The New York Times | https://nyti.ms/1hiVUgv

SCIENCE

Linking Genes to Diseases by Sifting Through Electronic Medical Records

Carl Zimmer

MATTER NOV. 28, 2013

The days of scrawled doctor's notes are slowly coming to a close. In the United States, 93 percent of hospitals are now using at least some electronic medical records and 2.2 percent have given up paper records completely, according to the consulting firm HIMSS Analytics.

The federal government has been pushing for electronic medical records for a decade, arguing that they will improve health care and bring down costs. That is still a matter of debate. Critics charge that the system is hobbled by poorly designed software and that some hospitals are using electronic medical records to bill more for the same services.

But a new study suggests that electronic medical records may have another, entirely different use: as a Rosetta Stone for our DNA. Researchers are using them to trace links between genes and disease.

It has been 13 years since scientists first published the rough draft of the

human genome and yet they are still just beginning to work out how our DNA influences our health. Most insights in recent years have come from so-called genome-wide association studies.

To run such a study, scientists find a lot of people with the same condition — diabetes, for example, or Alzheimer's disease. Then they take samples of DNA from the subjects and search them for mutations that are unusually common in people with the disease compared with people without it.

Since 2005, scientists have carried out more than 1,500 genome-wide association studies, discovering thousands of links between gene variants and various conditions. But these studies have only put a dent in the genome's mysterious complexity.

Many of the gene variants they have uncovered have only a tiny influence on the risk of getting a particular disease. And when scientists have replicated genome-wide association studies, some links between genes and diseases have faded away.

A study published this week in Nature Biotechnology opens up a new way to search for these links: by turning genome-wide associations on their head.

In the new approach — called phenome-wide association studies — scientists start with a gene variant and then search among thousands of conditions for a match.

To find those matches, scientists comb through electronic medical records.

The new study was carried out by scientists from a consortium of medical research institutions. Known as the Electronic Medical Records and Genomics Network — eMERGE for short — it was founded in 2007 and includes institutions like the Mayo Clinic and Vanderbilt University School of Medicine.

The eMERGE network has developed software to extract and analyze information from their electronic medical records. Those records rarely include

genetic information, but the researchers have also been setting aside blood samples from patients for genetic studies. After making the two sets of information anonymous, the researchers combined them into a single database.

The eMERGE team has now taken the database out for a test run. Looking at previously published genome-wide association studies, they identified 77 gene variants with strongly supported links to diseases. The scientists then tried to replicate the results.

To do so, they started with the gene variants and looked through the electronic medical records of their patients for any unusually common conditions. “We have everything they come to the doctor for,” said Dr. Joshua C. Denny, a biomedical informatics researcher at the Vanderbilt University School of Medicine and a co-author of the new study.

In 51 out of 77 cases, Dr. Denny and his colleagues ended up with the same link. With 1,358 different conditions to choose from in the electronic medical records, it was practically impossible for them to do so well simply by chance.

“It’s a phenomenal proof of concept,” said Robert C. Green, a geneticist at Harvard Medical School. The new study has persuaded him that electronic medical records are a reliable source of scientific information — even though they weren’t originally designed with such research in mind. “Warts and all, we can use them to do science,” Dr. Green said.

The eMERGE team then searched for new links of their own. They surveyed 3,144 gene variants identified in past genome-wide association studies to see if they also influenced any other diseases. They discovered 63 new links to diseases, ranging from skin cancer to anemia.

This finding also has exciting potential. Scientists have long known that a single gene can have a variety of effects on our health. That is because genes are biological multitaskers, helping carry out different functions in the body.

But genome-wide association studies — which are designed to examine a

single condition — have done a poor job of uncovering these effects. Electronic health records, the new study suggests, have the power to reveal them.

Such studies could help link seemingly unrelated symptoms, Dr. Denny suggested. They could also point to potentially harmful side effects of a drug. On the other hand, they could also guide research into new uses for drugs. “If you have a drug that targets a certain gene, you can understand what range of diseases you can use that drug to treat,” said Dr. Denny.

The eMERGE database now has information on about 51,000 patients and the researchers are currently expanding their phenome-wide association studies accordingly. Daniel MacArthur, a geneticist at Massachusetts General Hospital who was not involved in the new study, predicted that larger databases of electronic health records will allow scientists to find links between diseases and rare gene variants that are so far impossible to detect. “This is exciting stuff,” he said.

Even larger studies are on the horizon. The Department of Veterans Affairs, for example, has set up the Million Veteran Program, which will combine electronic medical records and a DNA database from volunteers. They now have 240,000 veterans enrolled.

“Unfortunately,” said Dr. MacArthur, “applying this on a larger scale in the U.S. is complicated by the dire state of electronic medical records.”
