

# Defogging Data

As researchers separate out health data from culturally distinct groups, they're finding ways to reduce health disparities BY JYOTI MADHUSOODANAN

**MANY OF THE PATIENTS** who come to Eugene Yang's cardiology clinic trace their origins back to India, China, Korea, and multiple parts of Southeast Asia. His clinic is in Seattle, a hub for the tech industry and home to thousands of immigrant workers. Yang had seen firsthand how people from each of these groups were at risk of heart disease and how their typical lifestyles differ.

Yet despite differences in their cultures and backgrounds, these patients have been lumped together with people from other communities in a single category: Asian American, Native Hawaiian and Pacific Islander, or AANHPI. So Yang and his colleagues created a study looking at how social stress factors affect heart health in the Asian American communities he treats. The researchers analyzed stressors such as food insecurity, delays in medical care and living in a neighborhood that didn't feel close-knit or safe. Then they correlated these issues with risk factors for heart disease among Chinese, Filipino and Asian Indian adults. Other Asian communities were grouped together into a single category.

The recently completed study showed that the same stressors manifest differently in people of different ethnicities. Across the board, those who experienced more social stress had poorer sleep, struggled to exercise and used more nicotine—all factors associated with higher rates of heart disease. But differences emerged between groups. In Chinese Americans high stress was associated with an increased risk of diabetes, whereas in Filipino adults it was linked to high blood pressure. Asian Indians were most likely to experience poor sleep and physical inactivity when bearing the brunt of social stress. "There are significant differences in how social determinants of health impact the different Asian subgroups," Yang says. Recognizing this variation is a first step toward helping physicians tai-

lor interventions more appropriately.

For decades such nuance had been all but invisible to scientists, clinicians and policymakers. The single AANHPI category, which was defined in the 1997 U.S. Census, is still used widely by hospitals, as well as by state and national health databases. Researchers and policymakers use these data to assess disease rates and peo-

ple's health needs and to decide how to allocate resources.

But the AANHPI category masks rich diversity. People in this group have ancestral links to more than 50 countries. They collectively speak more than 100 different languages, have widely variable ways of life that differentially affect their health risks and represent a diversity of genetic backgrounds. They're also the fastest-growing racial and ethnic minority in the U.S. By pooling their data, researchers end up with a potpourri that obscures population-specific health needs or health risks. "When you lump everybody together, you don't see that maybe there are important differences," Yang says.



Now efforts led by advocates, researchers and community organizers—most of them from AANHPI communities—are paving the way to data equity and better health.

Spurred in part by the realization that aggregated data masked stark health disparities during the COVID pandemic, researchers began studying disease risk in specific AANHPI cohorts such as Pacific Islander, South Asian and Vietnamese populations. They're finding that teasing apart data in community-specific ways lets them use race and ethnicity information without conflating it with biology. Policymakers are catching up, too, using data specific to individual communities to better understand how to allocate resources and communicate more effectively.

These efforts are improving AANHPI health outcomes, says epidemiologist Stella Yi of New York University Langone Health. In recent years disaggregating AANHPI data has helped health-care professionals improve hepatitis B vaccination rates, reduce the devastation that has been caused by COVID and wildfires among Hawaiian communities, and identify better diet strategies to help South Asian communities reduce their risk of heart disease. "It's been really exciting to watch," Yi says.

**T**ELLIE (CHANTELLE) MATAGI was a 20-inch, eight-pound, six-ounce bundle of newborn joy in a Utah hospital nursery when her identity vanished into the health system. On hospital forms Matagi, who is of Samoan ancestry, had been labeled Asian, a category that blurred racial lines so completely it rendered them meaningless. Matagi, a community health leader who managed the Pacific Islander Task Force within the Hawaii State Department of Health during the early days of the COVID pandemic, says the record bothered her parents. It also troubled Asian staff at the hospital, who recognized the incongruity of so many people being lumped together. Matagi ended up quitting her job in 2022 to address her own health. She had diabetes and high blood pressure, and her doctors suggested she just lose weight. But because she was familiar with the science and knew aggregated data were masking her Samoan ancestry, she realized they

couldn't understand her true health risks. "I knew I wasn't being seen," she says.

Grouping too much data blurs the reality of people's lives. For example, in the aggregate, the risk of cancer death among Asian Americans is about 40 percent lower than that for white people. But disaggregating data reveals important patterns. Within the AANHPI group, lung cancer is the leading cancer diagnosis among Vietnamese, Laotian and Chamorro (those with ancestry in the Mariana Islands) men, and colorectal cancer is highest among Laotian, Hmong and Cambodian men.

When data are pooled, these nuances vanish. "One group looks better than they really are, the other group looks worse than they really are, and you can't rely on those estimates anymore," says Joseph Kaholokula, a physician at the University of Hawai'i at Mānoa. "It's nonsense. It's not good science, yet people have been doing this for decades."

That's because for decades federal and state health databases have offered researchers only a high-altitude view. Early attempts to break population data down with greater granularity failed because there simply weren't enough people in each group. The effort sparked concerns that, although the people included in these health-related data samples should remain anonymous, there were so few they could be easily identified. And funding to look at AANHPI health has been limited—a 2019 study reported that over the previous 25 years, only 0.17 percent of all National Institutes of Health funding for clinical research supported projects focused on AANHPI communities.

This is in part the result of broader stereotyping of Asian Americans as a "model minority," a category in which everyone is assumed to be well educated, financially secure and generally healthy. The model-minority trope illustrates how race-based assumptions can bias scientific research, says Tina Kauh, a program manager at the Robert Wood Johnson Foundation. "It's important for people to recognize that systemic racism is really what's driving the fact that we don't disaggregate data." With so little NIH funding to support their work, scientists have struggled to dispel the model-mi-

nority myth. "It's like this hamster wheel you get stuck on," Kauh says.

KAUH FIRST BUMPED into that cycle in college during an undergraduate psychology class about how culture and ethnicity shape someone's behaviors and perceptions of social norms. Fascinated, she tried to dig deeper into the experiences of Asian Americans, yet she couldn't find the data. Kauh persisted, revisiting the topic in graduate school but says she found it "basically impossible" to get funders interested. Since then, she says, "it's been this mission of mine to try to push for collecting data about Asian Americans."

Kauh's parents were Korean immigrants who owned a convenience store in Philadelphia. Even as a teen, Kauh could tell that their grueling schedules, language issues and social isolation took a physical and mental toll. Their lives were hardly those of a model minority. "I could see the challenges they experienced on a daily basis, but no one ever really talked about that except to frame it as 'look how hardworking they are,'" she says.

The social stressors Kauh's parents experienced were financial and cultural, both of which can affect a person's health. Language barriers, racism, changes in diet with the move to a new country and the circumstances of that move—whether someone migrates to pursue a graduate degree or to flee from conflict—can add up. None of these factors are related to the biological basis of disease, but they determine what resources a person or community might need to achieve good health.

When researchers understand the links between social factors and people's health, they can begin to design tailored solutions. Food is one clear example. In the U.S., South Asian communities have disproportionately high rates of heart disease—an observation often explained by diet, says Alka Kanaya, a clinician who studies diabetes at the University of California, San Francisco. Researchers typically gather details about food habits using a list of standard questions based on Western diets that don't represent global cuisines. Advice about what constitutes a "healthy" food is also based on studies conducted with Western diets. "You have to be specific to what people may be eating

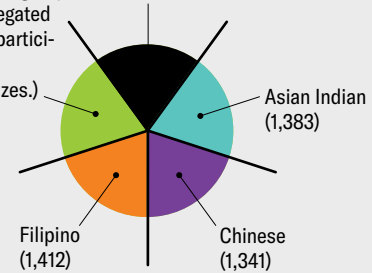


# Heart Health across Asian American Communities

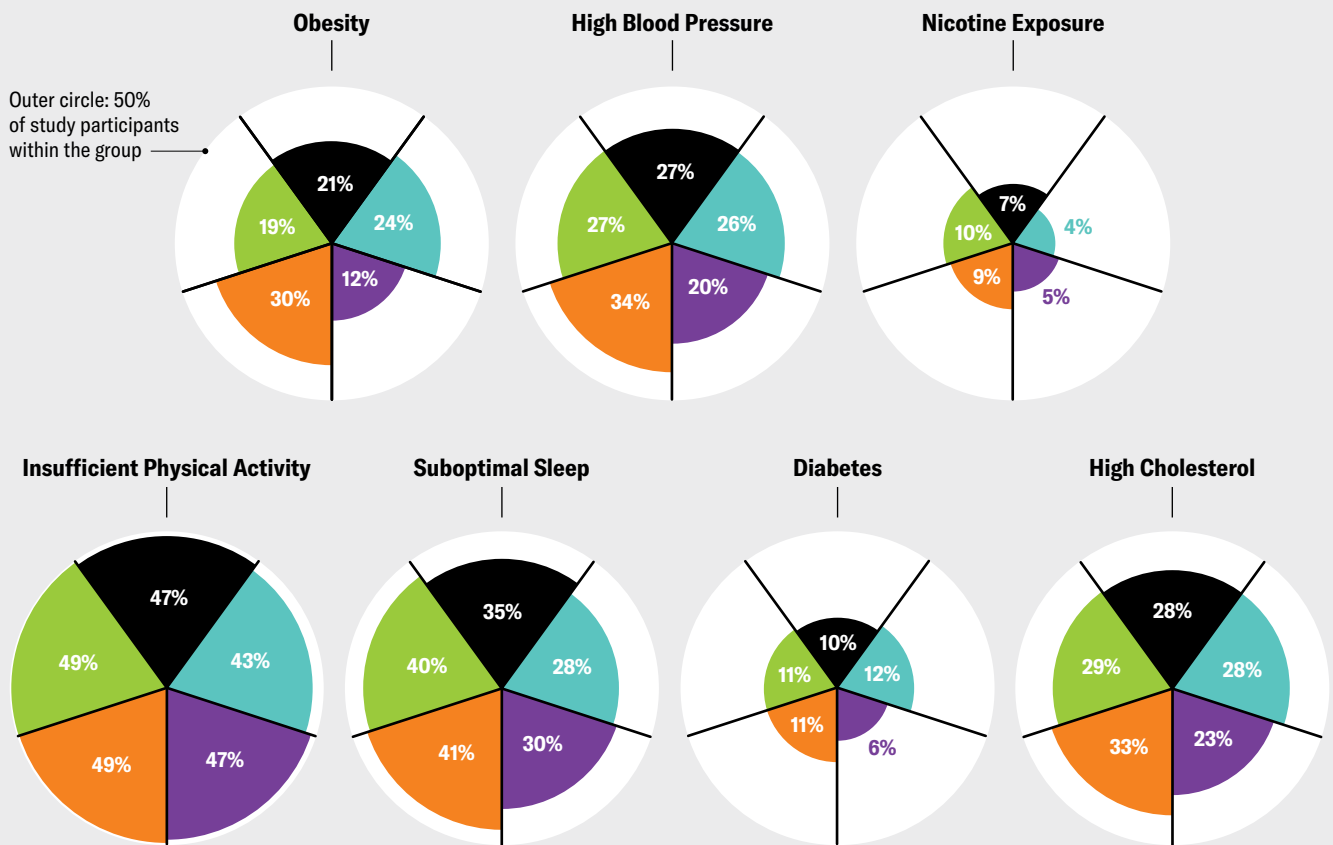
Food insecurity, living in an unsafe neighborhood, exposure to nicotine, and other social factors can increase risk of heart disease. But not everyone is affected the same way. Grouping data from diverse groups together hides the differences. A 2024 analysis by Eugene Yang and his colleagues found that Filipino, Chinese and Asian Indian adults experience different health outcomes in response to social and environmental stress.

All Asian Americans except for Asian Indian, Chinese, and Filipino. (This subgroup couldn't be disaggregated further because of participant confidentiality and small sample sizes.)

Prevalence values for all Asian American adults in the study (total of 6,395 people)



**Prevalence of Cardiovascular Risk Factors, by Asian American Subgroup** (age standardized, self-reported)



and how they may be cooking it. Having nonaccurate ways of measurement just gives you useless data,” Kanaya says.

For the past decade Kanaya and other researchers have run a study of heart health among South Asians living in the U.S. called Mediators of Atherosclerosis in South Asians Living in America (MAS-ALA). It includes a food-frequency questionnaire that lists many South Asian foods, such as dhokla (a savory cake), sambar (lentil stew), steamed fish, lamb curry and popular snacks. Last year the researchers analyzed the diets of nearly 900 people from the study and identified

foods correlated with a “South Asian Mediterranean-style diet”—one rich in fresh vegetables, fruit, fish, beans and legumes. They found that people who ate more of these foods had a lower risk of heart disease and diabetes than other people in the cohort.

Data such as these can help clinicians advise patients more effectively by offering dietary solutions that may be easier for them to follow rather than forcing a more Western lifestyle on them, Kanaya explains.

GETTING GRANULAR with community data proved to be a lifesaving strategy in

Hawaii during the worst of the COVID pandemic. The state health department’s infectious disease team was heavily focused on controlling the spread of the virus at the start in 2020. But the scientists were “thinking of it in terms of a purely biological system versus understanding what puts people at risk,” says Joshua Quint, an epidemiologist at the Hawaii State Department of Health. “Accurate measurement of social factors is so important.”

To gather those data, Quint teamed up with Matagi and Kaholokula, the University of Hawai‘i physician, to form a COVID investigation team. The group quickly

# When researchers understand the links between social factors and health, they can begin to design tailored solutions.

discovered there was no way to figure out which of the Native Hawaiian and 20 or more Pacific Islander communities needed resources or what those resources were. The data at hand were simply too sparse to base any estimates on. So the team began recording COVID deaths with more specific demographic details. When counts were low enough that they risked making individuals identifiable, the team noted these details in a separate section of the database to ensure that information from smaller communities was not lost in an aggregate, Matagi says.

The team members didn't just gather information—they shared it with the communities through hours of virtual visits and phone calls. As they talked, the carefully gathered and stored details helped communities see their own losses amid the sea of numbers. No one could deny the devastation they'd experienced, nor could their experiences be minimized by a database that didn't represent them and their needs. The strategy was especially effective among the Samoan, Marshallese and Chuuk (people originally from part of Micronesia), Matagi says, because they were the three Pacific Islander communities most affected by the disease.

The researchers worked with each community to identify specific requirements. Some needed a safe place to keep healthy family members distanced from those with COVID, others wanted more resources allocated to food or medical care, and still others sought a way to maintain social connections or attend religious gatherings virtually while observing COVID precautions.

The same approach helped the team customize care after the Maui wildfires by recognizing specific needs such as food, shelter and medicine. Its methods have since been highlighted by the World Health Organization as an effective way to reduce health disparities.

Identifying a community's needs and meeting them appropriately can make a range of infectious diseases more man-

ageable. In New York City in the early 2000s, routine hepatitis B vaccination was available only to children. Among adults the virus was typically seen as a sexually transmitted infection (STI), and testing and treatment were offered primarily at HIV clinics.

But the infection was common among Asian American immigrants because of high endemic rates in their countries of origin. In families the virus passed between married partners, from person to person through household contact such as the sharing of utensils, and from mother to child during childbirth. These adults were unlikely to seek care at an STI clinic. At the time, researchers reported rates of hepatitis B among Asian Americans that were about 50 times higher than those among non-Hispanic white people, as well as rates of liver cancer, a common consequence of infection, that were several times higher. In 2003 researchers at New York University teamed up with community organizers, politicians and clinicians in the city to help address the disparity.

The coalition's work helped to establish that the problem would not be stemmed by STI clinic screenings, because that "was not somewhere that we knew Asian American immigrant adults would feel comfortable going," says epidemiologist Simona Kwon of N.Y.U. Langone Health, who joined the effort a few years after it began. "The communities are very different," Kwon says, "and the health priorities are different." Western social norms and biased perceptions had been unintentionally driving health outcomes for hepatitis B.

The N.Y.U. team helped city officials implement community-based programs and offer adult vaccinations at primary care clinics and through community-based organizations. Recognizing that not just viral infection rates but social conventions guide people's choices about care was the key to driving down hepatitis B transmission.

Quint warns that in efforts to apply race and ethnicity data, researchers and policymakers should be careful not to conflate a person's health with these factors alone. Aggregated or not, race and ethnicity are always simple representations of broader social and cultural factors that affect health. But disaggregation, he says, can "help us get beyond race and talk about ethnicity in ways that are more meaningful and helpful."

**E**FFORTS TO CREATE community-specific solutions are what "actually move the disparities dial," Matagi says. Now, after the success of state- and community-level studies, policymakers are launching larger studies and investing more money in the hopes of better understanding the health of different groups under the AANHPI umbrella.

Last year the White House announced a national effort to prioritize equity for AANHPI communities, and earlier this year the National Heart, Lung, and Blood Institute launched a large epidemiological study to understand health trends in these populations. This seven-year project, named the Multi-ethnic Observational Study in American Asian and Pacific Islander Communities (MOSAIC), aims to track the health of 10,000 people who identify with various AANHPI subgroups. One challenge, Kanaya says, will be to find out how granular they can get—keeping the data anonymized but with sufficient detail to identify meaningful trends, yet without adding so many checkboxes that a long list leaves participants exhausted.

Establishing new categories of race and ethnicity may seem to contradict efforts to make medicine and health care equitable and free of racial bias. But done right, these endeavors can be complementary. "There's a push to avoid talking about race, and I think there are big risks associated with that if it's coming from a place of wanting to ignore problems," Quint says. "We need statistics that cut across all ranges of demographic factors so we can find out if we're building a more just and fair society."

**Jyoti Madhusoodanan** is a science journalist based in Portland, Ore. She covers health, medicine and the life sciences.