SCIENTIFIC AMERICAN™

Permanent Address: **http://www.scientificamerican.com/article.cfm?id=whole-genome-sequencing-predict-disease**

How Useful Is Whole Genome Sequencing to Predict Disease?

Few diseases have strong enough genetic components to make sequencing a solid way to assess individual risk

By Katherine Harmon | Monday, April 2, 2012 | 1 comments

A $1,000 genome sequence is close to being available. What will your sequence tell you about your actual risk for certain diseases?

Many companies advertise a laundry list of disease risks associated with your genes. But your genome is unlikely to reveal whether or not you will actually get one of these conditions, according to a study published online April 2 in *Science Translational Medicine*.

"Whole-genome testing is not a crystal ball," Bert Vogelstein, director of the Ludwig Center for Cancer Genetics and Therapeutics at Johns Hopkins University, and co-author of the new paper says. "It may become one important determinant in patient care, but certainly not the only one—and possibly not even a major one." Also, he doubts it will ever surpass more traditional tactics used by doctors to help shape care regimens, including preventative medicine and other information such as family history. Most common diseases, heart disease and stroke included, do not result from a single-point mutation—or even a combination of them; such diseases also involve one's own lifestyle and environmental exposures.

He and his fellow researchers examined genetic and health data from large existing registries of tens of thousands of identical twins. They searched for 24 common maladies, including some cancers, along with autoimmune, heart and neurological diseases. Although we do not yet have a perfectly clear picture of all of these conditions' potential genetic correlates, the researchers developed a model that acted as if we did already have that science down in an effort to test the upper boundaries of what we *might* be able to learn from the genome alone. If the tests were optimal in this way—and as Vogelstein says, presuming "that the interpreter of the genome would be omniscient" and already knew all of the potential links with these diseases—nine out of 10 people could get a tip-off about one or more of these two dozen diseases with robust enough results to be able to change their lifestyles or medical responses based on them.

But overall, based on their genome sequence alone, most people will wind up getting a "negative" (that is, "low risk") result for most diseases. "These negative test results will, in general, not be very informative," Vogelstein and his colleagues noted in their paper. Not only are low-risk readings less useful than high-risk ones (which could prompt lifestyle changes, more frequent screenings or prophylactic treatments) they also do not guarantee that an individual will not get a disease.

In fact, for about half of the diseases most of individuals from the registries who would have received a "low-risk" result from a whole-genome sequence would have ended up getting one of these diseases anyway. (Diseases that had a stronger link with the genome results included Alzheimer's disease, some autoimmune conditions, type 1 diabetes and, for men, heart disease.) And for most of the diseases the researchers studied, people who were classified as low-risk based on their genome still had a risk that was more than half that of the general population—thus, lower than average, but not exactly null.

"This level of risk reduction is probably not sufficient to warrant changes of behavior, lifestyle or preventative medical practices for these individuals," the researchers wrote.

Upside of negative news

Not all negative results are useless, however, some were relatively robust and could bring, if not a change in health behavior, at least some relief. Twins in the study whose genomes showed a low risk for Alzheimer's, for example, did have about a 12 percent (thus, much, much smaller) risk than that of the general population (which, to begin with, has a relatively small risk of getting Alzheimer's during their lifetimes). And that could be a big comfort, the researchers noted, "particularly to those with a family history of

Alzheimer's disease."

Muin Khoury, director of the Office of Public Health Genomics at the U.S. Centers for Disease Control and Prevention, who was not involved in the new study, notes that it is solid work and that the twin data is an excellent source.

At the same time, he says, "whole-genome sequencing is a great tool, but it's not ready for prime time—for a number of reasons."

Among those reasons, he notes, is that we have definitive genetic correlates for very few diseases—most of which are relatively rare in the general population. Most diseases are not inherently genetic in nature, and even if they seem to have some associated genetic hallmarks, those are not strong enough to be able to say for certain that a person will or will not get the disease at some point in his or her lifetime.

As Vogelstein notes, these genetic risk predictions are unlikely to get much stronger. Because they conducted their mathematical analysis based on an ideal scientific world, in which we already knew how gene variants were connected to diseases, even "1,000 years from now, with intense research, these numbers wouldn't change," he says. That is, they created a model that was as generous as possible in terms of genetic correlates to disease in order to create a sort of upper bounds of the utility of these tests.

Not all in the genes

One of the keys to improving disease risk prediction will be to collect even more comprehensive information from study subjects and individuals. "We need to integrate nongenetic factors because most factors are nongenetic anyway," Khoury notes. Variables such as diet, exposures to carcinogens (such as via first- or secondhand smoke) and family history have a strong influence on many disease risks. "That's part of the challenge we face," he says. The genome "really doesn't mean much by itself"—and for people seeking the most accurate picture of their risk for various diseases, a more nuanced and integrated picture is needed.

As Vogelstein and his co-authors pointed out in the study, most identical twins do not die from—or even get—the same diseases, which should be a first clue that genetics can only go so far in determining pathological outcomes.

Khoury cautions consumers that even if the price of a whole genome sequence becomes reasonable for them, it is worth spending some time asking, in turn, what they are going to get from it. Although the extra genetic information might not hurt, for an individual who is looking for educated estimations of disease risk, having strong family history and personal health and lifestyle information are some of the most valuable data points one can take to the doctor.

But that does not mean that whole-genome sequencing will not be useful in the future—or that is not already for some higher-risk individuals or well-characterized rare conditions. "I'm an optimist—but also a realist," Khoury says. "We'd all like for whole-genome testing to succeed—I think it's already succeeded in some areas. But for population health, I think we have a long way to go."

By better understanding the limitations of genome-wide sequencing, Vogelstein notes, researchers and policymakers might be better able to direct funding and efforts to areas, such as Alzheimer's disease, where a person's genetic profile might have a very real effect on their likelihood of getting the disease. With that information comes an opportunity to make a much larger impact on public health in general—and a way to ease consumers' minds as well.