

Mining Data from Electronic Records: Faster Way to Get Genetic Clues to Disease



Recruiting thousands of patients to collect health data for genetic clues to disease is expensive and time consuming. But that arduous process of collecting data for genetic studies could be faster and cheaper by instead mining patient data that already exists in electronic medical records, according to new Northwestern Medicine research.

In the study, researchers were able to cull patient information in electronic medical records from routine doctors' visits at five national sites that all used different brands of medical record software. The information allowed researchers to accurately identify patients with five kinds of diseases or health conditions -- type 2 diabetes, dementia, peripheral arterial disease, cataracts and cardiac conduction. "The hard part of doing genetic studies has been identifying enough people to get meaningful results," said lead investigator Abel Kho, M.D., an assistant professor of medicine at Northwestern University Feinberg School of Medicine and a physician at Northwestern Memorial Hospital. "Now we've shown you can do it using data that's already been collected in electronic medical records and can rapidly generate large groups of patients."

The paper is published April 20 in *Science Translational Medicine*. To identify the diseases, Kho and colleagues searched the records using a series of criteria such as medications, diagnoses and laboratory tests. They then tested their results against the gold standard -- review by physicians. The physicians confirmed the results, Kho said. The electronic health records allowed researchers to identify patients' diseases with 73 to 98 percent accuracy.

The researchers also were able to reproduce previous genetic findings from prospective studies using the electronic medical records. The five institutions that participated in the study collected genetic samples for research. Patients agreed to the use of their records for studies. Sequencing individuals' genomes is becoming faster and cheaper. It soon may be possible to include patients' genomes in their medical records, Kho noted. This would create a bountiful resource for genetic research.

"With permission from patients, you could search electronic health records at not just five sites but 25 or 100 different sites and identify 10,000 or 100,000 patients with diabetes, for example," Kho said. The larger the group of patients for genetic studies, the better the ability to detect rarer affects of the genes and the more detailed genetic sequences that cause a person to develop a disease.

The study also showed across-the-board weaknesses in institutions' electronic medical records. The institutions didn't do a good job of capturing race and ethnicity, smoking status and family history, all which are important areas of study, Kho said. "It shows we need to focus our efforts to use electronic medical records more meaningfully," he added. The institutions participating in the study are part of a consortium called the Electronic Medical Records and Genomics Network.

The research was supported by the National Human Genome Research Institute with additional funding from the National Institute of General Medical Sciences.

The above story is reprinted (with editorial adaptations by ScienceDaily staff) from materials provided by Northwestern University.

Journal Reference:

A. N. Kho, J. A. Pacheco, P. L. Peissig, L. Rasmussen, K. M. Newton, N. Weston, P. K. Crane, J. Pathak, C. G. Chute, S. J. Bielinski, I. J. Kullo, R. Li, T. A. Manolio, R. L. Chisholm, J. C. Denny. Electronic Medical Records for Genetic Research: Results of the eMERGE Consortium. *Science Translational Medicine*, 2011; 3 (79): 79re1 DOI: 10.1126/scitranslmed.3001807

www.sciencedaily.com

Published on : Mon, 2 May 2011