Disaggregation of Race and Ethnicity Group Data Research-to-Practice Issues in Clinical Environments

Farah Kader, MPH,
Section for Health Equity, Department of Population Health, NYU Grossman School of Medicine, 227 E. 30th St, 8th floor, New York, NY 10016

Perla Chebli, MPH, PhD
Section for Health Equity, Department of Population Health, NYU Grossman School of Medicine

Gaps in patient data have long hindered researchers' ability to quantify the extent of health injustices and disparities. This is attributed to several issues, one of which may be discordance between self-identity and the race and ethnicity labels provided in the minimum federal reporting requirements, determined by the Office of Management and Budget (OMB). Disconnects between the OMB classifications and self-perceptions of race and ethnicity can result in population miscounts and misclassification, particularly among Asian, Black, Latino, and Native American individuals. Further, an increasing number of people in the United States select "other race" or decline to self-identify on the census and other surveys.

Researchers and civil rights groups recommend making data more inclusive, reliable, and complete by disaggregating the standard OMB categories into more specific groups. However, health systems and patient populations are accustomed to the federal definitions of race and ethnicity, which are unchanged since 1997, making health care leaders, such as executives, clinicians, and policymakers, reluctant to deviate from the minimum standard on patient forms.

Nevertheless, the medical field stands to benefit from race and ethnicity data disaggregation more than any other sector. Individual OMB race and ethnicity categories currently homogenize diverse, multiethnic, and multiracial populations, potentially masking within group health differences due to social, biological, and cultural factors and thus possibly limiting the ability to respond effectively to emerging health needs.

Americans of Middle Eastern and North African Descent: A Case Study

With the rapid onset of the pandemic came a new sense of urgency among local health officials to follow equity driven data collection models. The case of Middle Eastern and North African American people, who federal agencies categorize as White, underscores this shift.

farah.kader@nyulangone.org .

Throughout 2020, there were reports of high COVID-19 positivity, hospitalization, and fatality rates in neighborhoods with large communities of Arab, Persian, and other ethnocultural groups who could benefit from a distinct Middle Eastern and North African identifier in vital statisticsdata.[1] These findings persuaded certain localities to collect and report Middle Eastern and North African individuals' COVID-19–related outcomes. For example, amidst pressure from the public and local media around Chicagoland vaccine uptake disparities, the Illinois Department of Public Health made Middle Eastern and North African an option with the race categories for vaccination registrants in 2021. In Michigan, intake forms at neighborhood COVID-19 testing sites offer the option to select Arab ethnicity in addition to another race.

Both states applied different terms and methods (racial vs ethnic category) to capture the Middle Eastern and North African population, an inconsistency that illustrates how accepted definitions and labels surrounding race and ethnicity shift overtime. For instance, *Arab* may be considered too narrow to encompass a pluralistic Middle Eastern and North African population. Yet, there may be some resistance to using the Middle Eastern and North African label due to its colonial origins, defining the region in relation to Europe, with some advocating for usage of *Southwest Asia and North Africa*. Study findings by Maghbouleh et al[2] also called into question whether a new Middle Eastern and North African category should be considered a race or ethnicity, the latter of which enables people from this population to select an ethnicity along with another race, as is the standard for Hispanic/ Latino(a) individuals.

Even though terminological consensus may not be achievable, both Michigan and Illinois implemented strategies that allowed for reaggregation of data back into the national standards, so that multiple reporting systems could maintain data harmonization. These more equitable COVID-19 monitoring efforts may lay the groundwork for broader health data policies. Clinical studies have reported that Arab American individuals experience other health risks, such as low birth weight and high incidence of metabolic disorders, compared with other non- Hispanic White individuals.[1] These and other possible disparities may go unnoticed without adequate data collection of race and ethnicity. The existing research mostly involves retrospective cleaning of current health registries using Arabic surname algorithms and other techniques, methods that are less accurate than the best practice of self-reporting. It is, therefore, vital that health governance systems incorporate inclusive data collection instruments for rapid detection of group-specific disparities and targeted health interventions.

A Changing Landscape

Populations known as Arab, Middle Eastern and North African, or Southwest Asian and North African are not the only ones affected by data limitations. The Asian race category, for instance, is defined as those having origins in the "Far East, Southeast Asia, or the Indian subcontinent." Civil rights organizations have long called for disaggregation of this diverse and populous group. Recently, New York passed a legislation package requiring agencies to provide disaggregated Asian and Native Hawaiian and Other Pacific Islander race options, such as Bangladeshi, Filipino, Hmong, Korean, and other traditionally homogenized

identities. In a medical context, this community driven legislation could allow systematic parsing of health differences associated with specific genotypes, cultures, language barriers, and other variables. Individuals who identify with Black and Hispanic/Latino OMB categories also may experience similar challenges due to category homogenization, which have been explored over decades of public health research.

State-level actions may have had a "trickle-up" effect on a more national focus toward data equity. The US Census Bureau announced in May 2022 that it will consider deviating from OMB standards for the Census questionnaire. The proposed changes include adding a Middle Eastern and North African race category and combining the 5 current race categories and the Hispanic/Latino ethnicity category into a single question.[3] One month later, the White House announced that the OMB would also examine these recommendations. A federal interagency working group has subsequently begun evaluating research and public comments to inform its recommended revisions to the federal minimum reporting requirements.[4]

These pending deliberations raise multiple research-to-practice implementation questions for the clinical setting, another key reason health professionals, patient advocates, and other leaders in health should not wait to coordinate their own granular approaches to data collection. Because the OMB is unlikely to revise the minimum standards with the level of granularity that is often recommended for biomedical research and clinical practice, medical practices and institutions should take community perspectives and local population demographics into account when making data collection decisions. Often overlooked is a recommendation included in the OMB's 1997 guidance to disaggregate locally populous groups as long as they can be rolled up into the standard categories, which is important for allowing some level of data interoperability and longitudinal comparability over time. Additionally, surveys could be locally tailored and collaboratively designed; this increases community buy-in and the likelihood that the terms used to describe communities are clear and culturally appropriate.

Starting these conversations as early as possible could expedite the operationalization of disaggregation practices and the planning it requires. For example, amending the necessary data infrastructure to collect and analyze disaggregated categories calls for administrative capacity, funding, and input from stakeholders such as insurers and electronic medical records companies. Data managers must also weigh the benefits of inclusivity against the statistical and confidentiality-related challenges of small population sizes; existing disaggregation guides offer numerous strategies that could help achieve this balance.

Borrell et al[5] also emphasized the distinction between perceived ("socially assigned") and self-reported race and ethnicity, noting that patient race and ethnicity recorded by a third-party observer, such as hospital staff, can lead to misclassification. Researchers argue that both socially assigned and self-identified race and ethnicity are valuable for evaluating differential health outcomes resulting from racism, regardless of ethnocultural identity.[5] This is especially important for multiracial populations, such as Arab and Hispanic/Latino individuals who also have African ancestry. These individuals may prefer labels that

more accurately reflect those who have multiracial/multiethnic backgrounds but experience additional discrimination due to anti-Black racism.

Thus, clinicians and administrative staff require training and patient education materials on how to explain definitions of race, ancestry, and ethnicity and encourage category selections that encompass both self-identified and perceived race.[6] Permitting patients to select multiple races and ethnicities is another best practice to obviate misclassification or nonresponse. Likewise, concerted training efforts for health professionals and staff should include tools for addressing patient uncertainty about more granular racial and ethnic inquiry early in the implementation period, not least among those who experience heightened government surveillance and profiling.

Conclusions

These are just a few of the many adjudications required for meaningful, equitable data collection procedures in health settings. Given that changes to the federal standards are an almost certain future, leaders in health care must begin planning for data collection and analysis methods that are participatory, evidence-based, and dynamic enough to shift with ever-changing social constructs of racial and ethnic formation. To overcome inevitable challenges with this work, health organizations and professionals could begin to use existing, evidence-based data disaggregation guides to inform institutional practices with a greater level of detail than broad national directives. Recommended tactics, particularly those that prioritize community buy-in, are essential for precision medicine and other medical models that acknowledge the combined influence of population-specific genetics, physical environments, and sociocultural influences on clinical outcomes. Thus, additions and alterations to racial and ethnic data management procedures must not be considered unnecessary burdens, but an investment in health data policies favoring racial equity.

Conflicts of Interest:

References

1. Awad GH, Abuelezam NN, Ajrouch KJ, Stiffler MJ. Lack of Arab or Middle Eastern and North African health data undermines assessment of health disparities. Am J Public Health. 2022;112(2):209–212. [PubMed: 35080949]

2. Maghbouleh N, Schachter A, Flores RD. Middle Eastern and North African Americans may not be perceived, nor perceive themselves, to be White. Proc Natl Acad Sci U S A. 2022;119(7):e2117940119. [PubMed: 35131945]

3. Schneider M. Census ready to study combining race, ethnicity questions. Associated Press. Accessed August 22, 2022. https://abcnews.go.com/Lifestyle/wireStory/census-ready-studycombining-race-ethnicity-questions-84550843

4. Orvis K. OMB launches new public listening sessions on federal race and ethnicity revision. Accessed September 18, 2022. https://www.whitehouse.gov/omb/briefing-room/2022/08/30/omb-launches-new-public-listening-sessions-onfederal-race-and-ethnicity-standards-revision/

5. Borrell LN, Elhawary JR, Fuentes-Afflick E, et al. Race and genetic ancestry in medicine: a time for reckoning with racism. N Engl J Med. 2021;384(5):474–480. [PubMed: 33406325]

6. Agency for Healthcare Research and Quality. Data improvement through education and training of hospital staff. Accessed August 22, 2022. https://www.hcup-us.ahrq.gov/datainnovations/raceethnicitytoolkit/data_improve_edu.jsp