Study highlights strategies for boosting accuracy of personal genetic risk scores

December 4 2020

Credit: CC0 Public Domain

As the consumer genetics industry rapidly expands, more and more people are turning to DNA-based services to learn their risk of developing a wide range of diseases.

However, the risk scores from these genetic tests are not always as

precise as they could be, according to a new study from Scripps Research. The scientists, whose research appears in the journal *Genome Medicine*, examine many approaches to calculating the scores and recommend that personal genomics organizations adopt standards that will raise the bar for accuracy.

"Polygenic risk scores can be incredibly useful and affordable tools to guide preventative health decisions," says the study's senior author Ali Torkamani, Ph.D., director of Genomics and Genome Informatics at Scripps Research Translational Institute and Associate Professor of Integrative Structural and Computational Biology at Scripps Research. "Given the increasing utility of these scores, it's important to remove variability as much as possible and ensure the data is updated regularly to reflect new knowledge from genomic studies."

Companies such as 23andMe and Ancestry—along with dozens of others—need only a small sample of a person's saliva to generate estimations known as "polygenic risk scores," which are determined by comparing snippets of an individual's DNA with findings from large-scale genome studies. Examining these small segments of genomic variation, known as SNPs, is faster and far more economical than sequencing a person's entire genome. But one drawback is that the results can sometimes vary unexpectedly.

Torkamani notes that in most cases, the degree of fluctuation in a risk score is small and doesn't change the overall interpretation of the result; most individuals remain in the same risk category. Also, fluctuations are especially minor at the extremes of the risk distribution, where disease risk implications are most critical. Yet any variation can lead to a loss of consumer confidence in these important scores, and that can undermine preventative health. And in some rare cases, the scores can fluctuate substantially.

For their study, Torkamani and his team used various methods of calculating polygenic risk scores for conditions including coronary artery disease, atrial fibrillation, type 2 diabetes, Alzheimer's disease, glaucoma and breast cancer. They found that regardless of the method used, the computational algorithms are prone to introduce random variability for individuals due to how the calculations incorporate data from population-level genetic studies.

To reduce variability in individual polygenic risk scores, the team recommends running the algorithms multiple times, which would help smooth out random imperfections. By paying attention to variable elements in a score calculation, actions can be taken to either eliminate those elements or create an average, Torkamani explains.

"While we identified some clear challenges in applying population genetics tools to individual-level genetic analysis, we also see ways to overcome these issues to produce results that will inspire confidence," he says.

He and his team decided to pursue the study after making updates to different project they developed called MyGeneRank, which predicts risk of heart disease and other conditions by looking at specific genetic markers. They sought to identify a computational process that would lead to the greatest degree of stability and accuracy in the face of a score that is expected to evolve over time as new data becomes available.

The recommendations put forth in the new study could help personal genomics companies improve the quality of the scores they deliver to customers—and ultimately, encourage more individuals take smart preventative health actions based on what they learn in genotyping reports.

  More information: Shang-Fu Chen et al, Genotype imputation and