



# Subtle signatures of risk

By accounting for the additive effect of multiple genetic variants, researchers can develop a scoring system that improves their ability to identify the people most vulnerable to heart disease.

By Michael Eisenstein

**B**y the time we reach middle age, many of us can look forward to serious talks from our physician about blood pressure, cholesterol levels, diet and exercise. These and other factors can meaningfully affect the risk of developing coronary heart disease. This condition arises from narrowing of the heart's arteries and can result in heart attack or heart failure, making it the leading cause of death worldwide.

But awareness of these risk factors becomes useful only later in life, when the likelihood of a heart attack has already begun to climb sharply. This means people who develop coronary heart disease earlier or without exhibiting a typical risk profile might slip through the net until a serious cardiac event occurs. This is especially unfortunate given that, with advance warning, an individual could potentially have ample opportunity to change the course of disease. "Pretty much everybody can optimize their diet, physical activity and

other health-related behaviours," says Pradeep Natarajan, a cardiologist at Massachusetts General Hospital in Boston. "And in preventive cardiology, we also try to identify individuals who are suitable for medicines."

Genetic testing could offer a powerful complement — especially because roughly half of the risk of coronary heart disease arises from genetic variation. "Genetics can be the first thing that is identified as abnormal for a patient, and that can start the conversation," Natarajan says. The genetic roots of coronary heart disease are complex, typically arising from the combined influence of many genome sequence variants rather than a single culprit mutation. But researchers are now developing polygenic risk scores that take the multifaceted origins of this complex disorder into account.

Early data suggest that such scores could help to manage the public-health burden of heart disease, by enabling clinicians to assess patients who might benefit most from early

interventions such as cholesterol-lowering statin drugs. However, clinical researchers are still waiting for robust proof that these scores can move the needle in the right direction and prevent outcomes such as heart attack and stroke.

## Death by a thousand cuts

The standard tactic for identifying genetic components of disease is a tool known as a genome-wide association study (GWAS). This entails surveying the genetic variation in many thousands of people, with one population affected by the condition of interest and the other representing healthy controls. Statistical analysis subsequently reveals which sequence variants have a statistically significant association with disease — meaning their presence can confidently be linked to that condition.

Researchers have published more than 100 such studies for coronary heart disease, revealing a few genes that amplify risk owing to factors such as elevated cholesterol. Those standout mutations, however, tend to be rare. Michael Inouye, a computational biologist with a joint appointment at the University of Cambridge, UK, and at the Baker Heart and Diabetes Institute in Melbourne, Australia, notes that clinical researchers now recognize that most individual variants have only a tiny influence on risk of disease. "It became clear very quickly that you had to combine those variants together to get any sort of potentially useful information," he says.

This can be achieved by selecting subsets of GWAS-identified variants that cross a threshold of statistical significance, and then assessing their combined influence on coronary heart disease. That analysis yields a metric — the polygenic risk score (PRS) — that can grade someone's relative risk based on the number and estimated impact of variants present in their genome. These measurements typically produce a bell-curve distribution showing how great the predicted risk of coronary heart disease is likely to be within a group of people. Most people fall in the middle, where their PRS offers little predictive value. "But for individuals at the tails of the distribution, you would be able to re-stratify them in terms of risk," says Iftikhar Kullo, a cardiologist at the Mayo Clinic in Rochester, Minnesota.

Notably, these scores tend to perform better when they are more inclusive rather than highly selective. A GWAS typically sets a high bar for statistical significance, such as a less than one in 20 million likelihood of an association arising by chance. This means that only a few dozen or so variants typically make the cut of being formally disease-associated. By contrast, the performance of a PRS can be improved by

SAM FALCONER

making use of large numbers of variants with much weaker associations, even if the statistical significance falls well below the typical GWAS cut-off. “We have evidence that having even millions of variants within a score seems to improve the predictive ability,” says Ioanna Tzoulaki, an epidemiologist at Imperial College London.

## A new layer of risk

On their own, these scores can achieve reasonable predictive power. For example, an influential 2018 study from a team led by Sekar Kathiresan at Massachusetts General Hospital in Boston used a PRS to identify a group of people in the UK Biobank database with at least threefold greater risk of coronary artery disease<sup>1</sup>. These people made up around 8% of the UK Biobank population.

But because the risk profiles produced by a PRS are distinct from those derived from conventional clinical risk scores, these two approaches can – and should – be combined to achieve even greater utility. A 2020 study<sup>2</sup> of a large Finnish group showed that a PRS could successfully identify individuals at high risk of coronary heart disease who would not have been alerted solely on the basis of their lifestyle or physiological metrics. “This was particularly clear for younger age groups, which is not super surprising since the [conventional] clinical scores were developed with middle-aged cohorts,” says Samuli Ripatti, a geneticist at the University of Helsinki, who led the study. Indeed, 13% of early-onset coronary heart disease cases were predicted only when their PRS was included in the assessment.

Another UK Biobank study<sup>3</sup> also showed that using the two risk-scoring approaches in parallel could potentially capture a larger swathe of individuals at risk for first onset of cardiovascular disease than using either one alone. “We estimate a roughly 7% increase in the number of events prevented if you drop a polygenic score in alongside conventional risk factors,” says Inouye, one of the lead authors of the study. If such screening were performed on the scale of the UK population, it could conceivably prevent tens of thousands of cardiovascular disease events over the course of a decade.

Other studies have found more limited gains in predictive power. For example, a 2020 study of UK Biobank participants by Tzoulaki and her colleagues showed that the addition of a PRS identified only 4.4% more people at high risk who would not have been detected based on gold-standard clinical metrics<sup>4</sup>. “It’s modest, but it could be useful in some settings or in risk stratification tools for further refining certain risk categories,” she says.

This variability can arise from a number of factors, including the algorithm used to generate

the score, the composition of the population used to develop and test the score, and the procedures used to assess the health of individuals in the cohort. This makes it difficult to directly compare individual scores. To address this impediment, researchers from Inouye’s laboratory have teamed up with other European collaborators to develop the Polygenic Score Catalog – an online resource for sharing and comparing PRS data. “It enables reproducible, open science and really helps support external validation of polygenic scores,” says Inouye. His team has also collaborated on a recent *Nature* publication highlighting best practices in PRS development<sup>5</sup>.

## A PRS for the people

Despite this heterogeneity, PRS studies have generally demonstrated the ability to boost the sensitivity of risk assessment for coronary heart disease. Inouye is enthusiastic about the prospect of adding PRSs to the screening arsenal, particularly as genetic testing becomes more commonplace. “We’ve got a potentially very useful risk factor that is not yet being implemented,” he says. “The challenge now is how we deploy it.”

**“We’ve got a potentially very useful risk factor that is not yet being implemented.”**

To build a solid case for PRS implementation, randomized controlled clinical trials must show that they can either safely reduce unnecessary treatment in low-risk populations, or prevent heart attacks by guiding timely intervention for high-risk patients. Early studies give some cause for optimism. In 2017, Natarajan, Kathiresan and their colleagues performed a retrospective analysis of data from a trial of statins for coronary atherosclerosis – a build-up of cholesterol in the arteries that can lead to coronary heart disease. They found that people in the highest risk category, identified by a PRS, experienced greater preventative benefits from treatment<sup>6</sup>. This led to a 44% risk reduction in cardiovascular disease events, compared with a 24% reduction in the remaining population.

PRS results might also convince people to take better care of themselves. In a study of more than 7,000 Finnish people, Ripatti and his colleagues found that individuals who were informed of their high risk for atherosclerosis based on a PRS were 27% more likely to take positive health steps, such as quitting smoking or consulting a physician, than those in lower-risk categories<sup>7</sup>.

But polygenic risk scores are also limited in their utility owing to the underlying ethnic and racial biases that continue to plague the field of genomics. Individuals of European ancestry are heavily over-represented in GWAS research, even for studies performed in countries with diverse populations. “A shockingly low number – something like 5 or 6% of these individuals – are of African ancestry,” says Kullo. In 2020, his team showed that a PRS that delivered strong predictive performance for coronary heart disease in people of European ancestry had much lower predictive power in people of African ancestry<sup>8</sup>. “It’s very uncomfortable for me as a clinician to have to turn to an African American individual and say, ‘I’m not sure how relevant this PRS is to you,’” says Kullo.

The good news is that genetic resources are rapidly growing for once under-represented populations. These include multiple Asian biobank initiatives and the H3Africa initiative, which aims to build genomic research capacity throughout Africa. Kullo is especially excited about the opportunity to collect richer data on African populations, which could essentially offer a Rosetta stone for understanding genetic variation throughout the world. “The out-of-Africa hypothesis says that we are all derived from Africa,” he says, “so if we had genetic markers from African populations, we would potentially have quite powerful pan-ancestry scores.” These could include variants that are rare in European populations but can exert a disproportionately strong influence on cardiovascular disease risk.

The value of a robust and broadly representative PRS could go beyond helping clinicians to manage their patients’ heart health. This approach also offers a road map for a better understanding of the biological processes that ultimately set the stage for cardiovascular disease, enhancing the potential for delivering personalized care. For example, Kullo predicts the possibility of identifying genetic profiles within a PRS that link elevated disease risk specifically to lipid transport, inflammation or other processes that might be selectively targeted with a particular drug. “That’s a whole field of exploration that is awaiting investigators,” he says.

**Michael Eisenstein** is a science writer based in Philadelphia, Pennsylvania.

1. Khera, A. V. et al. *Nature Genet.* **50**, 1219–1224 (2018).
2. Mars, N. et al. *Nature Med.* **26**, 549–557 (2020).
3. Sun, L. et al. *PLoS Med.* **18**, e1003498 (2021).
4. Elliott, J. et al. *J. Am. Med. Assoc.* **323**, 636–645 (2020).
5. Wand, H. et al. *Nature* **591**, 211–219 (2021).
6. Natarajan, P. et al. *Circulation* **135**, 2091–2101 (2017).
7. Widén, E. et al. Preprint at medRxiv <https://doi.org/10.1101/2020.09.18.20197137> (2020).
8. Dikilitas, O. et al. *Am. J. Hum. Genet.* **106**, 707–716 (2020).