COLLECTION, ANALYSIS, AND REPORTING OF ASIAN AMERICAN HEALTH DATA

Matthew K. Chin,
Yousra Yusuf,
Laura C. Wyatt,
Lan N. Đoàn,
Rienna G. Russo, Farah Kader,
Lloyd Feng, Lauren Fu,
Simona C. Kwon, Stella S. Yi

NYU CENTER FOR THE STUDY OF ASIAN AMERICAN HEALTH

CACF
Coalition For Asian American Children+Families

NYU Grossman School of Medicine at NYU Langone Health

Version 05/23/2023

TABLE OF CONTENTS

EXECUTIVE SUMMARY

The Center for the Study of Asian American Health (CSAAH) within the NYU Grossman School of Medicine at NYU Langone Health and the Coalition of Asian American Children and Families (CACF) have created a manual to promote best practices for collecting, analyzing, and reporting Asian American health data.

We recommend anyone who is working with quantitative data that pertains to Asian Americans to review this guide in order to promote an accurate and nuanced characterization of Asian American health.

Data Collection

The current national standards for collecting racial/ethnic data were issued by the Office of Management and Budget's 1997 Directive 15 though leaders in civil rights and academia have voiced concerns over their definition and application. CSAAH, CACF and the New York Academy of Medicine (NYAM) developed a race/ethnicity form for collecting disaggregated race/ethnicity information in the New York metropolitan area which was informed by three main sources: 1) national policy recommendations by the State Data Collaborative;[1] 2) American Community Survey 2019 data; and 3) feedback from community focus groups and trusted subject matter experts.

Our Revised Race/Ethnicity Form asks participants to identify as one or more of seven aggregate race/ethnicity groups – American Indian/Alaska Native (AI/AN), Asian, Black or African American, Hispanic/Latino, Middle Eastern or North African (MENA), Native Hawaiian or Pacific Islander (NH/PI), or White – and then asks participants to identify as one or more of 70+ detailed race/ethnicity categories. The form can be adapted in format and to include different detailed categories based on project or research focus and needs.

When conducting primary data collection among Asian American populations, there are several factors to consider regarding the study population of focus and survey design. These factors include but are not limited to the study population: age, income, immigration history, education level, and English proficiency. Data from the U.S. Census Bureau like the Decennial Census and American Community Survey can be used for information on the demographic makeup of the study population in a specific geographic region or neighborhoods where the population resides for targeted recruitment efforts.

Data Analysis

When constructing racial/ethnic group variables for analysis we recommend creating:

1) A mutually exclusive aggregate race/ethnicity variable where individuals who identify as more than one aggregate racial/ethnic identities are labeled 'Multiracial' (AI/AN, Asian, Black, Hispanic/Latino, MENA, NH/PI, White, Multiracial, Other);

2) A mutually exclusive detailed race/ethnicity variable where individuals who identify as more than one detailed racial/ethnic identifies are labeled 'Mixed' (e.g., Chinese, Asian Indian, Filipino, Japanese, Asian Mixed, Detailed Other, etc.);

3) If appropriate, for Asian populations, a mutually exclusive Asian regional race/ethnicity variable (e.g., East Asian, South Asian, Southeast Asian, Central Asian);

4) If appropriate, a non-mutually exclusive race/ethnicity variable for each unique aggregate and detailed race/ethnicity category; and

5) If appropriate, further granular variables for Multiracial/Mixed individuals.

When conducting secondary data analysis, it is important to consider how Asian American race/ethnicity data has been collected in the original dataset. Race/ethnicity data may be missing, unknown or other. Detailed race/ethnicity data may be limited or nonexistent. Additionally, many datasets are not always designed to be inclusive of non-English speakers, and sampling bias may result from higher rates of English proficient Asian Americans who tend to have more favorable socioeconomic indicators.[2] Data quality and therefore representation should be qualified in Methods, Discussion and Limitations sections. Sensitivity analyses should be conducted for more precise evaluation of the heterogeneity of the Asian American sample. Careful consideration of who was included or not included in the data being analyzed will help researchers evaluate biases in data collection and recruitment and the generalizability of the analyses to Asian American populations. Partial missingness of data using complete case analysis, imputation and advanced imputation and total missingness using standardization, simulation, and sensitivity analyses should be performed.

In the case studies, we examine the impact of undercounting in minority populations. The U.S. Census has a history of undercounting immigrant and minority populations, which can impact the accuracy of public health statistics. The 2020 U.S. Census may have had a greater undercount due to a potential citizenship question and the COVID-19 pandemic.[3,4] Correcting for the undercount requires specific birth, death, and migration data, making it challenging for most researchers. In the case of New York City (NYC) COVID-19 vaccination statistics, the implications of an undercount for Asian American and NH/PI populations were examined.

Data Reporting

When reporting Asian American data, descriptions should include all eligibility and exclusion criteria, including explicit details about the disaggregated Asian racial/ethnic group(s) sampled, the dates of the study period, the type of study design, and the number of individuals surveyed or enrolled. Having this information can help researchers better determine next steps for research and resource allocation by contextualizing the population sample.

The demographic details of the study population play a crucial role in preventing harmful stereotypes and generalizing data. These details should include racial/ethnic information, income, education, English language proficiency, language spoken at home, and nativity. It is important to be transparent about any groupings, such as aggregating into an "Other" category or combining smaller Asian race/ethnicity categories into regional groupings. When describing multiracial or Asian mixed populations, it is useful to include details of the combinations of race/ethnicities.

Reporting race and ethnicity to meet OMB guidelines is required for federal agencies as well as other state and local organizations that model their reporting guidelines on the OMB's 1997 directive. The Agency for Healthcare Research and Quality (AHRQ) provides a guideline for roll-up of detailed race/ethnicity categories while the Institute of Medicine (IOM) provides further granular guidance.

The process of rolling up detailed race/ethnicity categories into aggregate categories has gray areas and can lead to mismatches between aggregate categories and self-identification of race/ethnicity. This can result in individuals who identify as one aggregate category but are categorized differently based on their detailed race/ethnicity. It is recommended that participants' reported aggregate race/ethnicity should always reflect their selected aggregate category and not be changed. In cases where only detailed race/ethnicity information is provided, researchers should detail their assignment rules for ambiguous categories.

The representation of Asian American experiences in data can be distorted by three racialized stereotypes - the model minority, healthy immigrant effect, and perpetual foreigner. These stereotypes can be a result of implicit biases in researchers – embedded in their own lived experiences – and should be considered when reporting Asian American data.[5]

Recommendations for considerations when discussing results from Asian American data:

1) Discuss who is represented in the data to provide context to research findings and avoid creating generalizations about the Asian American population.

2) Be cautious when discussing Asian multiracial and mixed data as these categories are diverse.

3) Hypothesize the potential impact of missing data in the analysis.

4) Consider complementing statistics from large databases with community-based research to uncover health problems that may be hidden.

5) Seek and include community feedback to enhance trust and cultural resonance.

6) If Asian data is not available, discuss the implications and provide recommendations for additional research that includes Asian Americans.

This manual provides recommendations for public health practitioners, allied health professionals, health equity researchers, and data managers to employ in their own research with immigrant and minority groups in the U.S. Understanding and addressing the nuances of data collection, analysis, and reporting is the cornerstone to accurate representation of groups in data.

Suggested Citation: Chin MK, Yusuf Y, Wyatt LC, Đoàn LN, Russo RG, Kader F, Feng L, Fu L, Kwon, SC, and Yi SS. Collection, Analysis, and Reporting of Asian American Health Data. E-published by Center for the Study of Asian American Health at NYU Langone; 2023. Available from: https://aanhpihealth.org/resource/asian-american-manual-2023

Artwork by Isabel Lu

For more information, please contact:
Dr. Stella Yi, MPH, PhD
Center for the Study of Asian American Health
Email: Stella.Yi@nyulangone.org

1. INTRODUCTION

1.1 Purpose

The Center for the Study of Asian American Health (CSAAH) within the NYU Grossman School of Medicine at NYU Langone Health in collaboration with the Coalition of Asian American Children and Families (CACF) have developed this manual for public health practitioners, allied health professionals, and health equity researchers to promote best practices for data collection, analysis and reporting of Asian American health data.

Our recommendations are based on over 20 years of research experience and advocacy as well as draw from national guidelines for research on racial/ethnic minority populations. These best practices aim to inform Asian American population health research but may be useful for other racial/ethnic groups, like small or emerging populations. The application of these guidelines will vary depending on the study design, geography, population, and the specific needs of each research study.

2. DATA COLLECTION

2.1 Race/Ethnicity Data

2.1.1 Definition and Background

The definitions for race and ethnicity are frequently imprecise, with overlapping and often contradictory descriptions. Two useful working definitions for race and ethnicity are presented in the Institute of Medicine report, *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement.*[6]

Race – "A socioeconomic concept wherein groups of people sharing certain **physical characteristics** are treated differently based on stereotypical thinking, discriminatory institutions and social structures, a shared worldview, and social myths."[6]

Ethnicity – "A concept referring to a shared **culture and way of life**, especially reflected in language, religion, and material culture products."[6]

Both Race and Ethnicity:

- Are socially constructed or made-up categories; [7,8]
- Have no consistent biological or genetic basis; [9,10]
- Can be self-reported or socially assigned; [11]
- Can change over time in terms of how one self-identifies as well as how race is assigned to someone; and [12]
- Are frequently conflated and used interchangeably in public health literature. [6]

Throughout this manual, we intentionally use '**race/ethnicity**' to refer to these two terms as a singular combined concept and use 'race' and 'ethnicity' as individual terms only in the context of describing data collection and reporting practices of other researchers or government entities.

History of U.S. Census Race/Ethnicity Data Collection

Race and ethnicity data categories have not been consistent since the U.S. Census began collecting race/ethnicity information in 1790 (Figure 1[12]) and instead reflect prevailing political attitudes, advocacy, and immigration patterns of the time. For example:

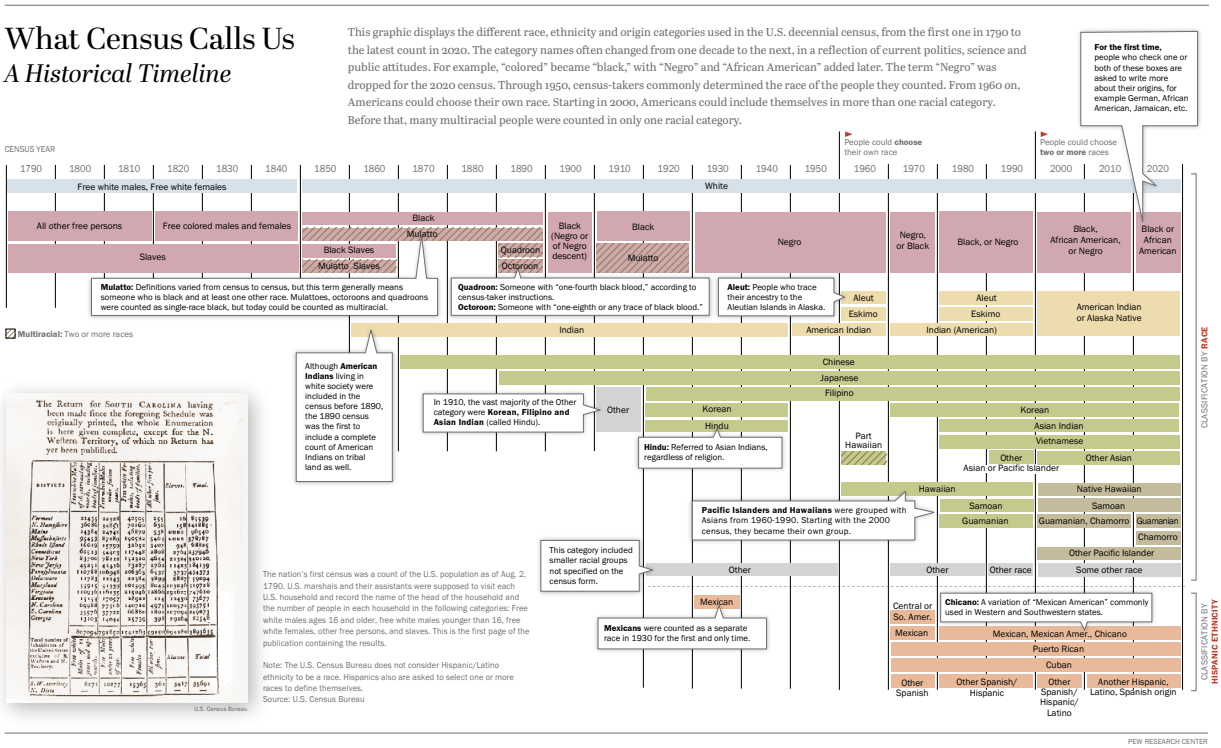
- Asian categories have been included on the Census since 1860, when Chinese was added as a category. In 1890, Japanese was added, and in 1920, Filipino, Hindu, and Korean were added. Hindu and Korean were removed in 1950, but Korean was added back in 1970. Asian Indian and Vietnamese were both added in 1980; Other Asian was added in 1990; and a write-in response was included beginning in 2020.[12]
- In 1980, a second question, representing Hispanic origin, was added to the Census.[12]

- In 2000, respondents were given the option to choose more than one race; this was a result of lobbying by advocates for multiracial individuals who wanted their identities to be recognized.[13,14]

In addition, the definitions in which racial/ethnic data has been collected and categorized have changed over time:

- Between 1790 and 1950, the census takers determined the race of the Americans that they counted. Beginning in 1960, Census respondents could choose their own race.[12]
- Prior to 1960, people of multiracial non-White backgrounds were categorized according to their father's race, and those who were white and another race were categorized as the minority race.[12]

Figure 1. U.S. Census Race/Ethnicity Timeline[15]



Office of Management and Budget – Race and Ethnicity Collection Standards

Currently, race and ethnicity data collection categories are largely guided by the federal Office of Management and Budget (OMB's) 1997 Directive 15,[16] which requires federal agencies to minimally inquire Hispanic ethnic origin and a minimum of five racial categories, with the use of "some other race" permitted on the Census.[17] See the categories and details of each below:

1) The OMB two minimum Ethnicity options include:

- Hispanic or Latino[18] – a person of Cuban, Mexican, Puerto Rican, Cuban, South or Central American, or other Spanish culture or origin, regardless of race; and
- Not Hispanic or Latino.

2) The five minimum Race categories required by the OMB are:

- American Indian or Alaska Native (AI/AN) – origins in any of the original peoples of North and South America (including Central America), and who maintains tribal affiliation or community attachment;
- Asian – origins in any of the original peoples of the Far East, Southeast Asia, or the Indian subcontinent;
- Black or African American – origins in any of the black racial groups of Africa;
- Native Hawaiian or Other Pacific Islander – origins in any of the original peoples in Hawaii, Guam, Samoa, or other Pacific Islands;
- White – origins in any of the original peoples in Europe, the Middle East, or North Africa; and
- Some other race (Optional; OMB permits the Census to use this category).

Key OMB Race and Ethnicity Data Collection Standard Considerations:

- Guidelines are only applicable to federal government agencies.
- Standards recommend only the minimum racial/ethnic categories for data collection and reporting – the collection of detailed race/ethnicity is encouraged.
- Additional racial/ethnic categories must be able to be aggregated (rolled up) into the minimum six race/ethnicity categories (Data Reporting) for federal reporting and standardization purposes.

CASE IN HISTORY: NEW YORK STATE DATA DISAGGREGATION BILL

For over a decade, the Coalition of Asian American Children and Families (CACF) has led an advocacy campaign for Asian and Pacific Islander data disaggregation in New York called Invisible No More. CACF convened a coalition of diverse partners consisting of Asian American and Pacific Islander-serving community-based organizations, including allies, unions, and philanthropic organizations to continually build public pressure on elected officials to pass data disaggregation legislation while concurrently building and strengthening relationships between the coalition members and elected officials to emphasize the importance and need for data disaggregation.

In 2016, Mayor Bill DeBlasio signed NYC's first data disaggregation bill into law, Local Law 2016/126. The law stipulated that the Administration for Children's Services, Department of Education, Department of Health and Mental Hygiene, Department for Youth and Community Development, Department for the Aging, Department of Social Services, and Department of Homeless Service must include a voluntary self-report form with questions on race and ethnicity for people to fill out. The law specified that data collection must include the top 30 ancestry groups and languages spoken according to U.S. Census data. The affected city agencies were also tasked with identification of forms that could be updated to include the expanded questions on race and ethnicity questions. Notably, this law was passed in the City Council and signed into law by the Mayor in conjunction with bills for data collection on multiracial identity (Local Law 2016/127) and sexual orientation/gender identity (Local Law 2016/128).

CACF's advocacy on implementation of the city data disaggregation law revealed a number of issues. First, the majority of forms administered by the affected city agencies were not eligible for updating to include demographic questions on race/ethnicity because the forms fell outside the jurisdiction of city agencies. This resulted in a shifted focus in the CACF-led Invisible No More campaign to target New York State legislators and the Governor to enact Asian American and Pacific Islander data disaggregation at the state level. CACF worked closely with legislative champions Assembly member Yuh-Line Niou and Senator Julia Salazar to pass an Asian American and NH/PI data disaggregation bill in both the Assembly and Senate with bipartisan support in 2019. Unfortunately, Governor Andrew Cuomo vetoed the bill on the grounds that the bill did not have an accompanying funding source and did not offer sufficient privacy protections to undocumented New Yorkers.

The following legislative cycle, CACF continued the collaboration with Assembly member Niou and Senator Salazar to pass a revised version of the bill again with bipartisan support in both chambers. While the bill awaited a signature to become law, there was increased public pressure on Governor Andrew Cuomo to sign the bill as the COVID-19 pandemic ravaged Asian American and NH/PI communities in NYC and New York State. However, Governor Cuomo only acted in support of Asian American and NH/PI data disaggregation in the aftermath of the tragic Atlanta shootings, allocating $3 million toward a data disaggregation project with the New York State Department of Health, out of which came this very manual.

Throughout 2021, CACF continued advocacy efforts with lawmakers on the urgency of signing the bill into law and keeping up the pressure amid a change in leadership. In December 2021, New York State Governor Kathy Hochul signed the state's first data disaggregation law focused on Asian Americans and NH/PIs. The law mandates that all state agencies, boards, commissions, and offices that collect demographic data now must collect and disaggregate data for the top 10 most populous Asian groups (Chinese, Asian Indian, Filipino, Korean, Bangladeshi, Pakistani, Japanese, Vietnamese, Taiwanese, and Nepali) and the top four Pacific Islander (Native Hawaiian, Samoan, Guamanian, and Chamorro) groups based on U.S. Census population data, all explicitly noted in the law.[19]

Such legislative victories were only possible with the broad and sustained support of diverse coalition of groups inside and outside the Asian American and NH/PI community, as well as by continued demonstrations on the importance of disaggregated data to elected officials, as the basis for developing good policies for long marginalized communities.

Why is Hispanic or Latino an ethnicity category and not a race category?

The standard for collecting and reporting Hispanic ethnicity was driven in the 1970s by an interest in monitoring civil rights data[20] and disaggregating Hispanic White and Hispanic Black populations.[14]

The separate Hispanic ethnicity category has been critiqued by the American Anthropological Association as well as members of the U.S. Commission on Civil Rights,[21,22] and it has led to inconsistent responses among Hispanic and Latino respondents. A Pew Research Center analysis of the 2010 U.S. Census found two-thirds of Hispanic adults say being Hispanic is part of their racial background.[23] An analysis of 2020 U.S. Census data found that 42% of Hispanics marked their race as "some other race" without marking any other response, followed by 33% selecting two or more racial groups, and 20% selecting White as their race.[24]

Middle Eastern or North African and Arab American communities

There is currently no standard MENA and Arab American racial category on the U.S. Census or many other data collection tools. As a result, MENA and Arab American individuals are currently grouped with White, and there is no direct way to count members of this group in official statistics.

The MENA region is racially, ethnically, and linguistically diverse. The colonial history of many countries, both Arab and non-Arab, means that individuals from this region may have European, African, and other ancestries produce constructs of Whiteness and Blackness in addition to other racializations. This is especially evident in North Africa, where there are larger groups of indigenous ethnic groups (e.g, Amazigh) and Afro-Arabs, or Arabs of African descent. However, the race standards in the U.S. dictates that all such individuals must identify as White.

Recent research has found that with respect to ancestry, MENA individuals reject identifying as exclusively White; however, individuals of North African ancestry were more likely to identify as Black even when offered a MENA category. Additionally, respondents identifying as Muslim or non-religious were more likely to choose the MENA category than those identifying as Christian.[17]

States with large MENA populations have employed different ways of collecting Arab or MENA ancestry. This can pose challenges for inter-state data harmonization. For example, Illinois uses the term MENA on its data collection forms and considers it a race, similar to White, Black, and Asian categories.[26] Alternatively, the Michigan Department of Health and Human Services uses the term 'Arab American' on its COVID-19 testing as an ethnicity option – similar to the OMB Hispanic/Latino category.[27] Both states applied different terms and methods to capture the population – an inconsistency that illustrates how accepted definitions and terms surrounding race and ethnicity shift over time. For instance, 'Arab' may be considered too narrow to encompass a pluralistic MENA population. Yet, MENA label itself also faces resistance due to its colonial origins,[28] defining the region in relation to Europe, with some advocating for increased usage of 'Southwest Asia and North Africa' (SWANA).[29]

MENA populations in particular are growing all over the U.S. due to immigration, which may be one reason why national institutions such as the American Medical Association adopted a policy supporting MENA race reporting on medical records, research surveys, and medical education.[30]

OMB Category Revision

In the summer of 2022, the White House announced that the OMB will be reconsidering national race and ethnicity standards for the first time since 1997.[31] In January 2023, the OMB released a set of proposals for updating its race and ethnicity standards, open to public comment until April 12th.[32] Notably, the proposal asks for comment on adding a MENA category to the required five race categories as well as including Hispanic/Latino with other race categories. It also is seeking comment on requiring the collection of additional detailed race/ethnicity categories. OMB plans to complete revisions to its standards by Summer 2024.[32]

2.1.2 Revised Race/Ethnicity Form – Collecting Data in New York City

Below, we illustrate the approach that we undertook in order to develop updated race/ethnicity questions for all racial/ethnic groups, and we describe in detail the Asian categories that are recommended for use when collecting data in the New York metropolitan area. We then share what this process could look like in your setting of interest.

For our updated race/ethnicity question, we referenced three main sources:

1) A template form developed to capture race/ethnicity data by a collaborative including the Robert Wood Johnson Foundation (RWJF) and Asian & Pacific Islander American Health Forum (APIAHF);[1]

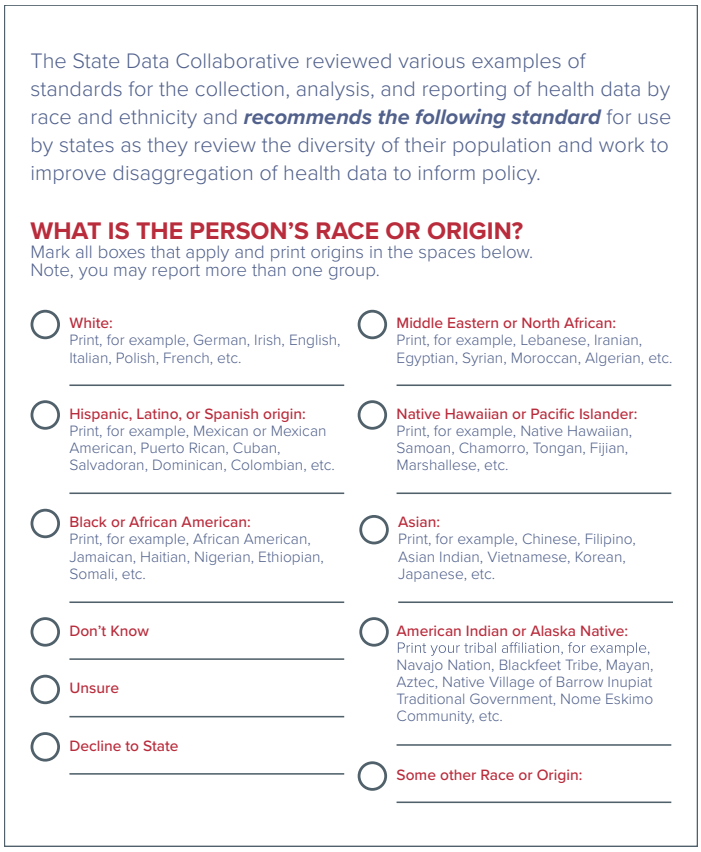
2) Detailed racial/ethnic categories for a particular geographic region using data from the 2015-2019 5-year American Community Survey (ACS); and

3) Recommendations from key stakeholders and local experts including community advocates, health systems administrators and IT, state department of health, and community members.

Format

The format of the Revised Race/Ethnicity Form was modeled after a Robert Wood Johnson Foundation (RWJF) effort led by our national partner, the Asian & Pacific Islander American Health Forum (APIAHF) in 2016. This race/ethnicity form (Figure 3) was developed in collaboration with four civil rights organizations and APIAHF– Arab Community Center for Economic and Social Services (ACCESS), UnidosUS, National Urban League, and the National Congress of American Indians—called the State Data Collaborative with the original purpose of serving as potential race/ethnicity question for the 2020 U.S. Census survey (Figure 2). This form was also informed by conversations with stakeholders that included state policymakers, state health agency staff, researchers, and community advocates.[33]

Figure 2. APIAHF Race/Ethnicity Form[33]



The State Data Collaborative reviewed various examples of standards for the collection, analysis, and reporting of health data by race and ethnicity and ***recommends the following standard*** for use by states as they review the diversity of their population and work to improve disaggregation of health data to inform policy.

WHAT IS THE PERSON'S RACE OR ORIGIN?

Mark all boxes that apply and print origins in the spaces below.
Note, you may report more than one group.

- White: Print, for example, German, Irish, English, Italian, Polish, French, etc.
- Middle Eastern or North African: Print, for example, Lebanese, Iranian, Egyptian, Syrian, Moroccan, Algerian, etc.
- Hispanic, Latino, or Spanish origin: Print, for example, Mexican or Mexican American, Puerto Rican, Cuban, Salvadoran, Dominican, Colombian, etc.
- Native Hawaiian or Pacific Islander: Print, for example, Native Hawaiian, Samoan, Chamorro, Tongan, Fijian, Marshallese, etc.
- Black or African American: Print, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc.
- Asian: Print, for example, Chinese, Filipino, Asian Indian, Vietnamese, Korean, Japanese, etc.
- Don't Know
- American Indian or Alaska Native: Print your tribal affiliation, for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat Traditional Government, Nome Eskimo Community, etc.
- Unsure
- Decline to State
- Some other Race or Origin:

Key differences on the OMB race/ethnicity standards on the APIAHF race/ethnicity form included:

- Combining Race and Ethnicity questions into a single question;
- Adding a Hispanic, Latino, or Spanish Origin group;
- Adding a Middle Eastern or North African racial group;
- Allowing participants to provide both aggregate race options and detailed race options; and
- Allowing users to select multiple race and some other race categories.

Detailed Categories

While the APIAHF form uses a fill-in-the-blank structure to collect detailed race/ethnicity information, we adapted our Revised form to include predetermined categories in order to allow for more efficient data analysis. To tailor these predetermined detailed race/ethnicity categories to the New York metropolitan area, we analyzed publicly available 2015-2019 ACS data[34] to find the top 10 ancestry categories for each of the seven aggregate race/ethnicity categories (AI/AN, Asian, Black, Hispanic/Latino, MENA, NH/PI, White) outlined in the APIAHF form. In descending order, for the Asian category, this included: Chinese, Indian (Asian), Korean, Filipino, Pakistani, Bangladeshi, Japanese, Taiwanese, Vietnamese, and Uzbek.

Feedback

Lastly, we sought feedback on initial drafts of our form through focus groups with community members and discussions with local subject matter experts.

Focus Groups : [35] Focus groups were conducted by the Center for Evaluation and Applied Research (CEAR) at the New York Academy of Medicine (NYAM). Thirteen focus groups and one key informant interview were conducted with individuals who self-identify as belonging to one of eight aggregate racial/ethnic groups— American Indian/Alaska Native, Asian, Black, Hispanic/Latino, Middle Eastern or North African, Multiracial/Mixed, Native Hawaiian or Pacific Islander, and White. Recruitment was conducted through local community-based organizations and online flyers, and participation was limited to English-speaking, New York State residents, aged 18 years or older who could use Zoom.

Focus groups aimed to gain perspective on racial/ethnic identity, preferences for reporting race/ethnicity, beliefs and other factors that underlie these perspectives and preferences, and recommendations regarding communications on race/ethnicity. Participants were specifically asked to provide feedback on both the OMB standard race and ethnicity questions as well as an early draft of the Revised Race/Ethnicity Form.

Key recommendations from focus group results:

- Combine Race and Hispanic Ethnicity questions into a singular question.
- Include a Middle Eastern or North African aggregate category.
- Allow for selection of multiple racial/ethnic identities.
- Add a Jewish detailed category.
- Alphabetize both aggregate and detailed race/ethnicity categories.
- Change wording of “Other” category to “Not Listed”.
- Remove word “Origin“ from “Hispanic, Latino, or Spanish Origin”.
- Add a “I do not identify as a more specific racial/ethnic group”.
- Add a “Don’t Know/Unsure” option.

Local experts: Initial drafts of the Revised Race/Ethnicity Form were iteratively circulated among community advocates, physicians, researchers, health systems administrators and IT, state department of health officials, and community members from the respective racial/ethnic groups. Specifically, experts were asked to review and supplement the detailed race/ethnicity categories based on their knowledge of the diversity of the New York metropolitan area.

A total of 24 additional detailed categories were recommended, including African American, Azerbaijani, Bhutanese, British Virgin Islands, Burmese, Chechen, Circassian, Dagestani, Dominican, Fijian, Georgian, Guyanese, Ingush, Jamaican, Jewish, Karakalpak, Kazakh, Kyrgyz, Lenape, Malian, Marshallese, Mongolian, Nepali, Panamanian, Russian, Salvadoran, Senegalese, Sri Lankan, Tajik, Tatar, Thai, Tibetan, and Trinidadian.

In addition to recommending expanded categories, several individuals advocated for the inclusion of certain detailed race/ethnicity categories in multiple aggregate race/ethnicity category. Specific

to the Asian aggregate category, several West Indian detailed race/ethnicities were suggested to be incorporated, including: British Virgin Islands, Guyanese, Jamaican, and Trinidadian.

Using the Revised Race/Ethnicity Form

Figure 3 below shows the Revised Race/Ethnicity Form's detailed categories for the Asian group. The full form with all categories can be found in the Appendix.

Supplemental detailed categories recommended by local experts are in red and can be incorporated at the researcher's discretion with consideration of what population(s) their project will be targeting.

When using the race/ethnicity form, participants should self-identify their racial/ethnic identity. Researchers or administrative staff should not attempt to guess or assign an individual's race/ethnicity based on their appearance, name, or any other factors during data collection.

Figure 3. Revised Race/Ethnicity Form for Use in Metropolitan New York Area – Asian categories only

What is your race/ethnicity? Please check all that apply.

- ☐ American Indian, Native, First Nations, Indigenous Peoples of the Americas, or Alaska Native
- ☐ Asian
- ☐ Black
- ☐ Hispanic, Latino, Spanish
- ☐ Middle Eastern or North African
- ☐ Native Hawaiian or Pacific Islander
- ☐ White
- ☐ Not listed (please specify): _____
- ☐ Don't know
- ☐ Prefer not to answer

[BRANCHING] If Asian was selected, do you identify as one or more of the specific racial/ethnic groups below?

- ☐ Bangladeshi
- ☐ Bhutanese
- ☐ British Virgin Islands
- ☐ Burmese
- ☐ Chinese
- ☐ Filipino
- ☐ Guyanese
- ☐ Indian (Asian)
- ☐ Jamaican
- ☐ Japanese
- ☐ Karakalpak
- ☐ Kazakh
- ☐ Korean
- ☐ Kyrgyz
- ☐ Mongolian
- ☐ Nepali
- ☐ Pakistani
- ☐ Sri Lankan
- ☐ Taiwanese
- ☐ Tajik
- ☐ Tatar
- ☐ Thai
- ☐ Tibetan
- ☐ Trinidadian
- ☐ Turkmen
- ☐ Uyghur
- ☐ Uzbek
- ☐ Vietnamese
- ☐ Not Listed _____
- ☐ I do not identify as a more specific Asian racial/ethnic group
- ☐ Don't know

Adapting to your setting

The Revised Race/Ethnicity Form was developed to be tailored to the New York metropolitan area; however, the same development process can be used to adapt a similar form for other target geographies in the U.S.

Key Recommendations for collecting disaggregated race/ethnicity data using the form:

- Format
  - Avoid segmenting race and Hispanic ethnicity into separate questions.
  - Allow researchers to provide both aggregate race/ethnicity and detailed race/ethnicity answer options.
  - Include 'Middle Eastern or North African' and 'Hispanic, Latino, or Spanish' as aggregate race/ethnicity options.
  - Allow participants to select multiple race/ethnicity identities.
  - Alphabetize both aggregate and detailed race/ethnicity categories.
  - Use wording "Not Listed" instead of "Other" to write-in additional race/ethnicity information.
  - Add a "Don't Know/Unsure" option.
- Detailed Categories
  - Use the latest available 5-year estimates ACS data to find the top 10 ancestries for each of the seven aggregate race/ethnicity categories (AI/AN, Asian, Black, Hispanic/Latino, MENA, NH/PI, White) in your geography of interest.
    - Note: The number of detailed categories listed is at the researcher's discretion, balancing consideration between racial/ethnic data disaggregation and practical limitations for the length of their form.
  - Seek feedback from local experts and community members regarding what categories are not being included and appropriate terminology used to describe categories; incorporate at the researcher's discretion with consideration to what populations are most likely to be engaged from their project.
  - Provide specific guidance/verbiage for why certain groups are listed.

[a] The terms Latinx and Latine, which are gender-neutral alternatives to "Latino" or "Latina", are becoming increasingly accepted, especially among younger generations. In 2019, the Pew Research Center found that "only 23% of U.S. adults who self-identify as Hispanic or Latino have heard of the term Latinx, and just 3% say they use it to describe themselves."[18] Thus for community-facing materials, we have opted to use the term "Latino."

2.2 Conducting Data Collection

2.2.1 Who is the population from whom you're collecting data?

For primary data collection efforts, it is important to understand who the study population of focus is in order to determine and tailor the study design. The table below provides several examples of factors to consider, specific to Asian American communities.

Factor to Consider	Specific Considerations

Demographic characteristics
 	
Examples
- Age
- Income level
- Educational attainment
	• Younger age groups
	  - May prefer or be the only groups reached by online survey administration;
	  - May require fewer language considerations (e.g., more likely to complete survey in English);
	  - Are more likely to be multiracial or Mixed; therefore, the race/ethnicity question should include Asian and other racial/ethnic groups.
	  - Are more likely to be born in the U.S., compared to older Asian Americans; therefore, questions related to acculturation and immigration could cause confusion and may not be applicable.
	• Older age groups
	  - May prefer or only be reached via paper surveys or require support completing online surveys;
	  - Are more likely to not speak English well and may require more involved 1:1 interpretation/translation support (e.g., interviewer vs. self-administered surveys); and
	  - May require recruitment by trusted community- and faith-based organizations and senior centers.
	• The income and education levels of your sample are important to understand.
	  - Those at lower income and education levels may have more language barriers.
	  - These measures provide important contextual information about whom you are actually characterizing in your sample/with your findings. Asian American communities have the largest wealth gap compared to other racial groups (e.g., income discrepancy by Asian ethnic group). Individuals captured by survey efforts are usually higher income.

Factor to Consider	Specific Considerations

Acculturation and immigration characteristics
 	
Examples
- Nativity
- Country of birth
- Citizenship status
- Generational status living in the U.S.
- Years lived in the U.S. English fluency
- Main language spoken at home
	• Many Asian American older adults were not born in the U.S. Capturing nativity, country of birth, generation, or years lived in the U.S. can provide additional context to the survey participants and to whom who the results are generalizable.
	• Capturing country of birth help to enumerate missing race/ethnicity data.
	• Information about main language spoken at home and English fluency can inform language and translation needs for future products and interventions.
	  - Many national, state, and local surveys that collect Asian racial/ethnic data are conducted in English only. English-only surveys exclude those with limited English proficiency who are likely of lower socioeconomic status, are older or have limited health literacy; this is an important consideration for the generalizability of the research findings.[19]
	  - Some considerations for offering your survey in Asian languages: who is the population of interest; what specific Asian communities do you have reach in; do you have community partners (or related experts) who can advise on what Asian survey languages are needed; and is there a budget and resources for translation and bilingual data collection support, as needed. Translations by vendors should be reviewed for accuracy and extra time provided for translations and IRB approval.
	• These questions are sensitive and may require additional training to mitigate participant concerns and justification in the survey explaining how the data will be used and addressing privacy concerns. Many Asian and other immigrant communities may harbor institutional distrust for sharing data due to political reasons including:
	  - Post 9/11 surveillance of South Asian, Muslim, Sikh, West Asian, and Arab/MENA populations;
	  - Increasing normalization of anti-China sentiment in U.S. and globally related to COVID-19 and geopolitical tensions among Chinese and other East Asian populations;
	  - Anxieties about lack of access to educational opportunity (race-based affirmative action policy) among Chinese populations;
	  - Undocumented status across all immigrant groups; and
	  - Residual trauma and distrust among certain groups coming from countries where government data collection has contributed to personal experiences with discrimination.

Factor to Consider	Specific Considerations

Geography
 	
Examples
- Demographics of the local Asian American population – know the groups of your location and setting
- Emerging populations
	• Knowing the demographic makeup of the local or geographic population can inform which racial and ethnic categories to include (i.e., the Asian American ethnic groups in California are different than in New York).
	  - There may be subgroups of Asian Americans that may be similar based on religion, ethnicity, culture, language, or related characteristics (e.g., indigenous groups).
	  - If these subpopulations are present in your geographic region and a main population of interest, it is important to include these ethnic categories in your race/ethnicity questions. For race/ethnicity categories that may be uncommon, consider noting this to trained interviewers so they know it is an answer choice, or provide a preview of the race/ethnicity categories to respondents so that the appropriate ethnic group is selected.
	• The U.S. Census and ACS can be used for information on the makeup of Asian American ethnic groups in a specific geographic region and specific neighborhoods or locations where Asian American communities reside for targeted recruitment efforts.
	• Who are the local community partners serving these communities who are potential paid collaborators on the project?

2.2.2 What is your process for data collection?

Other factors to consider for primary data collection are on survey administration and recruitment. The table illustrates a few considerations on the mode of survey administration and recruitment mechanisms.

Factor to Consider	Specific Considerations
Survey Administration Examples - Online survey administration - Interviewer administered by a trained interviewer (e.g., community health worker) - By participant on paper - Over the phone - By mail	• For online survey administration, branching logic can hide detailed race/ethnicity categories for aggregate race/ethnicity groups that were not selected. However, surveys only administrated online would systematically exclude participants who are not comfortable with or have access with technology (e.g., digital literacy, older adults) or are unable to read/write. • For paper survey administration, print outs may require additional staff support because the printed survey will include all of the branching questions. Consider how many questions are included to not overwhelm the survey respondent or interviewers. • Additional training may be needed for trained staff and interviewers to learn about the survey and flow of questions, have a consistent protocol to administer the survey, and prevent missing or incomplete survey responses. - Phone and mail survey administration for certain Asian American populations will require translated surveys and phone calls in their preferred language during survey recruitment and administration.

Factor to Consider	Specific Considerations
Projected survey recruitment Examples - Over the phone - By mail - Paper vs electronic adminsitration - Recruitment target	• Paper survey administration may allow for recruitment of more diverse populations with lower tech literacy (i.e., older adult populations) but will require additional time reserved for data entry. • The target recruitment numbers should consider the minimum sample size. Depending on the study purpose, the minimum sample size may be lower if descriptive data summaries; larger sample sizes would be required if more complex analyses are the end goal. • If more than one Asian racial/ethnic group is included, efforts should be made during recruitment and administration for a meaningful sample size for analyses and interpretations. • If the projected survey recruitment is small and/or interviewer administered, questionnaires can consider more open-ended and write-in questions. • For larger recruitment numbers, coding/data cleaning or analysis of open-ended and write-in questions may be more time consuming and difficult depending on the number of responses received.

Example additional survey question wording and considerations for Asian immigrant populations include:

- Were you born in the United States (U.S.)? (Yes, No); If no, what country were you born in? (Write-in response option);
- The language of survey administration;
- Do you speak a language other than English at home? (Yes, No); Write in language (Write-in response option) or language categories (depending on racial/ethnic group(s);
- How well do you speak English? (Very well, Well, Not well, Not at all);
- What is the total number of years you have lived in the United States (U.S.)? Write-in response option or categories (e.g., Less than one year, 1-3 years, 4-6 years, 7-9 years, 10+ years);
- What is your highest grade of schooling you completed? (Less than 8 years, 8-11 years, 12 years or completed high school, post high school other than college, some college, college graduate, post graduate);
- Do you currently rent or own your home? (Own, rent, occupied without paying monetary rent); and
- Thinking about members of your family living in your household, what is your combined annual income, meaning the total pretax income from all sources earned in the past year? ($0-$9,999, $10,000 to $14,999, $15,000 to $19,999, $20,000 to $34,999, $35,000 to $49,999, $50,000 to $74,999, $75,000 to $99,999, $100,000 to $199,999, $200,000 or more).

2.3 Case Studies

2.3.1 Case Study 1: Adapting Race/Ethnicity Data Collection for NYU Langone Health

In 2020, the NYU Langone Health hospital system began a concerted effort to update its processes for collecting Race, Ethnicity, and Language (REaL) data. CSAAH worked closely with hospital leadership to incorporate the Revised Race/Ethnicity Form approach and categories to meet institutional needs, navigating challenges unique to patient data collection in a medical system.

The first challenge to be addressed was the limited user experience design flexibility available in the hospital's data collection infrastructure. At NYU Langone Health, predetermined data input options are set by EPIC – the hospital's electronic medical record software. Unlike through a survey form, this system would not allow the use of branching logic to display detailed race/ethnicity categories based on the patient's initial aggregate race/ethnicity selection(s).

An initial draft design had patient's select an aggregate race/ethnicity category and then browse through a two-page list of detailed race/ethnicity options (93 categories total). Exploring this design further, hospital IT leadership found this would create challenges in ensuring previous patient race/ethnicity entry values – which followed a similar OMB-style design – could be pulled forward into the new system.

After reducing detailed categories to fit on a single page display (67 categories total), further feedback from administrative leadership raised concerns that too many options would increase patient registration time and overwhelm patients. To address this, the detailed category list was further reduced to include only the top 5 ancestry categories for each aggregate race/ethnicity category in the New York metro area plus 'African American,' 'Jewish,' and 'Lenape' categories based on local expert feedback (38 categories total).

Other considerations for adapting and updating the hospital's race/ethnicity data collection infrastructure included developing a communications plan to ensure both patients and staff understood why this data was being collected. Banner messaging at the top of each patient's online portal form was developed to link patients to the Q&A page, explaining why this data is important. Updated staff trainings will be developed for providers and patient-facing staff, detailing the ethos behind collecting race/ethnicity data and how to answer patient questions around why it is being collected.

2.3.2 Case Study 2: Cancer Community Health Resources and Needs Assessment (Cancer CHRNA)

In early 2020, the Community Outreach and Engagement Core (COE) at the NYU Perlmutter Cancer Center (PCC) and the NYU Section for Health Equity at the NYU Grossman School of Medicine, together with the Brooklyn Community Action Network, initiated the Cancer Community Resources and Needs Assessment Survey (Cancer CHRNA). The primary purpose of the Cancer CHRNA is to assess community-level health improvements and priorities for adults within the PCC catchment area with a focus on racial/ethnic minority and immigrant populations. The goal of the Cancer CHRNA survey is three-fold: 1) determine existing health issues for NYC residents in the PCC catchment area regarding cancer prevention and disparities; 2) assess resources available to

adults within the PCC catchment area with a focus on racial/ethnic minority and immigrant populations; and 3) identify best approaches to meet the needs of this population. Community Health Workers (CHWs) with extensive experience in building community partnerships and deep cultural knowledge engaged with clinical and community organizations to enhance data collection efforts.

Surveys were translated into 9 languages: Arabic, Bangla, Simplified Chinese, Traditional Chinese, Haitian Creole, Korean, Spanish, Russian, and Urdu. Race and ethnicity questions were adapted from the Revised Race/Ethnicity Form as follows: 1) Hispanic, Latino, or Spanish origin[b] and Middle Eastern or North African options were added as additional race/ethnicity categories; and 2) a don't know/not sure option. The additional detailed race categories only included groups relevant to our research and geography of interest. For example, under the detailed Asian origin or ancestry questions, we included Asian Indian, Chinese, Bangladeshi, Guyanese, Filipino, Japanese, Korean, Pakistani, and Vietnamese as options. Other considerations for this project: CHWs and survey staff were encouraged to engage under-represented and hard-to-reach communities through community and clinical partners for online, in-person or telephone-based data collection, and CHWs were trained to list all the options for race/ethnicity and detailed ancestry questions when asking participants to fill out the form.

[b]Earlier versions of the Revised Form included 'Hispanic, Latino, Spanish Origin' as a racial/ethnic category. Per focus group feedback, the category has been updated to no longer include the word 'origin'.

3. DATA ANALYSIS

3.1 Constructing Race/Ethnicity Variables

How racial/ethnic group variables are constructed will depend on how the question has been asked on the survey. When possible, recommendations for constructing race/ethnicity variables include:

- **Create a mutually exclusive aggregate race/ethnicity variable** for the seven aggregate race/ethnicity categories (American Indian/Alaska Native, Asian, Black or African American, Hispanic/Latino, Middle Eastern or North African, Native Hawaiian or Pacific Islander, and White) plus an eighth 'Multiracial' category for individuals who identify as more than one aggregate race/ethnicity group (i.e., White/Black, Hispanic/Asian, Middle Eastern or North African/White, etc.)
- **Create a mutually exclusive detailed race/ethnicity variable for all disaggregated race/ethnicity groups.**
    - For individuals who report multiple racial/ethnic identities within the same aggregate race/ethnicity group recode as '[their aggregate race/ethnicity + Mixed]' (i.e., Chinese/Vietnamese → 'Asian Mixed,' Italian/Irish →'White Mixed').
    - For detailed race/ethnicity categories that are too small, group them into an Other category corresponding to their aggregate race/ethnicity (i.e., Asian Other, Native Hawaiian or Pacific Islander Other, Middle Eastern or North African Other etc.). The cut off for 'too small' can be determined at the researcher's discretion but should be reported in any methods description. The groups represented in the Other category should be listed in the Methods and in tables/graphs (e.g., Asian Other includes individuals of Thai, Chinese, Korean descent).
    - Individuals who are categorized as 'Multiracial' for the aggregate race/ethnicity variable are also recoded as 'Multiracial' for the detailed race/ethnicity variable.
- **If appropriate, for Asian populations, a regional race/ethnicity variable can be created, grouping individuals into regional Asian categories.** Regional categories include East Asian, South Asian, Southeast Asian, and Central Asian. Geographic region definitions from the United Nations Statistical Division provide a framework for how Asian race/ethnicity categories can be grouped.
- **If appropriate, create a non-mutually exclusive race/ethnicity variable for each unique aggregate and detailed race/ethnicity category.** This is especially valuable when analyzing Multiracial or Asian Mixed populations. This step may not be necessary when race/ethnicity is collected via survey with a check-all-that-apply design.
- **Further granular variables for the Multiracial/Mixed populations can be created if needed.** Two potential options for creating more granular Multiracial/Mixed variables are:
    - Create multiple new variables for each unique race/ethnicity category in combination with other racial/ethnic groups, listed as '[race/ethnicity] in combination with other racial/ethnic groups' (i.e., Chinese in combination with other racial/ethnic groups, Filipino in combination with other racial/ethnic groups, etc.); or

- Create a single detailed Multiracial/Mixed variable that describes each unique combination of Multiracial/Mixed detailed race/ethnicity categories (i.e., Chinese-Italian, Filipino-Vietnamese, etc.).

Figure 4 illustrates an example of how these variables could be created for a hypothetical Asian American study sample.

Figure 4. Constructing Race/Ethnicity Variable Example

Mutually Exclusive Categories

Original Race/Ethnicity Value	Analysis Recode	Aggregate Race/Ethnicity	Detailed Race/Ethnicity	Asian Regional Race/Ethnicity
Asian/Chinese		Asian	Chinese	East Asian
Asian/Chinese/ Vietnamese		Asian	Asian Mixed	N/A
Asian/White/ Chinese/Italian		Multiracial	Multiracial	N/A
Asian/Bangladeshi		Asian	Bangladeshi	South Asian
Asian/Nepali		Asian	Nepali	South Asian
Asian/Black/ Chinese/Jamaican		Multiracial	Multiracial	N/A
Asian/Filipino		Asian	Filipino	South Asian
Asian/Mongolian*		Asian	Other*	East Asian
…		…	…	…

*Assumes Mongolian sample qualified too small for reporting.

Non-Mutually Exclusive Categories

Original Race/Ethnicity Value	Analysis Recode	Asian	Black	White	Chinese	Bangladeshi	…
Asian/Chinese		Asian	Not Black	Not White	Chinese	Not Bangladeshi	…
Asian/Chinese/ Vietnamese		Asian	Not Black	Not White	Chinese	Not Bangladeshi	…
Asian/White/ Chinese/Italian		Asian	Not Black	White	Chinese	Not Bangladeshi	…
Asian/Bangladeshi		Asian	Not Black	Not White	Not Chinese	Bangladeshi	…
Asian/Nepali		Asian	Not Black	Not White	Not Chinese	Not Bangladeshi	…
Asian/Black/ Chinese/Jamaican		Asian	Black	Not White	Chinese	Not Bangladeshi	…
Asian/Filipino		Asian	Not Black	Not White	Not Chinese	Not Bangladeshi	…
Asian/Mongolian		Asian	Not Black	Not White	Not Chinese	Not Bangladeshi	
…		…	…	…	…	…	…

3.2 Secondary Data Analysis

Most epidemiologic cohort studies have not included Asian Americans, and data containing disaggregated Asian American racial/ethnic groups is more likely to come from national health surveys.[27] It is important to consider how Asian American detailed race/ethnicity data has been collected on these surveys – disaggregated Asian racial/ethnic group data may be limited or nonexistent. Additionally, many surveys do not include Asian languages, and sampling bias may result from higher rates of English proficient Asian Americans.

Sensitivity analyses should be conducted for more precise evaluation of the heterogeneity of the Asian American sample. Careful consideration of who was included or not included in the data being analyzed will help researchers evaluate biases in data collection and recruitment and the generalizability of the analyses to Asian American populations. Sensitivity analyses, at a minimum, include understanding the distribution of income and educational status of participants by detailed Asian racial/ethnic group. We want to reiterate the importance of understanding who is in your sample (who was excluded from the research and data collection because of language support, the recruitment protocol – e.g., landline sampling, online only) and for whom the results of the analyses will or will not be generalizable. This is particularly important for Asian Americans, whose health outcomes might vary by social and demographic characteristics including geographic setting, immigration history, income, educational status, health literacy, and English fluency.[36]

Considerations to evaluate the heterogeneity of the Asian American sample:

- How were race and ethnicity asked?
- What languages was the survey offered in? What languages did the participants actually take the survey in?
- Were there questions about English language fluency or main language spoken at home?
- For electronic health records, is there a question for preferred language with primary care providers?
- Were there questions about nativity or years lived in the U.S.? Many surveys do not include a question for nativity (e.g., BRFSS, NHIS), and if the question is asked, country of birth is not always recorded.
- What was the distribution of income by the overall Asian American and within Asian American detailed race/ethnicity groups? For example, there might be a bimodal distribution of income within the Asian Indian group[29] that could identify different health outcomes.

For multivariate and multivariable analyses, it is important to understand the context and interpretations of the "effect of race."[38] By this, we consider how are the coefficient estimates being interpreted when race/ethnicity is added in a regression model. The goal of including race and ethnicity as social risk factors is to understand the effects of racism on health.[39]

Sources of datasets with disaggregated Asian American racial/ethnic group data for potential secondary analysis include:

- The U.S. Census;
- The American Community Survey (ACS);
- The Behavioral Risk Factor Surveillance System (BRFSS);
- The California Health Interview Survey (CHIS);
- The Healthy Chicago Survey (HCS);
- The Los Angeles County Health Survey (LACHS);
- The National Health Interview Survey (NHIS);
- The National Health and Nutrition Examination Survey (NHANES);
- The NYC Community Health Survey (NYC CHS);
- The Medical Expenditure Panel Survey (MEPS); and
- The New York City Health and Nutrition Examination Survey (NYC HANES).

Each survey may require different access options which may or may not include data use agreements or fees, include different available Asian American racial/ethnic groups, and be asked in different languages. Further details on many of these datasets and others can be found elsewhere.[32-34]

3.3 Strategies for Incomplete and Missing Data

Incomplete or missing data can take many forms. There may be partial or unit missingness, whereby an individual selectively responds to certain questions but not others, or total missingness, whereby an individual is not selected, unable to participate, or refuses to participate leading to non-response for all questions.

3.3.1 Partial missingness

What is partial missingness?

Partial missingness happens when an individual stops taking the survey before completion, and at least one measurement is missing, or when an individual does not answer some of the survey questions throughout the survey.

What can we do about partial missingness?

Complete case analysis is often used in partial missingness. When missing data account for less than ~5-10% of the data – a common rule of thumb, analyses are conducted on individuals for whom all relevant data are available. While this affects the statistical efficiency of estimates, there is no bias introduced if the data are missing completely at random. Issues arise with missing data when the characteristics of individuals who do not respond differ meaningfully from those who respond. Under these circumstances, measures of frequency and association estimated using complete case analysis will be both inefficient and biased.

Statistical inefficiency and bias are especially concerning when race and ethnicity data are missing. Many data reporting guidelines suggest that certain estimates be suppressed or cautious interpretation of measures of frequency or association due to small disaggregated racial/ethnic group sample sizes. Consequently, individuals in smaller racial/ethnic groups, (e.g., Korean or Vietnamese), are often aggregated into an overall Asian, Asian and Pacific Islander, or Other Race category, or they may be dropped from analyses entirely. In our case study of NHANES, we examine how small sample sizes caused Asian American data to be concealed in publicly available datasets for over a decade.

Analyzing the demographic characteristics of the individuals for whom data are and are not available is important to determine whether the data are missing systematically or at random and provide better insight into next steps for addressing the missing data. Under certain conditions, statistical methods can minimize missingness of race/ethnicity data post hoc and increase sample sizes to allow for better estimation and reduce potential bias.

Imputation is one solution that involves using available data to fill in the missing value to create a completed dataset. An example of informal imputation is using available responses to nativity (e.g., Were you born in the U.S.? Or, what country were you born in?) or language questions (e.g., Do you speak a language other than English at home? What is that language?) as proxies for missing detailed race/ethnicity data. Individuals reporting being born in China or speaking Chinese at home may be considered Chinese. Surname matching is another simple imputation method that relies on a validated name lists to assign race/ethnicity from participants' last names.[43] Lauderdale and Kestenbaum's (2000) Asian surname list is a commonly used name list developed from Social Security Administration data and validated against U.S. Census data for imputing the six largest detailed Asian race/ethnicity groups (Chinese, Asian Indian, Filipino, Vietnamese, Korean, and Japanese).[43]

Advanced imputation methods can involve algorithms, regression-based techniques, or machine learning models. Advanced methods often leverage individual surname data but also other inputs like sex, gender, place of birth, parental place of birth, or place of residence while using Bayesian probability to predict race/ethnicity. The Bayesian Improved Surname Geocoding (BISG) method is a validated and reputed name algorithm method that calculates the probability of an individual belonging to each of six race/ethnicity categories (AI/AN, Asian Pacific Islander, Black, Hispanic, White, Multiracial) and then summarizes the distribution of the race/ethnicity at the population level.[44] Additional details on these techniques can be found in a systematic review CSAAH conducted on methods for retrospectively predicting race/ethnicity.[45]

3.3.2 Total missingness

What is total missingness?

Total missingness may arise from a survey design issue. For example, offering the survey in English only or including a question about citizenship may lead to underrepresentation of foreign-born and limited English proficient populations. In the scenario of total missingness, in which data for certain individuals or groups are completely unavailable, there are few approaches to recover the true estimates for measures of frequency and association.

What can we do about total missingness?

As with partial missingness, an important first step is to analyze the demographic characteristics of the included groups to determine whether the study sample reflects the underlying target population. If a study reports that 100% of the included Asian American population is English proficient, then a portion of the Asian American population is missing. Data from the U.S. Census indicate 72% of the Asian population in the U.S. is English proficient, with important differences by ethnicity and nativity.[46] Estimates of measures of frequency and association derived in this study population may be internally valid but not be generalizable to the overall Asian American population in the U.S.

This problem has been previously defined as a failure to generalize, whereby the distribution of the outcome is expected to be different in the selected and non-selected populations with the variable influencing selection also influences the outcome. If language is associated with being selected into the study, e.g., the survey is only offered in English, and language is also associated with the outcome of interest, e.g., food security, then the estimate of food security prevalence among the Asian American population in the study cannot be generalized to represent the food security prevalence of Asian Americans in the U.S. overall.

Unlike with partial missingness, where there are a number of validated methods to impute data, methodology available for total missingness is limited. Most research for addressing selection issues are catered to the partial missingness scenario and focus on investigating bias in measures of association.[47] Potential opportunities for addressing this problem statistically require strict assumptions and depend on other data availability but may allow investigators to better approximate the true estimate.[48]

One option is to use standardization to identify the measure of frequency or association that we would expect in the total population. Standardization involves taking stratum-specific estimates and weighting them to the distribution of the covariate defining the stratum in the total population. Building on the example from above, to get an approximation of the true population estimate for Asian Americans, the estimate from the English-speaking study sample would be weighted by the proportion of Asian Americans who speak English in the U.S., and then an estimate of the non-English speaking population would be weighted by the proportion of Asian Americans who do not speak English in the U.S. The latter would require finding a valid estimate of the measure of interest within a non-English speaking population. This external data on the prevalence of the outcome in the non-selected population may be borrowed from other studies, with the assumption that the selected and non-selected populations are coming from the same source population. In the absence of external data, simulating plausible estimates based on published literature and expert knowledge may allow researchers to understand how influential selection is on measures of frequency or association for certain populations.

Populations may also be missing from denominators used for public health surveillance. Many health statistics rely on the U.S. Census counts for populations as the denominator, with recent examples including COVID-19 hospitalization rates and vaccination estimates. The U.S. Census has historically undercounted immigrant and racial/ethnic minority populations. Due to the potential citizenship question and the COVID-19 pandemic, research suggests that the 2020 U.S. Census had a greater undercount than prior years. Correcting for the undercount is possible but requires data on location- and demographic-specific birth and death rates as well as migration

patterns, making it impracticable for most researchers. In our case study of the NYC COVID-19 Vaccination statistics, we examine the implications of an undercount for Asian American and NH/PI populations.

3.4 Case Studies Related to Data Analysis

3.4.1 Case Study 1: NHANES Race/Ethnicity Variables

Prior to 2011, the only race/ethnicity variable publicly available in the NHANES dataset was RIDRETH1, with the following categories: Mexican American, Other Hispanic, Non-Hispanic White, Non-Hispanic Black, and Other Race including Multiracial. Asian Americans, American Indian or Alaska Native, and Native Hawaiian or Pacific Islander were grouped into the 'Other Race including Multiracial' category in the RIDRETH1 variable due to small sample sizes and statistical efficiency concerns. Consequently, Asian Americans, American Indian or Alaska Native, as well as any further disaggregated racial/ethnic groups, remain absent from analyses examining racial/ethnic disparities that use NHANES data prior to 2011.

Beginning in the 2011-2012 cycle, the NCHS began oversampling Asian Americans, and a new race/ethnicity variable was created for public use, RIDRETH3, which added a separate category for Non-Hispanic Asian. Some analyses investigating racial/ethnic disparities that pool data or analyze trends from before and after 2011 still use the RIDRETH1 variable, which remains available in more recent data releases. We recommend future analyses utilizing the NHANES public data use the RIDRETH3 variable, so that data on Asian Americans can be presented when possible.

A change in survey methods, involving a purposeful oversampling of Asian Americans, enabled the disaggregation of this group from the 'Other' category. When changes in survey methods are either infeasible or impossible, statistical methods such as those mentioned above could be employed to disaggregate data from the 'Other' category or impute missing data to power disaggregated racial/ethnic group analyses.

Of note, the data for Asian American populations in NHANES may not represent the underlying population of Asian Americans in the U.S. NHANES is administered in English and Spanish and approximately a quarter of Asian Americans in the U.S. have limited English proficiency. Thus, researchers should take into consideration when interpreting results.

3.4.2 Case Study 2: Vaccination Uptake in NYC

During the COVID-19 pandemic, the NYC Department of Health and Mental Hygiene (DOHMH) began releasing data on vaccination uptake by race and ethnicity. During the fall of 2021, vaccination data for Asian Americans and NH/PIs began exceeding 100%. The U.S. Census Bureau American Community Survey (ACS) population estimates, which are derived from the most recent U.S. Census, were used as the denominators for the vaccination statistics. As previously mentioned, the U.S. Census has historically undercounted immigrant and racial/ethnic minority populations. An undercount of Asian American and NH/PI populations in NYC may explain the implausible estimates of vaccination uptake in these groups.

Unfortunately, data were unavailable at the disaggregated level for Asian American and NH/PI communities. But from community-based research, we found that COVID-19 vaccination prevalence varied greatly by Asian American and NH/PI disaggregated racial/ethnic group, using needs assessments from NYC, Salt Lake City, Phoenix, and Atlanta; among Asian Americans, high vaccination rates were seen among Japanese (89%), Vietnamese (87%), Filipinos (84%), and Mixed Asians (85%), while low vaccination rates were seen among Nepalis (16%) and Bangladeshis (55%), and among NH/PIs, all groups had low vaccination rates (Marshallese – 12%, Tongan – 37%, Samoan – 46%, and other NH/PI - 35%). COVID-19 vaccine hesitancy also varied by disaggregated racial/ethnic using largely national needs assessment data; among unvaccinated Asian Americans, the highest hesitancy (reporting unsure, somewhat unlikely, or very unlikely to get the COVID-19 vaccine) was seen among Japanese (27%), followed by Koreans (21%), and Mixed Asians (21%), and among unvaccinated NH/PIs, the highest hesitancy was seen among Tongans (69%) and Samoans (67%), followed by Marshallese (58%), other NH/PIs (50.7%), and multiracial NH/PIs (22%). This exemplifies the need for accurate enumeration and inclusion of participants in all datasets, especially those used in public health surveillance efforts and for health policy decision-making.

4. DATA REPORTING

In the following section, we provide recommendations and considerations for reporting Asian American data, organized into three bucket topics: Methods, Results, and Discussion. These three topics serve as a way of categorizing our recommendations; not all Asian American data reporting will need to follow a formal academic reporting structure.

4.1 Methods

Details of study design regarding how data was collected can contextualize the population sample and help researchers better determine next steps for research and resource allocation.

The methods section should include:

- All eligibility and exclusion criteria, including explicit details about the disaggregated Asian American racial/ethnic group sampled (different Asian American racial/ethnic groups would likely be sampled in different studies based on geography);
    - What were the demographics of your sample?
    - Which languages other than English were used in your study or dataset?
    - How many participants were recruited from each disaggregated racial/ethnic group?
    - What recruitment methods were used?
    - Of the participants recruited, how many participants were in the analytic sample and what were the exclusion criteria and why?
- The dates of the study period (what dates did you run the study/were data collected-- mention any secular trends that occurred during the study period, e.g., COVID-19 pandemic);
- The type of study design (what study design was used, e.g., cross-sectional, cohort, randomized-controlled design, what methods were used to collect and analyze data, any strengths and limitations of the design used);
- The number of individuals surveyed/enrolled.

4.2 Results

4.2.1 Demographic Details

Demographic details of the study population are especially relevant for preventing the perpetuation of harmful Asian American stereotypes and describing to whom your data can be generalized.[29]

When available key characteristics to report should include:

- Detailed racial/ethnic information;
- Income;
- Education;
- English language proficiency;
- Language(s) spoken at home; and
- Nativity.

When grouping either aggregate or detailed race/ethnicity groups into an 'Other' category, detailed description of what race/ethnicity categories are included in that grouping should be provided. Frequently Asian American data is subsumed into 'Other' due to small sample sizes without clear reporting of what detailed Asian race/ethnicities were being captured. If any disaggregated racial/ethnic groups were too small to include, explain who they were and why they were excluded.

When combining smaller Asian detailed race/ethnicity categories into regional groups (South Asian, East Asian, Southeast Asian), detail the groups that were included in these new combined subgroups.

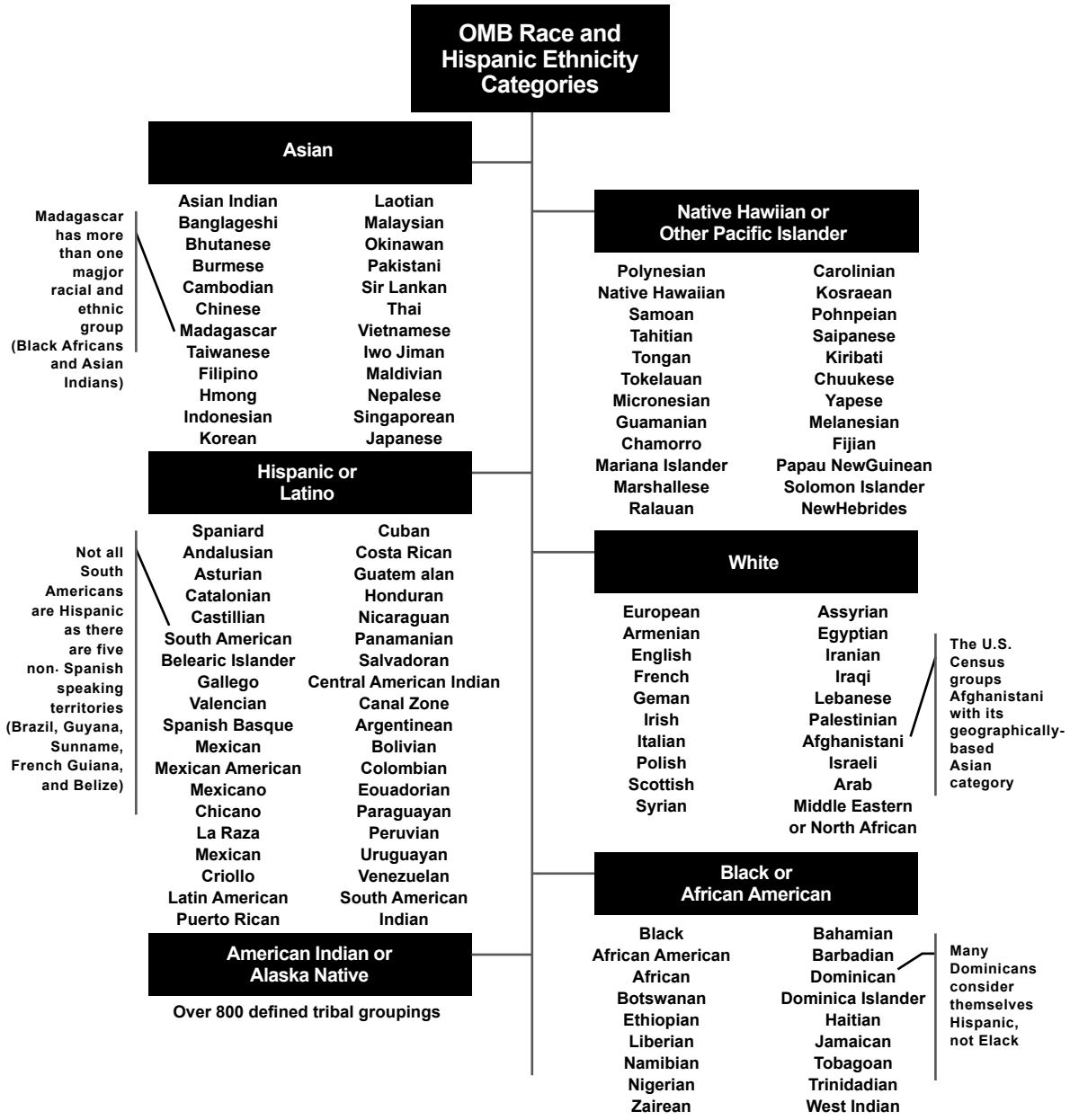
When describing Multiracial or Asian Mixed populations, including descriptions of what combinations of aggregate or detailed race/ethnicities can also be useful for illuminating trends in reported data.

4.2.2 Aggregating Race/Ethnicity Data and Roll-up Reporting

OMB Roll-up

In some circumstances, race and ethnicity reporting must meet OMB guidelines. For example, compliance with OMB standards is required required for federal agencies as well as other state and local organizations who model their reporting guidelines on the OMB's 1997 directive.[49] In these cases, Figure 4 illustrates a roll-up guideline provided by Agency for Healthcare Research and Quality for detailed race/ethnicity categories. Further granular roll-up guidance is can be found in Appendix E, Table-E1of the Institute of Medicine's report Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement.[20]

Figure 5. OMB Race and Hispanic Ethnicity Categories[49]



Further race/ethnicity aggregation roll up guidance can be found in Appendix E, Table E-1 of the Institute of Medicine's report *Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement*.

Middle Eastern or North African

When reporting aggregate race/ethnicity is not required to be limited to the six standard OMB race and ethnicity categories (AI/AN, Asian, Black, Hispanic/Latino, NH/PI, White), we recommend including MENA as a seventh aggregate reporting category.

There is currently no standard for how detailed race/ethnicities roll-up into MENA. At CSAAH we use the United Nations Statistical Division's definition of North African and West Asian countries/regions. These countries/regions include: Algeria, Egypt, Libya, Morocco, Sudan, Tunisia, Armenia,

Azerbaijan, Bahrain, Cyprus, Georgia, Iraq, Israel, Jordan, Kuwait, Lebanon, Oman, Qatar, Saudi Arabia, Palestine, Syria, Turkey, and United Arab Emirates.

Roll-up Grey Areas

The roll-up process still has several grey areas (as noted in Figure 4) emphasizing the potential mismatches between the aggregate racial/ethnic categories and self-identification of detailed race/ethnicity. Examples include but are not limited to:

- Many Dominicans may consider themselves Hispanic and not Black;
- Afghan is classified under White and not Asian (which is in conflict with the U.S. Census);
- Not all South American countries are categorized under the Hispanic OMB category;
- Countries like Madagascar may individuals who self-identify with more than one major racial/ethnic group (e.g., Black Africans and Asian Indians).[50]

We recommend a participant's reported aggregate race/ethnicity should always reflect their selected aggregate race/ethnicity. In situations where an individual self-identifies as an aggregate race/ethnicity that does not align with their detailed race/ethnicity, their reported aggregate race/ethnicity should not be changed. For example, if someone identifies as 'Native Hawaiian or Pacific Islander' and then writes-in their detailed race/ethnicity as 'Filipino', their reported aggregate race/ethnicity should remain as 'Native Hawaiian or Pacific Islander'.

In roll-up situations where only detailed race/ethnicity information is provided, researchers should detail their aggregate race/ethnicity assignment rules for ambiguous categories (i.e., Dominican, Guyanese, etc.)

4.3 Discussion

Below are several recommendations for considerations when discussing results from Asian American data:

Discuss who is represented in the data. As described in the Data Collection section, the Asian American category includes different people in different studies with regard to geography, Asian racial/ethnic group, and socioeconomic status. Presenting data for Asian Americans only and not providing context to the research findings and who results are generalizable to fuel scientific and societal perceptions that Asian Americans do not experience health disparities.[36]

Be cautious when discussing Asian Multiracial and Asian Mixed data. These categories are not monoliths. Data may be generalizable to a specific Asian multiracial group combination but not all multiracial individuals.

Hypothesize how data may look different if data were available for populations absent from their study. Even without formal sensitivity analyses, researchers can theorize the direction of bias based on selection, (i.e., whether the study estimates may be an under- or over-estimate of the true population estimate).

Consider complementing statistics from large administrative databases with findings from community-based research. Doing so can help elucidate health problems for populations that may be hidden when data is only available in aggregate or missing entirely.

Seek and include community feedback. Community-partnered research has been shown to be a successful model to enhance trust, participation, and cultural resonance. Community-based organizations or local partners serving Asian American communities can bring unique perspective to interpreting findings.[36]

If Asian data are not available, cite reasons and discuss or contextualize what implication this may have (e.g., Asian Americans were excluded from the analysis because of inadequate sample size, but an increasing diabetes prevalence in this group has been observed by others). Also include a recommendation for additional research inclusive of Asian American populations.

As mentioned in the considerations for discussion section, contextualizing findings is important when reporting results to avoid perpetuating stereotypes about Asian Americans. Three racialized stereotypes that exacerbate the misrepresentation of Asian American experiences at the data-reporting phase are the model minority, healthy immigrant effect, and perpetual foreigner. These stereotypes represent potential areas for implicit bias that may be embedded in a researcher's own lived experience and should be considered when reporting any findings for Asian data.

Model Minority	Healthy Immigrant Effect	Perpetual Foreigner
The perception of high academic and economic achievement among Asian Americans. This stereotype is rooted in anti-Blackness through the false proximity of Asian Americans to whiteness.[51] The model minority stereotype was used to illustrate that individual underperformance explains racial inequality in American society and to pit Black Americans and Asian Americans against one another to uphold the narrative of white supremacy.[52]	The concept that immigrants display better health than their U.S.-born counterparts, despite overall lower socioeconomic status and limited access to resources. The assumption that duration in the U.S. affects the health of all Asian immigrants in a uniform way is inaccurate and misleading for both research and practice. The healthy immigrant effect also places disproportionate emphasis on individual health behaviors rather than the environmental and political factors that shape health (e.g., role of structural racism and how it can block socioeconomic mobility and integration of immigrants into their new environments).[42]	The stereotype that assumes Asian American are permanently foreign regardless of how long they or their families have lived in the U.S. The perpetual foreigner stereotype has allowed for widespread scapegoating of Asian Americans during disease outbreaks and economic downturn (e.g., yellow peril).[54] It has also manifested in racist and discriminatory perceptions about Muslim and South Asian populations as a threat to national security.

4.4 Case Studies

4.4.1 Case Study 1: NYC COVID-19 Community Health Resources and Needs Assessment

The NYC COVID-19 Community Health Resources and Needs Assessment (NYC COVID-19 CHRNA) examined the impact of the COVID-19 pandemic on the NYC Asian American community (n=1,270).[55] The NYC COVID-19 CHRNA COVID-19 included questions on COVID-19 vaccination, food hardships and access to food, language access, economic impact, housing, access to public benefits, healthcare and healthcare access, anti-Asian hate and discrimination, and support received from community-based organizations. This CHRNA was developed with community partners, and recruitment/completion of the survey was led by the 25+ community partners.

We had an overall sample of 1,353 adults in the NYC COVID-19 CHRNA, with 1,270 adults who self-identified as Asian American. This is a community-based sample and not representative of the NYC Asian American population as a whole. However, we want to emphasize that the NYC COVID-19 CHRNA is differentiated from other NYC population-based surveys and needs assessments because 54% of survey participants took the survey in an Asian language. The survey was offered online in Bangla, Burmese, Simplified and Traditional Chinese, English, Japanese, Korean, Nepali, Punjabi, Tagalog, Urdu, and Vietnamese. Total survey administration by language was: English (46%), Simplified Chinese (31%), Traditional Chinese (4%), Bangla (1%), Japanese (10%), Korean (4%), and Nepali (4%).

Other NYC-based surveys during the COVID-19 pandemic that have mainly been administered in English, Spanish, and occasionally Mandarin. The unintended consequence of conducting surveys in only these languages is overrepresentation of Asian American New Yorkers with higher education and income levels; who tend to have better social and health outcomes. When data collection in Asian American NYC communities is done only in English, social needs stay invisible. These data are a representation of the cultural and linguistic diversity of Asian American communities in NYC that are not captured in existing data.

Below are challenges that arose when reporting results from the NYC COVID-19 CHRNA and how we navigated the issues:

What is the best way to present disaggregated Asian American ethnic group data?

The NYC COVID-19 CHRNA quantitative data were supplemented with community quotes to provide context to the findings and additional detail to what community members were experiencing (that may not have been captured from the CHRNA data). We provided preliminary updates to community partners so that they could: 1) identify discrepancies between the CHRNA data and what they were hearing within the community; 2) better interpret data findings; and 3) have ongoing communication about for whom the NYC COVID-19 CHRNA was capturing information.

Who was or was not represented in the NYC COVID-19 CHRNA data?

There was a large sample of Japanese Americans and small sample of South Asian Americans in the NYC COVID-19 CHRNA; this was largely due to community partner capacity to outreach to their respective communities during the COVID-19 pandemic amid their regular work. The larger proportion of Japanese Americans and smaller proportion of other groups is not reflective of the

Asian American community as a whole. When presenting results from the NYC COVID-19 CHRNA, we were clear about describing which Asian ethnic groups were included and the sample size; how data were not representative what Asian detailed race/ethnicity groups were collapsed into an 'Other Asian' category due to small sample sizes; and any results that should be interpreted with caution.

We also examined demographic characteristics by detailed race/ethnicity group, including English proficiency fluency, education attainment, age group, and neighborhood of residence, to evaluate the diversity of each Asian detailed race/ethnicity group (e.g., Japanese participants were older and had higher education attainment).

Some Asian groups are too small to present disaggregated information for. Should they be included in the results?

All participants who completed the NYC COVID-19 CHRNA were included in the results. When Asian detailed race/ethnicity groups were too small (n<35 people) to present disaggregated results, they were collapsed into an 'Other Asian' category. We described the detailed race/ethnicity groups included in the 'Other Asian' group. These data can inform populations who might require additional outreach for future work and for what populations we should gather more qualitative data.

How should the data be interpreted and what figures should be presented?

We were cautious in presenting disaggregated data from the NYC COVID-19 CHRNA because of the polarized context of the COVID-19 pandemic, and we did not want to shame specific Asian detailed race/ethnicity groups. For example, there was high COVID-19 vaccination among the overall Asian American NYC COVID-19 CHRNA sample, similar to the high vaccination rates on the NYC DOHMH online dashboard during that time period. However, when the NYC COVID-19 CHRNA vaccination data were disaggregated by detailed race/ethnicity group, we found that there was low COVID-19 vaccination but high acceptance among Bangladeshi and Nepali adults (e.g., Bangladeshi and Nepali adults were not hesitant to get the COVID-19 vaccination, so there may be issues with vaccination access).

In our reports and presentations, we presented data for Asian detailed race/ethnicity groups only and did not include data for the overall Asian American sample. We did not want to perpetuate generalizations about high COVID-19 vaccination among Asian Americans that could be used to justify that COVID-19 prevention efforts were not needed in the Asian American community.

4.4.2 Case Study 2: Health Atlas

The Health Atlas presents a visual comparison of Asian American and NH/PI health behaviors, risk factors, and lifestyle factors across national and local datasets, while stratifying by all available disaggregated Asian American and NH/PI racial/ethnic groups.[42] Aggregated years of data were used for each dataset in order to increase smaller cell counts available for single years. Datasets included: The NHIS 2012-2017, the BRFSS in Arizona, Texas, and Utah (2013-2017), and Hawaii (2013-2018), the NYC CHS 2013-2017, the NYC Community Health Resources and Needs Assessments (NYC CHRNA) 2013-2016, the CHIS 2013-2018, the LACHS 2015 & 2018, and the Healthy Chicago Survey (HCS) 2015-2017.

The NHIS, HCS, and BRFSS surveyed in English and Spanish, but no Asian languages. NYC CHS included Chinese, while the NYC CHRNA included all disaggregated racial/ethnic group languages. CHIS and LACHS both surveyed in Mandarin, Cantonese, Korean, and Vietnamese, while CHIS also surveyed in Tagalog.

When health outcomes among Asian Americans from the NHIS, which is surveyed in English and Spanish but no Asian languages, were examined, more positive outcomes were often seen in both overall Asian Americans and the three available disaggregated racial/ethnic groups (Chinese, Asian Indian, and Filipino), when compared to data from regional datasets. For instance, 8% of Asian Americans overall reported a diabetes diagnosis nationally, compared to 12% of Asian Americans in NYC CHS, 10% in CHIS, 10% in Hawaii BRFSS, and 9% in Texas BRFSS. Among Filipinos, 9% of Filipinos reported a diabetes diagnosis nationally, compared to 21% of Filipinos in NYC CHS, 14% in CHIS, 12% in LACHS, 14% in Arizona BRFSS, and 13% in Hawaii BRFSS.

When examining demographics of the available datasets used in the Health Atlas, Asians in NYC had the lowest education levels compared to the other datasets. For instance, among Asian Americans overall in the Health Atlas, the percentage of college graduates was 31.3% in NYC CHS and 40.5% in NYC CHRNA, compared to 17.9% in the NHIS sample, 58.4% in CHIS, 46.4% in LACHS, 59.6% in Arizona BRFSS, 31.0% in Hawaii BRFSS, 64.6% in Texas BRFSS, 41.7% in Utah BRFSS, and 67.2% in the HCS. However, differences were also seen by Asian racial/ethnic group; the percentage of college graduate was 23.7% in NYC CHS and 23.8% in NYC CHRNA, compared to 44.7% in NHIS, 61.9% in CHIS, 45.4% in LACHS, 71.7% in Arizona BRFSS, 42.4% in Hawaii BRFSS, 77.6% in Texas BRFSS, 63.5% in Utah BRFSS, and 61.5% in the HCS. In addition to potential difference by languages offered, Asian Americans also differ in terms of detailed racial/ethnic group and geography, and this should be considered with any data collection efforts. Nativity was asked nationally, as well as in NYC and California; compared to 22.2% of individuals born in the U.S. on the NHIS, 11.1% were U.S.-born in NYC CHS, and 28.6% and 27.2% were born in the U.S. on the CHIS and LACHS, respectively, showing additional geographic differences.

5. CLOSING STATEMENT

Data are a critical starting point to policy, resource allocation, and priority setting. Herein, we have described best practices for the collection, analysis, and reporting of Asian American health data that we believe embodies an anti-racist approach for combatting systemic racism embedded in data infrastructure and research institutions. Together, we can collectively incorporate these recommendations into our day-to-day research and data practice, to help move the dial towards a more accurate representation of the health status of Asian Americans.

6. CENTER FOR THE STUDY OF ASIAN AMERICAN HEALTH

The **Center for the Study of Asian American Health (CSAAH)** is a National Institutes of Health (NIH) National Institute on Minority Health and Health Disparities (NIMHD) funded Specialized Center of Excellence based in the Section for Health Equity within NYU Grossman School of Medicine's Department of Population Health at NYU Langone Health.

Established in 2003 through an NIH NIMHD Project EXPORT (Excellence in Partnership, Outreach, Research, and Training) Center grant, CSAAH is an NIHMD-funded Center of Excellence dedicated to research and evaluation on Asian American health and health disparities. CSAAH's work is guided by a population health equity framework. In close collaboration with over 75 local and national community partners, we have evolved our mission and goals to advance health disparities research within a health equity framework.

CSAAH's guiding principles are as follows:

- We believe in systemic change through multi-pronged strategies and working with diverse stakeholders;
- We believe in equitable collaboration and partnerships;
- We believe in action-oriented research;
- We believe in strengthening the research capacity of both community and academic partners to fully engage in the research process;
- We believe in multi-cultural evaluation.

For more information, please visit us at: https://med.nyu.edu/asian-health.

Find us on Facebook: @NYU.CSAAH and on Twitter: @NYU_CSAAH

7. COALITION FOR ASIAN AMERICAN CHILDREN AND FAMILIES

The **Coalition for Asian American Children and Families (CACF)** is the nation's only pan-Asian children and families' advocacy organization bringing together community-based organizations as well as youth and community allies to fight for equity for Asian Americans and Pacific Islanders (AAPI). CACF was founded in 1986 by a group of concerned social service providers about the City's inability to serve the rapidly growing numbers of AAPI children and families.
CACF is building a community too powerful to ignore. One where the challenges of the AAPI community are understood and addressed. We fight to ensure that the AAPI community has equitable access to education, healthcare, and supportive services they need to thrive. Our vision is that all New York City children and families are safe, healthy, and able to reach their full potential.

CACF LISTENS so we can identify our community's needs.

CACF TRAINS the next generation of leaders for our community.

CACF UNITES our diverse community and our allies.

CACF FIGHTS for equitable access to quality education, services, and funding for our community.

For more information, please visit us at: https://www.cacf.org/

Find us on Facebook: @CACFnyc, Twitter: @cacf, or Instagram: @cacfnyc

8. FUNDING ACKNOWLEDGEMENTS

This manual is a part of the Innovations in Data Equity for All Laboratory (IDEAL) initiative led by the CSAAH and CACF with support by colleagues from the New York Academy of Medicine (NYAM), NYU Langone Health, and the New York State Department of Health. Work was supported by the NIH National Institute on Minority Health and Health Disparities (NIMHD) Award No. U54MD000538 and in part by New York State, the National Heart, Lung, Blood Institute (NHLBI) Community Engagement Alliance (CEAL) Non-Federal 1OT2HL156812–01, Westat Sub-OTA No. 6793–02-S013 and R01HL141427, and Department of Health and Human Services (DHHS) Centers for Disease Control and Prevention (CDC) Award Nos. NU38OT2020001477, CFDA No. 93.421 and 1NH23IP922639–01–00, CFDA No. 93.185.

9. APPENDIX—REVISED RACE ETHNICITY FORM

9.1 Instructions

This document provides a template for collecting disaggregated race/ethnicity information via survey questionnaire tailored to the New York metropolitan area population.

Question format and aggregate categories were derived from APIAHF recommendation. Disaggregated categories were chosen by selecting the top 10 ancestry categories for each aggregate race or ethnicity group according to 2019 ACS - 5-year estimates (retrieved from IPUMS).

Trusted local experts and community partners provided additional disaggregated category recommendations which are highlighted in red. **These additional community suggested categories can be incorporated at the project's discretion.**

For projects that prefer to collect fewer disaggregated categories, the appendix of this document lists, for each aggregate race or ethnicity group, the ranking of the top ten 2019 ACS-5 year ancestry categories by population size. The appendix ranking list may allow projects to choose their own appropriate top category threshold cutoff (ie top 5, top 3, etc.). Categories recommended by local experts and community partners are listed separately.

Disclaimer: These categories are a suggested guide for collecting racial/ethnic data in the New York metropolitan area. Further modifications to these categories is encouraged with consideration for the populations of interest and their respective contexts. Categories should be updated over time.

[c] There are four disaggregated racial/ethnic groups that appear in multiple aggregate categories: British Virgin Islander (Black, Asian), Trinidadian (Black, Asian), Dominican (Black, Hispanic/Latino), Jewish (White, MENA), and Guyanese (Black, Hispanic/Latino, Asian).

9.2 Form

What is your race/ethnicity? Check all that apply.

- ☐ American Indian, Native, First Nations, Indigenous Peoples of the Americas, or Alaska Native
- ☐ Asian
- ☐ Black
- ☐ Hispanic, Latino, Spanish
- ☐ Middle Eastern or North African
- ☐ Native Hawaiian or Pacific Islander
- ☐ White
- ☐ Not listed (please specify): _____
- ☐ Don't know
- ☐ Prefer not to answer

Branching: If selected **American Indian, Native, First Nations, Indigenous Peoples of the Americas, or Alaska Native**, do you identify as one or more of the specific racial/ethnic groups below? Check all that apply.

- ☐ Apache
- ☐ Cherokee
- ☐ Chickasaw
- ☐ Inupiat
- ☐ Iroquois or Haudenosaunee
- ☐ Lenape
- ☐ Mexican American Indian (For example, Mixteco, Nahua, Otomi, Tlapaneco, among others)
- ☐ Navajo
- ☐ Sioux
- ☐ Southern American Indian (For example, Quechua, Kichwa, Shuar, Aymara, among others)
- ☐ Not Listed _____
- ☐ Don't know
- ☐ I do not identify as a more specific American Indian, Native, First Nations, Indigenous Peoples of the Americas, or Alaska Native racial/ethnic group
- ☐ Prefer not to answer

Branching: If selected **Asian**, do you identify as one or more of the specific racial/ethnic groups below? Check all that apply.

- ☐ Bangladeshi
- ☐ Bhutanese
- ☐ British Virgin Islands
- ☐ Burmese
- ☐ Chinese
- ☐ Filipino
- ☐ Guyanese
- ☐ Hong Kong
- ☐ Indian (Asian)
- ☐ Jamaican
- ☐ Japanese
- ☐ Karakalpak
- ☐ Kazakh
- ☐ Korean
- ☐ Kyrgyz
- ☐ Mongolian
- ☐ Nepali
- ☐ Pakistani
- ☐ Sri Lankan
- ☐ Taiwanese
- ☐ Tajik
- ☐ Tatar
- ☐ Thai
- ☐ Tibetan
- ☐ Trinidadian
- ☐ Turkmen
- ☐ Uyghur
- ☐ Uzbek
- ☐ Vietnamese
- ☐ Not Listed _____
- ☐ I do not identify as a more specific Asian racial/ethnic group
- ☐ Don't know
- ☐ Prefer not to answer

Branching: If selected **Black**, do you identify as one or more of the specific racial/ethnic groups below? Check all that apply.

- ☐ African American
- ☐ Anguilla
- ☐ Bajan (Barbadian)
- ☐ British Virgin Islands
- ☐ Dominican
- ☐ Ghanaian
- ☐ Grenadian
- ☐ Guyanese
- ☐ Haitian
- ☐ Jamaican
- ☐ Malian
- ☐ Nigerian
- ☐ Senegalese
- ☐ Trinidadian
- ☐ Not Listed _____
- ☐ I do not identify as a more specific Black racial/ethnic group
- ☐ Don't know;
- ☐ Prefer not to answer

Branching: If selected **Hispanic, Latino, or Spanish**, do you identify as one or more of the specific racial/ethnic groups below? Check all that apply.

- ☐ Colombian
- ☐ Cuban
- ☐ Dominican
- ☐ Ecuadorian
- ☐ Guatemalan
- ☐ Guyanese
- ☐ Honduran
- ☐ Mexican
- ☐ Panamanian
- ☐ Peruvian
- ☐ Puerto Rican
- ☐ Salvadoran
- ☐ Uruguayan
- ☐ Venezuelan
- ☐ Not Listed _____
- ☐ I do not identify as a more specific Hispanic, Latino, or Spanish racial/ethnic group
- ☐ Don't know;
- ☐ Prefer not to answer

Branching: If selected **Middle Eastern or North African**, do you identify as one or more of the specific racial/ethnic groups below? Check all that apply.

- ☐ Armenian
- ☐ Egyptian
- ☐ Georgian
- ☐ Iranian
- ☐ Israeli
- ☐ Jewish
- ☐ Lebanese
- ☐ Moroccan
- ☐ Syrian
- ☐ Turkish
- ☐ Yemen
- ☐ Not Listed _____
- ☐ I do not identify as a more specific Middle Eastern or North African racial/ethnic group
- ☐ Don't Know
- ☐ Prefer not to answer

Branching: If selected **Native Hawaiian or Pacific Islander**, do you identify as one or more of the specific racial/ethnic groups below? Check all that apply.

- ☐ Chamorro/Guamanian
- ☐ Fijian
- ☐ Marshallese
- ☐ Native Hawaiian
- ☐ Samoan
- ☐ Tongan
- ☐ Not Listed _____
- ☐ I do not identify as a more specific Native Hawaiian or Pacific Islander racial/ethnic group
- ☐ Don't know
- ☐ Prefer not to answer

Branching: If selected **White**, do you identify as one or more of the specific racial/ethnic groups below? Check all that apply.

- ☐ English
- ☐ French
- ☐ German
- ☐ Greek
- ☐ Hungarian
- ☐ Irish
- ☐ Italian
- ☐ Jewish
- ☐ Polish
- ☐ Romanian
- ☐ Ukrainian
- ☐ Not Listed _____
- ☐ I do not identify as a more specific White racial/ethnic group
- ☐ Don't know;
- ☐ Prefer not to answer

9.3 Categories by Population Size and Recommended Categories

American Indian, Native, First Nations, Indigenous Peoples of the Americas, or Alaska Native

	TOP 10 ANCESTRY CATEGORIES 2019-ACS 5 YEAR BY POPULATION SIZE (LARGEST TO SMALLEST)	COMMUNITY AND LOCAL EXPERT RECOMMENDED CATEGOREIS
1	Southern American Indian (For example, Quechua, Kichwa, Shuar, Aymara, among others)	Lenape
2	Mexican American Indian (For example, Mixteco, Nahua, Otomi, Tlapaneco, among others)	
3	Iroquois or Haudenosaunee	
4	Cherokee	
5	Apache	
6	Sioux	
7	Navajo	
8	Inupiat	
9	Chickasaw	
10	Blackfoot	

Asian

	TOP 10 ANCESTRY CATEGORIES 2019-ACS 5 YEAR BY POPULATION SIZE (LARGEST TO SMALLEST)	COMMUNITY AND LOCAL EXPERT RECOMMENDED CATEGOREIS
1	Chinese	Bhutanese
2	Indian (Asian)	British Virgin Islands
3	Korean	Burmese
4	Bangladeshi	Guyanese
5	Filipino	Hong Kong
6	Pakistani	Jamaican
7	Japanese	Karakalpak
8	Vietnamese	Kazakh
9	Taiwanese	Kyrgyz
10	Uzbek	Mongolian
		Nepali
		Sri Lankan
		Tajik
		Tatar
		Thai
		Tibetan
		Trinidadian
		Turkmen
		Uyghur

Black

	TOP 10 ANCESTRY CATEGORIES 2019-ACS 5 YEAR BY POPULATION SIZE (LARGEST TO SMALLEST)	COMMUNITY AND LOCAL EXPERT RECOMMENDED CATEGOREIS
1	Jamaican	African American
2	Haitian	Anguilla
3	Guyanese	Bajan (Barbadian)
4	Trinidadian	British Virgin Islands
5	Nigerian	
6	Ghanaian	
7	Bajan (Barbadian)	
8	Grenadian	
9	Anguilla	
10	British Virgin Island	

Hispanic, Latino, Spanish

	TOP 10 ANCESTRY CATEGORIES 2019-ACS 5 YEAR BY POPULATION SIZE (LARGEST TO SMALLEST)	COMMUNITY AND LOCAL EXPERT RECOMMENDED CATEGOREIS
1	Puerto Rican	Guyanese
2	Dominican	Panamanian
3	Mexican	Peruvian
4	Ecuadorian	Puerto Rican
5	Colombian	
6	Salvadoran	
7	Peruvian	
8	Cuban	
9	Honduran	
10	Guatemalan	

Native Hawaiian or Pacific Islander

	TOP 10 ANCESTRY CATEGORIES 2019-ACS 5 YEAR BY POPULATION SIZE (LARGEST TO SMALLEST)	COMMUNITY AND LOCAL EXPERT RECOMMENDED CATEGOREIS
1	Native Hawaiian	Fijian
2	Chamorro/Guamanian	Marshallese
3	Samoan	
4	Tongan	

[d] Only four disaggregated categories listed

Middle Eastern or North African

	TOP 10 ANCESTRY CATEGORIES 2019-ACS 5 YEAR BY POPULATION SIZE (LARGEST TO SMALLEST)	COMMUNITY AND LOCAL EXPERT RECOMMENDED CATEGOREIS
1	Egyptian	Jewish
2	Turkish	Palestinian
3	Israeli	
4	Iranian	
5	Armenian	
6	Lebanese	
7	Syrian	
8	Moroccan	
9	Yemen	
10	Georgian	

White

	TOP 10 ANCESTRY CATEGORIES 2019-ACS 5 YEAR BY POPULATION SIZE (LARGEST TO SMALLEST)	COMMUNITY AND LOCAL EXPERT RECOMMENDED CATEGOREIS
1	Italian	Azerbaijani
2	Irish	Chechen
3	German	Circassian (including Adyghe, Kabardian, and others)
4	Polish	Dagestani (including Avar, Dargin, and others)
5	Romanian	Georgian
6	English	Jewish
7	Greek	Ingush
8	Hungarian	Russian
9	Ukrainian	
10	French	

[e] MENA categories derived from United Nations Statistical Division definition of North African and West Asian countries/regions

10. REFERENCES

1. Asian & Pacific Islander American Health Forum. Policy Recommendations: Health Equity Cannot Be Achieved Without Complete and Transparent Data Collection and the Disaggregation of Data. 2021. Available from: https://www.apiahf.org/wp-content/uploads/2021/02/APIAHF-Policy-Recommendationas-Health-Equity.pdf. Accessed December 15, 2021.
2. Urban Insitute. Many Asian American and Native Hawaiian/Pacific Islander Adults May Face Health Care Access Challenges Related to Limited English Proficiency. 2022 Available from: https://www.urban.org/research/publication/many-asian-american-and-native-hawaiianpacific-islander-adults-may-face-health2022. Accessed on March 14, 2023.
3. US Census Bureau. Predicting the Effect of Adding a Citizenship Question to the 2020 Census. 2019. Available from: https://www2.census.gov/ces/wp/2019/CES-WP-19-18.pdf. Accessed on March 14, 2023. 10.1007/s13524-019-00803-4.
4. Urban Insitute. Simulating the 2020 Census: Miscounts and the Fairness of Outcomes. 2021. Available from: https://www.urban.org/research/publication/simulating-2020-census-miscounts-and-fairness-outcomes. Accessed on March 14, 2023.
5. Yi SS, Kwon SC, Suss R, Đoàn LN, Iyanrick John, Islam NS, Trinh-Shevrin C. The Mutually Reinforcing Cycle Of Poor Data Quality And Racialized Stereotypes That Shapes Asian American Health. Health Aff (Millwood). 2022;41(2):296-303. doi:10.1377/hlthaff.2021.01417.
6. Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement. Washington, D.C.: The National Academies Press; 2009.
7. Coates T. What We Mean When We Say 'Race is a Social Construct'. In. The Atlantic May 15, 2013.
8. Gannon M. Race Is a Social Construct, Scientists Argue. In. Scientific American Febuary 5, 2016.
9. Cooper R, David R. The biological concept of race and its application to public health and epidemiology. J Health Polit Policy Law. 1986;11(1):97-116.
10. Cavalli-Sforza LL, Menozzi P, Piazza A. The History and Geography of Human Genes. Princeton, NJ: Princeton University Press; 1994.
11. Lopez N, Vargas ED, Juarez M, Cacari-Stone L, Bettez S. What's Your "Street Race"? Leveraging Multidimensional Measures of Race and Intersectionality for Examining Physical and Mental Health Status Among Latinxs. Sociol Race Ethn (Thousand Oaks). 2018;4(1):49-66.
12. United States Census Bureau. Measuring Race and Ethnicity Across the Decades: 1790–2010 Mapped to 1997 U.S. Office of Management and Budget Classification Standards. Accessed September 6, 2022.
13. Institute of Medicine. Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement. 2009. Washington, DC: The National Academies Press. https://doi.org/10.17226/12696.

14. Wallman KK, Evinger S, Schechter S. Measuring our nation's diversity: developing a common language for data on race/ethnicity. Am J Public Health. 2000;90(11):1704-1708.
15. Pew Research Center. Koreans in the U.S. Fact Sheet. https://www.pewresearch.org/social-trends/fact-sheet/asian-americans-koreans-in-the-u-s/. Published 2017. Accessed February 12, 2021.
16. Office of Management and Budget. Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity. October 30, 1997. Available from: https://obamawhitehouse.archives.gov/omb/fedreg_1997standards.
17. The White House Office of Management and Budget. Revisions to the Standards for the Classification of Federal Data on Race and Ethnicity. https://obamawhitehouse.archives.gov/omb/fedreg_1997standards. Published October 30, 1997. Accessed September 7, 2022.
18. Noe-Bustamante L, Mora L, Lopez MH. About One-in-Four U.S. Hispanics Have Heard of Latinx, but Just 3% Use It. Pew Research Center;August 11, 2020.
19. United States Census Bureau. B02015 Asian Alone by Selected Groups. https://data.census.gov/table?q=%09+B02015&g=0400000US36&tid=ACSDT5Y2020.B02015. Published 2020. Accessed December 8, 2022.
20. Institute of Medicine. Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement. Washington, D.C.: The National Academies Press; 2009.
21. American Anthropological Association. American Anthropological Association Response to OMB Directive 15: Race and Ethnic Standards for Federal Statistics and Administrative Reporting. https://s3.amazonaws.com/rdcms-aaa/files/production/public/FileDownloads/pdfs/cmtes/minority/upload/AAA_Response_OMB1997.pdf. Published September 1997. Accessed September 6, 2022.
22. U.S. Commission on Civil Rights. Racial Categorization in the 2010 Census. https://www2.law.umaryland.edu/marshall/usccr/documents/cr182r122.pdf Published March 2009. Accessed September 6, 2022.
23. Pew Research Center. Multiracial in America. June 11, 2015.
24. Pew Research Center. Majority of Latinos Say Skin Color Impacts Opportunity in America and Shapes Daily Life. November 4, 2021.
25. 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).
26. Assad S, Tucker D, Hacker C. Arab Americans in Illinois will finally be counted when they get their vaccine. CBS Chicago August 2, 2021.
27. Disparities Data Committee. Update of Racial Disparities in Michigan & COVID-19. https://www.michigan.gov/-/media/Project/Websites/coronavirus/Folder20/Racial_Disparities_in_Michigan__COVID-19_Biweekly_Report_20210604.pdf. Published June 4, 2021. Accessed April 21, 2023.
28. Danforth N. How the Middle East was invented. The Washington Post May 19, 2016.
29. SWANA Alliance. https://swanaalliance.com/about. Accessed September 13, 2022.
30. Proceedings of the National Academy of Sciences. Ethnoracial identity of MENA Americans [Internet]; March 14, 2022. Podcast. Available from: https://www.pnas.org/post/podcast/ethnoracial-identity-mena-americans

31. Orvis K. Reviewing and Revising Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity. In: The White House Briefing Room, ed. Vol 2022 June 15, 2022.
32. Office of Management and Budget. Initial Proposals For Updating OMB's Race and Ethnicity Statistical Standards. January 27, 2023. Available from: https://www.federalregister.gov/documents/2023/01/27/2023-01635/initial-proposals-for-updating-ombs-race-and-ethnicity-statistical-standards.
33. Asian & Pacific Islander American Health Forum. Policy Recommendations: Health Equity Cannot Be Achieved Without Complete and Transparent Data Collection and the Disaggregation of Data. https://www.apiahf.org/wp-content/uploads/2021/02/APIAHF-Policy-Recommendationas-Health-Equity.pdf. Published February 2021. Accessed September 6, 2022.
34. Ruggles S, Flood S, Goeken R, Schouweiler M, Sobek M. IPUMS USA: Version 12.0 [dataset]. https://doi.org/10.18128/D010.V12.0. Published 2022. Accessed November 7, 2022.
35. The Center for Evaluation and Applied Research (CEAR) at the New York Academy of Medicine (NYAM). Data Disaggregation Community Focus Group Report. December 2022.
36. Yi SS. Taking Action to Improve Asian American Health. Am J Public Health. 2020;110(4):435-437.
37. Yi SS, Kwon SC, Sacks R, Trinh-Shevrin C. Commentary: Persistence and Health-Related Consequences of the Model Minority Stereotype for Asian Americans. Ethn Dis. 2016;26(1):133-138.
38. VanderWeele TJ, Robinson WR. On the causal interpretation of race in regressions adjusting for confounding and mediating variables. Epidemiology. 2014;25(4):473-484.
39. Chokshi DA, Foote MMK, Morse ME. How to Act Upon Racism—not Race—as a Risk Factor. JAMA Health Forum. February 24, 2022;3(2).
40. Nguyen KH, Lew KP, Trivedi AN. Trends in Collection of Disaggregated Asian American, Native Hawaiian, and Pacific Islander Data: Opportunities in Federal Health Surveys. Am J Public Health. 2022:e1-e7.
41. Shimkhada R, Scheitler AJ, Ponce N. Capturing Racial/Ethnic Diversity in Population-Based Surveys: Data Disaggregation of Health Data for Asian American, Native Hawaiian, and Pacifc Islanders (AANHPIs). Population Research and Policy Review. 2021;40:81-102.
42. Wyatt LC, Russo R, Kranick J, Elfassy T, Kwon SC, Wong JA, Đoàn LN, Trinh-Shevrin C, Yi SS. 2012-2018 Health Atlas for Asian Americans, Native Hawaiians, and Pacific Islanders: A comprehensive look at AA and NH&PI health in the U.S.
43. Lauderdale DS, Kestenbaum B. Asian American ethnic identification by surname. Population Research and Policy Review. 2000;19(3):283-300.
44. Elliott MN, Morrison PA, Fremont A, McCaffrey DF, Pantoja P, Lurie N. Using the Census Bureau's surname list to improve estimates of race/ethnicity and associated disparities (vol 9, pg 69, 2009). Health Serv Outcome. 2009;9(4):252-253.
45. Chin MK, Đoàn LN, Russo RG, Roberts T, Persaud S, Huang E, Fu L, Kui KY, Kwon SC, YI SS. Methods for Retrospectively Improving Race/Ethnicity Data Quality: A Scoping Review [published online ahead of print, 2023 Apr 12]. Epidemiol Rev. 2023;mxad002. doi:10.1093/epirev/mxad002.

46. Budiman A, Ruiz NG. Key facts about Asian Americans, a diverse and growing population. Pew Research Center. 2021. Available from: https://www.pewresearch.org/fact-tank/2021/04/29/key-facts-about-asian-americans/. Accessed on March 14, 2023.

47. Lesko CR, Buchanan AL, Westreich D, Edwards JK, Hudgens MG, Cole SR. Generalizing Study Results: A Potential Outcomes Perspective. Epidemiology. 2017;28(4):553-561.

48. Smith LH. Selection Mechanisms and Their Consequences: Understanding and Addressing Selection Bias. Current Epidemiology Reports. 2020;7(4):179-189.

49. Agency for Healthcare Research and Quality. Figure 3-3. Centers for Disease Control and Prevention (CDC) ethnicities rolled up to the Office of Management and Budget (OMB) minimum categories for race and Hispanic ethnicity with subcommittee annotations. https://www.ahrq.gov/research/findings/final-reports/iomracereport/reldata3fig3-3.html. Published Content last reviewed May 2018. Accessed September 7, 2022.

50. Agency for Healthcare Research and Quality. Race, ethnicity, and language data: Standardization for health care quality improvement. https://www.ahrq.gov/research/findings/final-reports/iomracereport/index.html. Published April 2018. Accessed December 8, 2022.

51. Wong F, Halgin R. The "Model Minority": Bane or Blessing for Asian Americans? Journal of Multicultural Counseling and Development. 2006;34(1):38-49.

52. Kawai Y. Stereotyping Asian Americans: The Dialectic of the Model Minority and the Yellow Peril. Howard Journal of Communications. 2005;16(2):109-130.

53. John DA, de Castro AB, Martin DP, Duran B, Takeuchi DT. Does an immigrant health paradox exist among Asian Americans? Associations of nativity and occupational class with self-rated health and mental disorders. Soc Sci Med. 2012;75(12):2085-2098.

54. Asian American Bar Association of New York. A Rising Tide of Hate and Violence against Asian Americans in New York During COVID-19: Impact, Causes, Solutions. https://cdn.ymaws.com/www.aabany.org/resource/resmgr/press_releases/2021/A_Rising_Tide_of_Hate_and_Vi.pdf. Published 2021. Accessed September 7, 2022.

55. Đoàn LN, Wyatt LC, Cowen C, Leung V, Yusuf Y, Wong JA, Chen A, Ghosh M, Tablada E, Kwon SC, Yi SS. (2022). New York City COVID-19 Community Health Resources and Needs Assessment (CHRNA). https://aanhpihealth.org/resource/nyc-covid-19-community-resources-and-needsassessment-nyc-covid-19-chrna.