However, many questions remain that require additional input from the public before the Working Group as a whole would adopt this proposal and recommend a change in the standard to OMB.

The current standards already allow for the breakout of *racial* subcategories from the required minimum standards when the quality of the collected data can be assured. However, many public commenters signaled that an *ethnic* classification for the MENA population should be added because those with roots, origins, and heritage from the Middle East and North Africa are racially diverse. Although results from the Census 2015 National Content Test show that those in the Census Bureau's working classification of MENA largely identify with the MENA response category when it is available, the 2015 test did not break out MENA as an *ethnicity*. Rather, it was presented as a checkbox where respondents could check multiple boxes, such as MENA and White or MENA and Black or just MENA, all with the ability to also write in a country of origin or ethnicity.

In addition, some of the groups proposed for inclusion under a MENA classification were also ethnoreligious groups. A challenge to ethnicity measurement can be the intersection of ethnicity with religious affiliation. The race/ethnicity standards are not intended to measure religion (see P.L. 94-521), and it is unclear how to address inclusion of ethnoreligious groups while clearly maintaining the intent and use of the resulting measure as not indicating religion. Further, the cost and burden of requiring this additional reporting category when race/ethnicity is measured across the Federal government is unclear. Also, the size of the overall group is small as a percentage of the total population, and reporting out the data could encounter many of the problems encountered with other small population groups such as Native Hawaiians and Pacific Islanders in surveys or data collections where there are small samples. These cost and quality issues need to be explored further.

---

[1] The rationale for using these examples is to include the two largest Middle Eastern Arab nationalities (Lebanese and Syrian), the two largest North African Arab nationalities (Egyptian and Moroccan), and the two largest non-Arab nationalities within the Middle Eastern / North African region (Iranian and Israeli) as the first six examples. This is followed by the next largest Middle Eastern Arab nationality (Iraqi), the next largest North African Arab nationality (Algerian); as well as an example of a transnational, non-Arab group (Kurdish).

3. Additional Minimum Reporting Categories

During OMB's review of Federal standards on collecting and reporting data on race/ethnicity in the 1990s, approximately 10,000 letters and postcards were received on the issue of classifying data on Native Hawaiians. The Interagency Committee recommended to OMB that data on Native Hawaiians continue to be classified in the Asian or Pacific Islander category. This recommendation was opposed by the Hawaiian congressional delegation, the 7,000 individuals who signed and sent preprinted yellow postcards, the State of Hawaii departments and legislature, Hawaiian organizations, and other individuals who commented on this recommendation. Based largely on evidence presented at public hearings in Hawaii, OMB did not accept the Interagency Committee recommendation and decided to break apart the Asian or Pacific Islander category into two categories—one called "Asian" and the other called "Native Hawaiian or Other Pacific Islander."

The initial review of the 1997 standards did not identify additional, minimum reporting categories for detailed race/ethnicity groups as an element for evaluation. However, during the public comment period for September 30, 2016's Federal Register Notice, the Working Group received more than 1,200 comments expressing the need for further disaggregated data for Asian communities and Native Hawaiian or Other Pacific Islander communities. Other comments express a similar need for disaggregated data, including 10 comments advocating for the disaggregation of the Black or African American category.

There are numerous examples of Federal agencies collecting detailed race and ethnicity data in their statistical reporting; these are not limited to decennial censuses or the American Community Survey (ACS). Nonetheless, OMB has learned that the minimum reporting categories as described in the current standards are often misinterpreted as the only permissible reporting categories rather than allowing more detailed categories, provided that detail group data can be aggregated in a systematic way to allow for comparison across data sources.

The Subgroup proposes that rather than changing the standards, OMB should issue guidelines for the collection of detailed data for American Indian or Alaska Native, Asian, Black or African American, Hispanic or Latino, Native Hawaiian or Other Pacific Islander, and White groups for self-reported race and ethnicity collections. By providing these guidelines, comparable detailed race and ethnicity data will be supported across Federal agencies. However, many questions remain, such as under what conditions detailed data should *not* be collected. In addition, the Subgroup will continue to discuss whether OMB should require or, alternatively, strongly support but not require Federal agencies to collect detailed data.

4. Relevance of Terminology

Although many respondents report within the race and ethnicity categories specified by the 1997 standards, the standards themselves may not be well understood by the public. Over time, some terms may be preferred over others. For example, from comments received, the Subgroup noted some confusion about the opportunity to select more than one category, and about the use of the terms of race, origin, and ethnicity. Some respondents do not identify themselves within current Federal definitions of the race and ethnicity minimum categories.

Additionally, references to geographic location in descriptions of race and ethnicity categories may be incomplete, unclear, inconsistent, or confusing. Given both the low prevalence of these geographic locations appearing as write-in responses, as well as public comments to the September 30, 2016 Federal Register Notice advocating against use of geographic regions in race/ethnicity group definitions, the Subgroup recommends no changes be made to the current standards to specifically incorporate the following geographic locations into any existing race or ethnicity category: Australian (including the original people of Australia/the Aborigines), Brazilian, Cape Verdean, New Zealander, Papua New Guinean.

The Subgroup observed a lack of clarity in the classification of several groups in the current standards. Further, there is an error in the 1997 OMB classification standards with "Cuban" being listed twice in the reporting category of "Hispanic or Latino." The Subgroup noted that population size currently does not govern the listing of detailed groups across all minimum reporting categories but recommends that population size should be used to determine which duplicate initial category mention of "Cuban" should remain. The Subgroup also considered whether the current ordering of the classification list should be updated to reflect current population size. As a next step, the Subgroup plans to apply this rationale to the classification listing and determine the magnitude and benefit of any resulting changes.

The Subgroup noted that some terms used in the current standards to describe race/ethnicity may not (or no longer) resonate with the public, such as "Negro" and "Far East." Because respondents may find certain terms to be outdated and/or offensive, the Subgroup recommends that these terms be removed from the standards.

The Subgroup examined reporting patterns of South and Central American Indian respondents to determine if self-identification and reporting behavior is consistent with the current standard, which includes them in the American Indian or Alaska Native reporting category. The Subgroup noted more research and public input would be necessary before adding specific South or Central American subgroups to the current description of the AIAN category.

The Subgroup considered the use of "Principal Minority Race" in the current standard. Given the changes in the US population and the language in some legislation and Executive Orders, the Subgroup examined whether the current standards still refer to and provide guidance on reporting "principal minority race," and if the term is still meaningful, whether the designation should be guided by minority group population size, historical disadvantage, or some other principle. The Subgroup considered whether "principal minority race" be expanded to include ethnicity. In particular, the Subgroup considered whether referring to Black or African American as the "principal minority race" is still relevant, meaningful, accurate, and acceptable. The Subgroup also plans to consider if "Hispanic" should be among the groups considered, and the salience of alternative terms (e.g., "principal minority race/ethnicity").

The above is an excerpt only. The remainder of this interim report provides detail supporting the initial findings and proposals of each Subgroup.

Author Biographies

Victor Rubin

Victor Rubin, PhD, Vice President for Research, leads, designs, and conducts knowledge-building activities to create a strong research base for PolicyLink. An urban planner with broad experience in community development, education, and social policy, Dr. Rubin guides the PolicyLink analyses of healthy communities, economic growth, infrastructure, and other areas. He previously directed the U.S. Department of Housing and Urban Development's Office of University Partnerships, and served as adjunct associate professor of city and regional planning at the University of California, Berkeley.

Danielle Ngo

Danielle Ngo, MA/MS, Program Associate, identifies policy opportunities, develops strategies and tactics, and coordinates across partner organizations to lift up community-led efforts for health equity, food security, and water security. Prior to joining PolicyLink, she conducted research on food sovereignty and sustainable agriculture in New Orleans and anti-displacement measures in Boston.

Ángel Ross

Ángel Ross, MCP, is a Program Associate at PolicyLink, where they research and collaborate with local and national partners through the National Equity Atlas and the All-In Cities initiative. Ángel specializes in leveraging data to support advocacy and to build the case for equity and racial economic inclusion.

Dalila Butler

Dalila Butler, MPH, Associate Director, works to promote social, economic, and health equity through environmental and policy change. She provides technical assistance to communities across the country and supports research for health equity projects. She was the coordinator of Equity Summit 2018, a national gathering of over 4,000 local leaders, policymakers, and researchers.

Nisha Balaram

Nisha Balaram, BA, Program Coordinator, provides programmatic support for health equity projects through research analysis, budget management, and editorial feedback. Currently, she works on issues of police violence and systemic bias, and movement building in the public health field. She has previously worked as a policy coordinator for the Greenlining Institute.

PolicyLink

Lifting Up What Works®

Headquarters
1438 Webster Street
Suite 303
Oakland, CA 94612
t 510 663-2333
f 510 663-9684

Communications
75 Broad Street
Suite 701
New York, NY 10004
t 212 629-9570

Washington, DC
1200 18th Street, NW
Suite 1200
Washington, DC 20036

www.policylink.org
Facebook: /PolicyLink
Twitter: @policylink