Addressing Asian American Misrepresentation and Underrepresentation in Research

Nilay S. Shah, MD, MPH[1,2];
Namratha R. Kandula, MD, MPH[1,3]

Asian Americans are the fastest growing racial/ethnic minority group in the United States and have unique, heterogenous health status and outcomes across a range of conditions between disaggregated Asian subgroups. Despite the rapid growth of this group, clinical and epidemiologic research lags considerably in adequately and appropriately representing Asian Americans. Too often, Asian American participants and populations are inappropriately aggregated into a single race category in research, masking important differences between ethnic subgroups. In this commentary, actionable recommendations are provided to investigators in order to enhance inclusion and representation of Asian Americans in a broad scope of research programs. Incorporating these recommendations in research planning and conduct will support health and promote health equity for these populations. *Ethn Dis.* 2020;30(3):513-516; doi:10.18865/ed.30.3.513

[1] Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL
[2] Division of Cardiology, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL
[3] Division of General Internal Medicine and Geriatrics, Department of Medicine, Northwestern University Feinberg School of Medicine, Chicago, IL

Address correspondence to Nilay S. Shah, MD, MPH; Department of Preventive Medicine, Northwestern University Feinberg School of Medicine, 680 N. Lake Shore Drive, Suite 1400, Chicago, IL 60611; nilay.shah@northwestern.edu

Introduction

Asian Americans are the fastest growing racial/ethnic minority group in the United States. This group's approximately 70% increase in population since 2000 to current estimates of 20.9 million[1] has primarily been fueled by immigration. There are significant knowledge gaps about the health of Asian Americans. Yet, National Institutes of Health (NIH) funding on health and disease in Asian Americans lags behind the remarkable growth of this group, comprising only .17% of the NIH budget between 1992-2018.[2] Deficiencies in research in Asian Americans extend beyond inadequate funding. Asian Americans are grossly underrepresented as participants in epidemiologic and clinical research in the United States. When included, they are often aggregated into one "Asian" race category. As Asian Americans are projected to become the largest immigrant group in the United States by 2065,[3] understanding the current and future health of the nation requires a more granular, comprehensive, and contemporary understanding of health and disease outcomes in Asian American subgroups.

The largest Asian American subgroups are Chinese, Indian, Filipino, Vietnamese, Korean, Japanese, Pakistani, Cambodian, Hmong, Thai, Laotian, Bangladeshi, Burmese, Nepalese, Indonesian, Sri Lankan, Malaysian, Bhutanese, and Mongolian.[1] Accumulating evidence demonstrates that these Asian American subgroups are not uniformly affected by noncommunicable and communicable diseases. Aggregating Asian Americans into a single Asian race category masks important heterogeneity and obscures health disparities of Asian subgroups. Aggregation may consequently lead to disease under-recognition and undertreatment, exacerbating disparities and threatening health equity. For instance, the prevalence of type 2 diabetes mellitus appears similar in aggregate Asian Americans compared with non-Hispanic Whites. However, a 2019 study reporting the first national estimates of diabetes prevalence in Asian American subgroups showed that South Asians had the highest diabetes prevalence (23.3%) compared with 12.1% in non-Hispanic Whites, followed by Southeast Asians (22.4%), and East Asians (14.0%).[4] This study also found that Asian Americans were

most often under-diagnosed, potentially because of misperceptions that they are not at higher risk for diabetes resulting from concealed variability of diabetes prevalence in aggregated Asians. The reason for disease heterogeneity in disaggregated Asian American subgroups is likely multifactorial, reflecting not only cultural, lifestyle, socioeconomic, and genetic differences between Asian American subgroups, but also indicating early life and

---

Aggregating Asian Americans into a single Asian race category masks important heterogeneity and obscures health disparities of Asian subgroups.

---

environmental exposures and health care access in immigrant Asian Americans' countries of origin.

In the last decade, large population-based studies in the United States have started oversampling Asian Americans and collecting country of origin and ethnic identification data, providing the opportunity to report differences in health and disease in some Asian subgroups. For instance, disaggregated Asian American subgroup identification has been collected since 2013 in the

Behavioral Risk Factor Surveillance System (BRFSS). Similarly, the National Health and Nutrition Examination Surveys (NHANES) have identified participants of Asian race and disaggregated Asian subgroup since the 2011-2012 survey cycle. Yet, while BRFSS and NHANES data are publicly available, access to disaggregated Asian American ethnic identification remains restricted to the National Center for Health Statistics Research Data Center because of small sample sizes and disclosure concerns,[5] not only limiting their availability to inform research questions but suggesting that oversampling efforts remain inadequate for appropriate evaluation of Asian Americans. Despite efforts to enhance Asian American representation in federal data, publicly available data from US governmental sources – alongside most population-based local data – remain inadequate to study health and disease in Asian Americans.

RECOMMENDATIONS TO IMPROVE ASIAN AMERICAN REPRESENTATION

For existing data in which identification of Asian race or disaggregated Asian subgroups is limited, we provide the following two recommendations:

1) When Asian race is not available, investigators should explicitly identify this limitation in their discussion.
2) When Asian race is available but not disaggregated, the justifica-

tion that imperfect representation using an aggregate Asian category is more favorable than no representation at all has become increasingly outdated, as evidence mounts for significant differences in health and disease between Asian subgroups. Publishing aggregated data on Asian Americans may misrepresent the health status and health care needs of this diverse group. Thus, if investigators publish data using an aggregated Asian category, they should acknowledge that the findings do not reflect the health of all Asian Americans and avoid making conclusions about health indicators or disparities within this group. Such a recommendation also applies to other aggregated groups, such as Hispanic Americans.

Although we are encouraged that disaggregated Asian American subgroup identification is increasing in some epidemiologic studies, clinical research, including trials, registries, local cohorts, and with electronic health records (EHRs), recruiting and appropriately identifying Asian American participants lags behind. Pharmacologic and other intervention studies would be informed by identifying Asian subgroups to evaluate treatment effectiveness in heterogenous populations. Community-level identification of disaggregated Asian Americans is important for understanding local disease patterns. Health system collection of disaggregated Asian subgroup data is essential for quality improvement programs, especially since health system engagement, experiences of discrimination,[6] health-related cultural attitudes, and social determinants

such as English-language proficiency and income vary substantially between Asian American subgroups.[7]

Representation and reporting of Asian Americans in prospective clinical research may be improved by considering the following three recommendations:

1) Clinical investigators should routinely seek to understand the demographics of their populations of interest, including assessment of the Asian (and other) race/ethnic composition of their patient population and local community.

Pharmacologic and other intervention studies would be informed by identifying Asian subgroups to evaluate treatment effectiveness in heterogenous populations.

nity. These statistics may be available via local health departments, Asian health or social service nongovernmental organizations, or less commonly, EHRs. Where an Asian American population exists, investigators should develop plans to recruit Asian-race participants.

2) The relative dearth of data on disaggregated Asian Americans likely makes hypotheses of differences in relevant outcomes between disaggregated Asian subgroups challenging. Therefore, we recommend assuming that differences exist between subgroups until proven otherwise, ie, disaggregated subgroup identification should always be collected when Asian American participants are recruited. We acknowledge that disaggregated subgroup categories may have limited statistical power to detect differences across groups, thus it may be reasonable to aggregate Asian Americans for statistical comparison. However, the hypothesis-generating benefit of disaggregation – such as to estimate effect sizes – is of substantial value to close knowledge gaps and advance health equity, and disaggregated data should be published. Recruitment of Asian-race participants and collection of disaggregated subgroup identification will support and target future research for this minority group.

3) Commonly, disaggregated Asian ethnicity is conflated with county of origin, which on the surface may seem reasonable given relative similarity within Asian subgroup categories in lifestyle, behaviors, and language. However, we caution against asking Asian American research participants their country of origin to identify subgroups, given a long history of Asian diaspora to other countries prior to immigrating to the United States and the growing population of US-born Asian Americans. Instead, we recommend asking participants to self-identify as one or more of the aforementioned Asian subgroups. As the prevalence of each of these populations varies by region, only some categories may be relevant to an investigator's local population. Listing options for self-identification should therefore be guided by local demographics.

Conclusion

The five recommendations we offer will not only support adequate and accurate representation of Asian Americans in prospective data collection, but will also optimally leverage existing data to promote research in this rapidly growing population. Ultimately, inclusion and appropriate identification of Asian Americans in clinical research is critical to understand health disparities in the United States and develop clinical, public health, and policy strategies that improve the health of all Americans.

Conflict of Interest
No conflicts of interest to report.

Author Contributions
Research concept and design: Shah, Kandula; Manuscript draft: Shah, Kandula; Acquisition of funding: Shah; Administrative: Shah, Kandula; Supervision: Shah, Kandula

References
1. *Asian Americans: A Diverse and Growing Population*. Pew Research Center. 2017. Last accessed January 11, 2020 from https://www.pewresearch.org/topics/asian-americans/.
2. Doàn LN, Takata Y, Sakuma KK, Irvin VL. Trends in Clinical Research Including Asian American, Native Hawaiian, and Pacific Islander Participants Funded by the US National Institutes of Health, 1992 to 2018. *JAMA*

Netw Open. 2019;2(7):e197432. https://doi.org/10.1001/jamanetworkopen.2019.7432 PMID:31339543

3. Cohn D. *Future Immigration Will Change the Face of America by 2065*. Pew Research Center. 2015. Last accessed January 20, 2020 from https://www.pewresearch.org/fact-tank/2015/10/05/future-immigration-will-change-the-face-of-america-by-2065/.

4. Cheng YJ, Kanaya AM, Araneta MRG, et al. Prevalence of diabetes by race and ethnicity in the United States, 2011-2016. *JAMA*. 2019;322(24):2389-2398. https://doi.org/10.1001/jama.2019.19365 PMID:31860047

5. Paulose-Ram R, Burt V, Broitman L, Ahluwalia N. Overview of Asian American data collection, release, and analysis: National Health and Nutrition Examination Survey 2011-2018. *Am J Public Health*. 2017;107(6):916-921. https://doi.org/10.2105/AJPH.2017.303815 PMID:28426300

6. McMurtry CL, Findling MG, Casey LS, et al. Discrimination in the United States: experiences of Asian Americans. *Health Serv Res*. 2019;54(S2)(suppl 2):1419-1430. https://doi.org/10.1111/1475-6773.13225 PMID:31657465

7. Kim W, Keefe RH. Barriers to healthcare among Asian Americans. *Soc Work Public Health*. 2010;25(3):286-295. https://doi.org/10.1080/19371910903240704 PMID:20446176