THE CAUSATION DETECTOR

A technique called Mendelian randomization has become the go-to for drawing lessons from epidemiological data. But are scientists overdoing it? **By David Adam**

In 1812, the British ophthalmologist James Ware relayed a curious finding to the members of the Royal Society in London. Of thousands of young men recruited to regiments of the British army, only six had been turned away for poor vision in 20 years. But up to one-quarter of students about the same age going to the University of Oxford, UK, relied on a hand glass or spectacles[1].

Ware didn't draw any conclusions about cause and effect: that poring over books might contribute to poor eyesight, for example, or that the bespectacled are naturally drawn to academic pursuits. And just as well. Epidemiologists have long been frustrated by observations that link environmental exposures and health. Myopia is a classic example. Decades of studies show that children who spend the most time at school have the worst eyesight. But the data don't reveal whether schooling makes children myopic or whether myopic kids spend more time at school. Or whether something else, such as socio-economic status, drives both.

Fed up with this logical cul-de-sac, by the turn of this century some epidemiologists had begun suggesting that their field should call it a day. Advances in genetics, they said, could do a better job.

ILLUSTRATION BY ACAPULCO STUDIO

They were half-right. Two decades on, genetics has transformed how people untangle correlation from causation. But it has come to raise epidemiology, not bury it. Genetic differences, it turns out, can help remove confounding variables from analyses, by standing in as proxies for environmental exposure. The technique is called Mendelian randomization.

Scientists have used it to re-evaluate observational data and draw fresh, firmer conclusions on long-standing questions of cause and effect. The analyses have affirmed that low cholesterol levels do not cause cancer[2], for example, that drinking small amounts of alcohol does not protect the heart[3] and that — yes — schooling can make children short-sighted[4].

"Mendelian randomization in principle is a really, really cool idea. It attempts to solve one of the most daunting challenges in epidemiology," says Philipp Koellinger, a social-science geneticist at the Free University of Amsterdam.

Gathering momentum

George Davey Smith, a clinical epidemiologist at the University of Bristol, UK, who helped to pioneer the technique, says: "It came about because we were getting desperate and looking for ways of getting better causal inference in epidemiology." But, he says, there is a downside, too. "The issue is that it became very simple to do."

He has been urging colleagues not to get carried away with Mendelian randomization. It's a powerful tool, but one that must be used properly. As genetic data have piled up, a flurry of Mendelian randomization studies have emerged that don't make the grade. Some have relied on misleading data, and others have failed to sufficiently test the assumptions on which Mendelian randomization relies. It's time, many in the field say, to tighten things up.

Davey Smith was one of the scientists who suggested that epidemiology might have run its course. Writing in an editorial in the *International Journal of Epidemiology*, he and a co-author pointed out that the observational data on the possible harm or benefits of environmental exposures would repeatedly fail when interventions were tested in randomized controlled trials[5].

A few years after the article was published, that point came through loud and clear in the high-profile failure of a US$100-million trial

> **"THE REPUTATION OF EPIDEMIOLOGY WAS COMING UNDER SCRUTINY."**

called SELECT, which found that eating selenium supplements did not protect against prostate cancer — despite mountains of epidemiological evidence suggesting that it would[6].

"It was all rather depressing, and the reputation of epidemiology was coming under scrutiny," Davey Smith says. Researchers had suggested[7] as early as 1986 that genetics could improve the interpretations. But it took the growth of genome-wide association studies (GWAS), which link genetic variants to specific traits, for the approach to gain traction. Last year, Davey Smith turned to Mendelian randomization to revisit the selenium–prostate cancer connection.

Using genotype data for tens of thousands of men, the researchers found almost a dozen genetic variants that were associated with naturally higher levels of selenium in the blood[8]. From birth, these people had lived as if they were taking selenium supplements. The scientists could then compare the incidence of prostate cancer in people with these variants to that in a control group without them. That allowed the researchers to focus more squarely on selenium levels and to ward off the influence of lifestyle factors, such as a healthy diet, that might influence both selenium levels and cancer risk. And, because the tendency to have high or low selenium levels was fixed in DNA, the analysis was less troubled by the likelihood of reverse causation: the possibility that early stages of prostate cancer might influence selenium levels.

The analysis found no benefit from selenium[8], just as the SELECT trial had done[6].

Data bounty

Such results can feed into decisions about whether to launch full clinical trials, Davey Smith argues. And Mendelian randomization can test hypotheses for which it would be unethical or impractical to carry out a trial.

In principle, a Mendelian randomization analysis can be done wherever a genetic variant can be found to naturally mimic the effects of an environmental exposure. And more of those are found every year — especially as millions of people around the world sign up to have their genomes analysed and their health tracked. That gives geneticists the statistical power to identify genetic associations with everything from alcohol consumption to cholesterol levels.

Now, epidemiologists and others are feeding these findings into more Mendelian randomization tests. Data from Scopus and the Web of Science list fewer than 100 papers published per year on the topic in 2010, growing to about 200 by 2015. In 2019, so far, more than 500 papers have used or discussed the method. Researchers have used the tests to tackle a number of questions typically confounded by life's many variables. Studies have helped to show more definitively that drinking alcohol can increase the risk of cancer[9]. Meanwhile, having low cholesterol does not[2], despite some observations to the contrary (see 'Genes as proxy').

As a prime example of how Mendelian randomization can help, many researchers point to myopia, which is a rapidly growing public-health issue. It's been impossible to test its connection to schooling with a randomized controlled trial, because it's unethical to deliberately keep some children out of school.

> "IT DOESN'T TAKE MUCH TO JUST HIT THE BUTTONS AND SAY 'HERE'S A NUMERIC ANSWER'."

A chance to use Mendelian randomization came in 2016, after geneticists published data from two separate GWAS: one looking for genetic signatures related to educational attainment[10]; the other looking for genes associated with myopia[11]. The studies looked at hundreds of thousands of people and found dozens of genetic variants robustly associated with myopia and years spent in school.

The next year, epidemiologists used these variants to explore one of the biggest population data sets around — 488,000 middle-aged and older people who had signed up to the UK Biobank project. Volunteers have their genomes analysed and answer questions on dozens of personal details, including their education and eyesight. When the researchers — at the University of Cardiff and the University of Bristol in the United Kingdom — used Mendelian randomization to analyse the data, they found that being genetically prone to myopia made no difference to how many years people had spent at school. Those who carried the genes associated with educational attainment, however, were significantly more likely to be short-sighted[4].

Whether through time spent reading, lower levels of natural light or some other factor, time spent in school clearly influences vision, says Denize Atan, an ophthalmologist at Bristol who led the project. The link is so strong, she says, that policymakers and schools should do more to address it.

The open secret

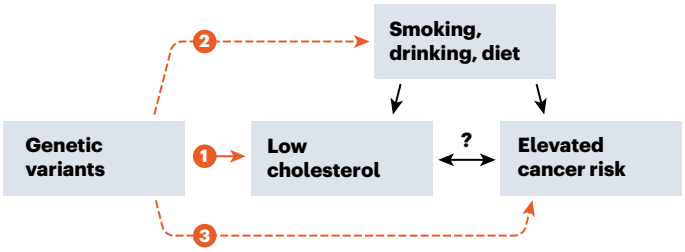
The problem is, critics argue, that not all Mendelian randomization studies are as sound. "You need to have a robust hypothesis and some supporting evidence before you start," Atan says. A growing number, she says, do not. "You think, 'Where did they get that idea from?' It just seems to come out of the blue.'" This is a big problem, she adds, because Mendelian randomization allows researchers to seek, find and publish relationships between unfamiliar data sets without any specialist knowledge of the relevant field.

As Sonja Swanson, an epidemiologist at the Erasmus University Medical Center in Rotterdam, the Netherlands, puts it: "It doesn't take much to just hit the buttons and say 'here's a numeric answer to my question'."

Several epidemiologists say that it's an open secret in the field that many published Mendelian randomization studies are problematic. "You can get papers published very easily," says Davey Smith. "Some of the very poor papers are from researchers who don't understand epidemiological principles." In 2016, a Mendelian randomization study claimed to have found that high blood levels of C-reactive protein, a liver enzyme associated with inflammation, caused schizophrenia. It suggested that drugs capable of lowering levels of the enzyme in the blood might help to treat people with the disease[12]. Davey Smith's group and

GENES AS PROXY

Mendelian randomization uses gene variants to interrogate cause and effect in epidemiological data. For example, researchers can ask whether having low cholesterol increases cancer risk — as some data suggest — by looking at people genetically predisposed to low cholesterol. This strategy can help rule out reverse causation (that cancer lowers cholesterol), and it can bypass some variables that might influence both cancer risk and cholesterol. However, it relies on several assumptions, which must be tested.



Assumption 1
The link between the variants and having low cholesterol has to be strong and stable over time.

Assumption 2
The variants must not influence variables that might affect both cholesterol levels and cancer risk.

Assumption 3
The variants must not be associated with cancer risk in any way other than through the relationship to cholesterol.

another did similar analyses, and found that, in fact, C-reactive protein had a protective effect against schizophrenia[13]. Davey Smith and his co-authors suggested that there were issues with how the initial group had combined genetic data sets, and the 2016 paper was eventually retracted.

Epidemiologists have also criticized[14] Mendelian randomization analyses that claimed to have found that smoking while pregnant causes dramatic declines in baby birthweights[15] and substantially increases the risk of orofacial clefts in offspring[16].

The problem, they say, is that the genetic variants used as proxies for smoking behaviour were identified in what are known as candidate-gene studies, in which researchers evaluate a few genes that they suspect are involved in a behaviour such as smoking. Results from such studies can be unreliable because they are biased towards finding some effects in the genes being examined. The variants that the authors used in their Mendelian randomization haven't shown up in larger, more comprehensive GWAS.

George Wehby, a health-policy researcher at the University of Iowa in Iowa City, who led the smoking projects, says that the work was done before better data were available. "I agree that these wouldn't be the first choice," he says, "given current knowledge about genetics of smoking from large GWAS."

Against common sense

To an economist, Mendelian randomization looks a lot like something called instrumental variable analysis, in which a variable referred to as the instrument is used to help unpick hidden relationships between two other observations. "When we saw that epidemiologists were using genes as instrumental variables, we were both intrigued and said, 'Wait a moment!'," says Koellinger. Such analyses are built on assumptions that need to be carefully scrutinized.

One central assumption in Mendelian randomization is that the genetic variants must not affect the outcome in any other way. For example, there is a variant of the gene that encodes the enzyme aldehyde dehydrogenase (ALDH2) that disrupts metabolism of alcohol. When people with this variant drink, they tend to feel nauseous, and so it's associated with lower levels of alcohol consumption. That might seem a plausible way to test, for example, whether drinking raises blood pressure, because those who carry the variant generally drink less than those who don't.

The problem is that ALDH2 also influences how likely someone is to smoke[17], which independently influences blood pressure. This phenomenon, known as genetic pleiotropy, can invalidate Mendelian randomization results. And that creates a problem, because the extent of pleiotropy isn't fully realized for many genes.

Another assumption is that a given genetic variant has a strong effect. As bigger and more powerful GWAS dredge up weaker genetic links to different traits, this assumption becomes harder to test.

A 2015 review by epidemiologists in the Netherlands of 178 published Mendelian randomization studies found that fewer than half adequately discussed these assumptions[18]. "As these assumptions are crucial for the validity of Mendelian randomization studies, they should always be discussed in the specific context of the study," the researchers argue[18].

Mendelian randomization is also subject to a distinct source of bias — one that's a matter of life and death. People can die only once. This issue would complicate, for example, an analysis of deaths from stroke. Such deaths tend to occur in older people, so a study of strokes will typically recruit people who have already survived conditions that affect younger people, such as heart disease. Because stroke and heart disease have common causes, such as high cholesterol (and therefore common therapies, including statins), this survivor bias can throw up some misleading results.

To demonstrate the effects of this bias, Mary Schooling, a public-health epidemiologist at the City University of New York, ran Mendelian randomization tests in which gene variants linked to reduced cholesterol stood in for statin use. People with this beneficial inheritance have fewer heart attacks early in life and live to an age at which the risk of stroke rises. So, the study concluded that cholesterol-lowering statins would actually cause strokes[19].

"It didn't make any sense," Schooling says. Proper randomized controlled trials aren't confused in this way: they show that statins protect against stroke. But Mendelian randomization shows a survivor bias that must be identified and corrected for.

Battling biases

"Every single method can be biased," says Davey Smith. Mendelian randomization, he says, is not intended to replace randomized controlled trials, but, alongside other sources, including observational studies, they can add to the evidence available to help make an informed decision. Now, researchers are looking for ways to improve them.

One way is to identify and correct for some of the biases, and to apply statistical tools for testing the strength of the assumptions. Davey Smith points to papers that can help researchers to assess the quality of Mendelian randomization studies for themselves[20].

Better organization of data can help, too. Unbiased analyses assume that genes are randomly distributed, but some genes are known to cluster in geographical regions[21]. Already, genotype data sets are becoming available that are grouped by extended families, and Mendelian randomization studies of these data are identifying, for example, that height and body mass index might not influence educational attainment as much as previous studies had suggested[22].

By comparing the results of within-family and population-based studies, geneticists can help to distinguish the roles of nature and nurture in a given trait. "Particular genes are correlated with particular features of the local environment. And if you want to use genes for causal inference you need to break that link," says Koellinger.

This kind of accuracy is important if researchers want to harness the growing torrent of genetic information for public-health and policy recommendations. But even these tools need to be improved and supplemented.

Ware's observations 200 years ago on the eyesight of students and soldiers have been explained through a genetic lens that no one could have imagined at the time. Ironically, it took the British army another century to accept recruits who need to wear glasses, and to change its standards for what it considers adequate vision. Even during the First World War, some authorities argued it didn't matter if a British soldier couldn't see clearly what he was shooting at, as long as he could "fire in the right direction"[23].

Statistical tools for epidemiology are improving. And although Mendelian randomization does not always offer perfect clarity, it might, at least, point researchers in the right direction.

David Adam is a freelance journalist based near London.

1. Ware, J. *Phil. Trans. R. Soc. London* **103**, 31–50 (1813).
2. Benn, M., Tybjærg-Hansen, A., Stender, S., Frikke-Schmidt, R. & Nordestgaard, B. G. *J. Natl Cancer Inst.* **103**, 508–519 (2011).
3. Millwood, I. Y. *et al. Lancet* **393**, 1831–1842 (2019).
4. Mountjoy, E. *et al. Br. Med. J.* **361**, k2022 (2018).
5. Davey Smith, G. & Ebrahim, S. *Int. J. Epidemiol.* **30**, 1–11 (2001).
6. Lippman, S. M. *et al. J. Am. Med. Assoc.* **301**, 39–51 (2009).
7. Katan, M. B. *Lancet* **327**, 507–508 (1986).
8. Yarmolinsky, J. *et al. J. Natl Cancer Inst.* **110**, 1035–1038 (2018).
9. Pierce, B. L., Kraft, P. & Zhang, C. *Curr. Epidemiol. Rep.* **5**, 184–196 (2018).
10. Okbay, A. *et al. Nature* **533**, 539–542 (2016).
11. Pickrell, J. K. *et al. Nature Genet.* **48**, 709–717 (2016).
12. Inoshita, M. *et al. Sci. Rep.* **6**, 26105 (2016).
13. Hartwig, F. P., Davies, N. M., Hemani, G. & Davey Smith, G. *Int. J. Epidemiol.* **45**, 1717–1726 (2017).
14. Taylor, A. E. *et al. Econ. Hum. Biol.* **13**, 99–106 (2014).
15. Wehby, G. L. *et al. Biodemography Soc. Biol.* **57**, 3–32 (2011).
16. Wehby, G. L. *et al. Health Serv. Outcomes Res. Methodol.* **11**, 54–78 (2011).
17. Masaoka, H. *et al. Drug Alcohol Depend.* **173**, 85–91 (2017).
18. Boef, A. C. G., Dekkers, O. M. & le Cessie, S. *Int. J. Epidemiol.* **44**, 496–511 (2015).
19. Schooling C. M. *et al.* preprint at bioRxiv https://doi.org/10.1101/716621 (2019).
20. Davies, N. M., Holmes, M. V. & Davey Smith, G. *Br. Med. J.* **362**, K601 (2018).
21. Abdellaoui, A. *et al. Nature Hum. Behav.* https://doi.org/10.1038/s41562-019-0757-5 (2019).
22. Brumpton, B. *et al.* Preprint at BioRxiv https://doi.org/10.1101/602516 (2019).
23. Cubitt, B. B. *Br. J. Ophthalmol.* **2**, 35–40 (1918).