The Race for Ancestral Genetics in Clinical Trials

A seasoned cancer researcher would never set up a study in which all the ill patients were, say, Canadian, and all the healthy controls were Japanese. And yet cancer researchers risk making a similar mistake if they overlook genetic information that fleshes out what many of us like to think of as race or ethnicity, some experts say.

Fortunately, awareness of how ancestral genetics might contribute to risk of disease and drug response in people has risen over the last several years. Studies that look directly at the problem are on the rapid rise, and this increased interest has biotechnology companies lowering the cost of tests that determine genetic ancestry, thanks to a little competition. However, experts have yet to decide on how to genetically define ancestry, suggesting examining anywhere from a handful to hundreds of gene variants.

"We've gone from a very controversial and understudied issue 5 years ago, to lots of papers coming out recently," says geneticist Timothy Rebbeck, Ph.D., of the University of Pennsylvania School of Medicine in Philadelphia.

Racial Tensions in Our Genes

To understand how the genetic underpinnings of people from different backgrounds might skew clinical trial results, he says, imagine trying to determine whether a link exists between telephone usage and cancer. Comparing the health and habits of people in the United States with a country that has nonexistent cancer rates and few telephones would result in a frighteningly high risk of cancer for all those phone calls made last week.

"Replace telephone with 'genes' and you're set," says Rebbeck.

Not including information on the race or ethnicity of study volunteers could skew disease risks as stronger or weaker than they really are. But geneticists aren't sure exactly how big of a problem that lack of information represents in current clinical studies.

"In most situations, it's probably small," says Rebbeck. "But we don't know *when* it's a [big] problem." Some researchers think the big study to define the problem has yet to be undertaken.

Researchers first got a sense of how race could confound genetic studies of diseases in the late 1980s. Epidemiologists found what appeared to be a genetic variant in Pima Indians in the southwest United States that correlated with type II diabetes. But the Pimas had substantially higher rates of both the variant and diabetes than white people. In the end, the variant ended up correlating with being Pima, not with having diabetes. So comparing the American Indian population to whites resulted in a strong genetic link in the Pimas that wasn't really there.

But conducting genetic studies within one group or controlling for ethnicity can still result in higher risk on the basis of race. "Alzheimer's disease is the poster child for this problem," says pharmacogeneticist Esteban Burchard, M.D., of the University of California in San Francisco. A variant of the gene ApoE4 is a strong genetic risk factor for early-onset Alzheimer's disease, and the characteristic most likely to raise or lower that risk is race.

"It occurs in about 20% of the African American population, and it means nothing. It occurs in about 6% of the Japanese, and it makes their risk six times higher [than that for white people]," Burchard said. "Something about being Japanese unleashes the wrath of the gene, and something about being African American attenuates it."

A Cancer Confounder

More recent studies look at the problem in cancer. In the January 26 issue of the *New England Journal of Medicine*, researcher Christopher Haiman, Sc.D, and colleagues showed how the incidence of lung cancer varied among African Americans, Hawaiians, whites, Japanese Americans, and Latinos. In an accompanying editorial, genetic epidemiologist Neil Risch, Ph.D., also of UCSF, analyzed the relative risk of the different populations.

Among nonsmokers, race made no difference. "But for every level of cigarette smoking, African Americans are at much higher risk of developing lung cancer than whites, Japanese, or Latinos," Burchard says.

Race and ethnicity are proxies for many unidentified characteristics, Risch says, and these studies will ultimately help ferret out biological or environmental causes of cancer or other diseases among people. "Any time a variable is predictive, it's useful," Risch says.

Researchers led by molecular geneticist Georgia Dunston of Howard University in Washington, D.C., found that similar results disappeared for genes

thought to be involved in prostate cancer. Lumping together the incidence of prostate cancer in whites and African Americans revealed a link between a CYP3A4 gene variant and disease. But the variant showed up more often in blacks—it was present in only 10% of the European population but 66% of African Americans and in nearly 90% of a Nigerian population. When volunteers were classified on the basis of ethnicity, the white patients carried the variant about twice as often as healthy white controls did, indicating a link between the variant and prostate cancer. But the variant split itself more evenly between the cancer cases and controls in the other two groups. Correcting for this frequency difference wiped out the risk of disease for the African and African American groups.

"Stratification was a big problem," says molecular anthropologist Mark Shriver, Ph.D., of Pennsylvania State University in University Park. "If you know there's a lot of variation and you don't measure it, then you miss the boat on why stratification is important."

To avoid losing the fruit of your labors, says Burchard, "when you do genetic studies you want to make sure you're comparing apples to apples."

Assessing Ancestry

To assess the effects of ancestry, Shriver uses a set of 40 genetic markers shown to reveal ancestral data for each of at least five ancestral groups, including West Africans, Europeans, Native Americans (from both North and South America), East Asians, and South Asians.

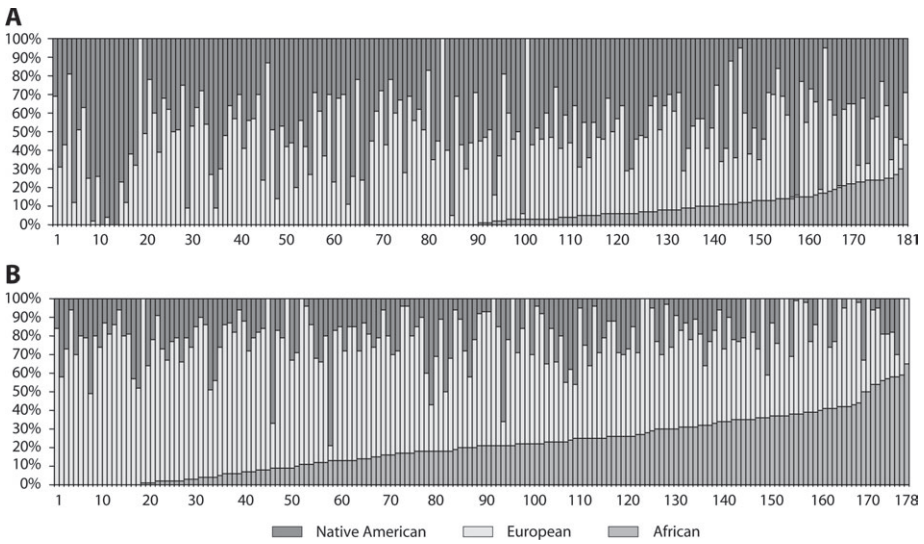
"We don't talk in terms of races. It's not that they don't exist. It's just that the terms are too complex," he says.

Instead, they look at these so-called ancestral informative markers (AIMs).

"One of the main things to remember is that race is not equal to genetics," says Rebbeck. "The markers don't define the race. They are just a hallmark."

Researchers are quick to point out that finding that race, ethnicity, or AIMs affect the risk of disease or how well drugs will kick the diseases are merely steps to learning more about the cancer.

"If you find racial differences, that doesn't say if it's genetic," Risch says. "It's a first clue—and you have to do studies to disentangle whether the differences have biological or environmental underpinnings."



What Race Means: Nearly 400 people consider themselves either purebred Mexican or Puerto Rican (individual bars) but 44 genetic markers reveal their mix of Native American, European, and African ancestry.

Efforts to include ancestral information in statistical analyses have historically relied on people's recollections. But this approach has been unreliable because most people are more mutt than they realize, Burchard has found (*see* graph).

In one study of Mexicans and Puerto Ricans, he required volunteers to have both parents and grandparents of the same national origin. Both of these groups are part Native American, part European, and part African. Of about 90 volunteers in each category, two to four were purebreds on the basis of AIMs—and they were either purebred Native American or purebred European. The rest ranged widely in their mix. The same incorrect reporting may obscure research findings. "In a case–control study, any positive or negative results might be due to the heterogeneity of the subjects," he says.

Statistical Fixes

Statistics make room for such variation, however. Using genetic markers to probe ancestry in cancer studies allows researchers to include biologically ill-defined notions of race or ethnicity as a continuous variable rather than a single choice in their statistical analyses. The difference is like being able to include the actual height of a person (in a study investigating height and shadow length, for example) rather than characterize everyone as either "tall" or "short."

"You get more statistical power that way," says Shriver.

The only way to be sure to not miss important variation is to "study all the variation," says Rebbeck. Shriver and Risch say one way to make sure you get it is to oversample minority populations in clinical trials.

"African Americans, Asians, and Hispanics might not be underrepresented in trials with respect to their percentage in the general population, but their [disease risks and response to drugs] are not going to be addressed because they haven't been included in adequate numbers," says Risch.

As diagnostic tests for ancestry get less expensive and can be included regularly in clinical trials, scientists will be able to get a handle on what risks our ancestors have left us with. Rebbeck says there is still a lot of work to be done to determine how race is relevant to cancer. "No one would say we've figured it out."

And Shriver reminds us that biology isn't everything in disease risk and how well we survive it: We inherit not only genes but also our family's wealth, the educational opportunities, and other factors. "We don't know the right questions to ask. At least, we have some recognition that all people are not the same."

—Mary Beckman

 DOI: 10.1093/jnci/djj402