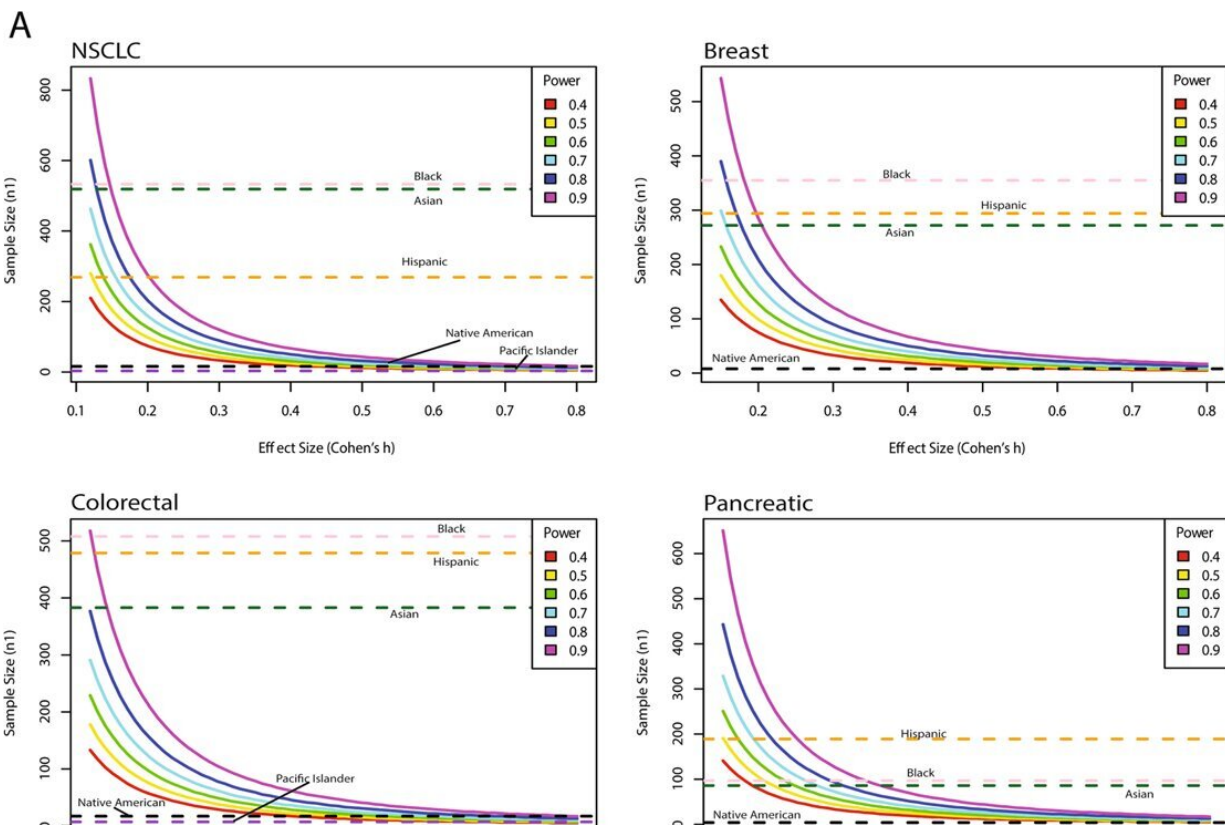# Researchers uncover racial and ethnic disparities in a widely used precision oncology data registry

February 1 2023, by Katie Marquedant

Sample size estimation to detect differences in mutational proportions between white and racial/ethnic minority patients at varying power increments. Estimated number of samples (y-axis, n1) required to detect differences in mutational proportions between white and non-white patients at varying statistical power (solid curved lines) for different effect sizes (x-axis). All horizontal dashed lines indicate the current actual number of samples in GENIE for a given racial/ethnic

Biorepositories created to support precision cancer research through their vast stores of genomic data may lack sufficient representation of cancer distribution among racial and ethnic minorities, researchers from Massachusetts General Hospital (MGH), a founding member of Mass General Brigham, have found.

In a paper published in *npj Precision Oncology*, the team reported that this omission may impact the validity of studies that attempt to directly compare mutational profiles between ethnic and racial groups for the most common types of cancer, thus limiting any generalization of biomarker discoveries to all patient populations.

"Our findings suggest that these data registries do not reflect the true landscape of cancer patients in the U.S., and may therefore misrepresent the disease burden in many racial/ethnic minority populations," says senior author Sophia Kamran, MD, with the MGH Department of Radiation Oncology, and director of Diversity, Equity, and Inclusion for the department. "The need is greater than ever for diversity to be a mission-critical priority for precision medicine so that breakthroughs or new findings from biorepositories can broadly apply to and be safe for diverse cancer populations."

Large biorepositories, designed to help scientists better understand and treat cancers, have historically included mostly white patients.

For their study, MGH researchers focused on the American Association of Cancer Research (AACR) Project Genomics Evidence Neoplasia Information Exchange (GENIE), a publicly available, international cancer registry populated with clinically relevant genomic sequencing information. Multiple research organizations have used Project GENIE to describe novel genomic variations between racial categories of different tumors. These discoveries, however, have often conflicted with other findings and not been amenable to replication or biological explanations.

"Our work showed that GENIE is not sufficiently powered to detect small yet potentially clinically meaningful differences between white and non-white patients in even the most common cancer types," explains Kamran.

To get a better reading of the GENIE tumor landscape relative to the broader cancer population, researchers used the Centers for Disease Control and Prevention's Wide-Ranging Online Data for Epidemiologic Research (CDC-WONDER) which, specially, defined the proportion of cancer patients expected from each racial/ethnic group for a given cancer type.

What they found was that white patients were adequately or overly represented for all cancer types except lung and melanoma, and that Asian patients had significant overrepresentation across several tumor types, including colorectal, lung, and breast.

Black patients, however, were significantly underrepresented for most cancers, except melanoma and cervical, while Hispanics were consistently underrepresented among all cancers. Native American and Pacific Islander populations had zero representation for many cancer types.

The reasons behind low registry representation of minorities are varied, ranging from ingrained institutional exclusion to mistrust of research agendas by minority groups themselves.

The MGH research team expressed hope, however, that greater recognition of the scope of the problem will lead to sustainable improvements.

"Even basic changes in methodology and recruitment for future biorepositories can dramatically improve their ability to address health disparities," notes Kamran.

"For instance, standardized methods to obtain demographic information of participants should be implemented by all institutions, and self-identified ethnicities should have more options than the binary 'Hispanic' or 'Non-Hispanic' present in the current GENIE database."

More information: Alexander T. M. Cheung et al, Racial and ethnic disparities in a real-world precision oncology data registry, *npj Precision Oncology* (2023). DOI: 10.1038/s41698-023-00351-6

Provided by Massachusetts General Hospital