EHR-Based Research Yields Insights on CKD

By Bridget M. Kuehn

Patients with chronic kidney disease who also have chronic obstructive pulmonary disease (COPD) have a 41% increased risk of death, according to a recently published study that relied on electronic health records (EHRs) (Navaneethan SD, et al. *Am J Nephrol* 2016; 43:39–46).

The finding is part of a growing body of evidence demonstrating the power of EHR-based studies to help elucidate the many factors that contribute to poor outcomes for patients with CKD. The technology is also being used to help test ways to improve their care.

Joseph Nally, MD, a coauthor of the COPD study and Director of the Center for Chronic Kidney Disease at the Cleveland Clinic, and his colleagues began exploring electronic data for CKD research nearly a decade ago. There were a few "false starts" using billing data before the multidisciplinary team eventually created a CKD registry using EHRs, Nally said. Their registry now includes 110,000 patients, he said.

One of the reasons EHRs are particularly useful for CKD research is that objective laboratory values found in EHRs can identify patients even if they haven't been diagnosed, said Jesse D. Schold, PhD, of the Cleveland Clinic's Quantitative Health Sciences Department.

"That allows you to identify patients who haven't been recognized through other traditional means," Schold said.

This is particularly important in CKD patients because only about 10% have CKD listed as a condition in their records (Jolly SE, et al. *Am J Nephrol* 2014; 39:288–96). Patients who aren't listed as having CKD are less likely to see a nephrologist and receive recommended testing, and are more likely to receive contraindicated medications, Nally noted.

Another factor that aids EHR-based kidney disease research is the ability to access data from the United States Renal Data System, which provides a census of all patients with end stage renal disease in the US, Schold said. Nally noted that his team also taps data from the Ohio Death Index, which can provide cause-specific deaths for CKD patients (Navaneethan SD, et al. *J Am Soc Nephrol* 2015; 26:2512–2520).

Having a CKD registry also makes study recruitment more efficient. "The CKD registry is a spectacular vehicle for recruitment into research studies and randomized controlled trials," Nally said.

Currently, Nally and his team are using their registry for continuous quality improvement efforts (CQI).

For example, they have compared the outcomes of 500 CKD patients who receive care in a CKD clinic, which uses nurse practitioners and algorithm-based care, with 1500 matched CKD patients receiving standard nephrology care. So far, the CKD clinics have performed better on patient processes of care and patient education, but whether it translates into better outcomes is unknown, Nally noted.

The team is also exploring the use of EHR-linked technologies for patient engagement through an ongoing randomized trial funded by the National Institute of Diabetes and Digestive and Kidney Diseases. The Cleveland Clinic's CKD-enhanced patient portal allows patients to access their lab results, request appointments or prescriptions, or send messages to their clinicians. It also provides CKD-stage–specific information. For example, stage 4 patients may receive information about managing anemia or dialysis options.

"There is momentum in health care to facilitate transparency, which allows for joint decision-making [among] patients, caregivers, and providers," Schold said. "The proliferation of [electronic] data and technology will facilitate that and [help us] understand best practices." ●

Commercial Online Health Data Research: Weighing Privacy Concerns against Potential for Medical Insights

By Bridget M. Kuehn

Search engines are one of the first places many Americans turn when looking for health information, according to a 2013 survey by the Pew Research Center. But what they may not know is that the data from these searches is collected by the search engine and is increasingly being used for health research and public health surveillance.

The data has enormous potential to help researchers better understand pressing public health issues and perhaps even to identify individuals at risk of developing serious disease. But this emerging venue for health research also poses new questions about what constitutes consent for research use of online health information and what role corporations, who own the data, should play in the process.

"Innovation is crucial in our world, and these approaches that have shown promise should be pursued if we develop appropriate methods to ensure the benefit to society and individual patients," said Mauricio Santillana, PhD, a member of the faculty at Boston Children's Hospital and an associate at Harvard's Institute for Applied Computational Science.

Emerging field

Epidemiologists have been at the leading edge of using search data, often combined with data from electronic health records or social media sites, Santillana said. His group at Harvard University has partnered with Google to use its data for tracking and forecasting epidemics of the flu and other infectious diseases.

While initial attempts to develop a Google search–based flu-tracking system were stymied, methods have improved substantially since then (Yang S, et al. *Proc Natl Acad Sci USA* 2015; 112:14473–14478). Now, Santillana and his colleagues can produce very accurate outbreak estimates in real-time and accurate forecasting of flu trends about 1 to 2 weeks ahead.

"The field has evolved quite a bit," Santillana said. "We basically show data from Google searches may be noisy and may not be straightforward to interpret, but by developing robust methods we can minimize the effect of the noise and produce accurate forecasts."

Other types of research are also being explored that are more longitudinal and focus on individuals. For example, researchers from Microsoft recently showed that Bing search data might be useful to identify individuals with symptoms of pancreatic cancer even before diagnosis (Paparrizos J, et al. *J Oncol Pract.* pii: JOPR010504 [Published online June 7, 2016]). Such early identification might help improve patient outcomes, because many patients with pancreatic cancer receive diagnoses too late to be treated effectively, wrote lead author John Paparrizos, MSc, a computer scientist at Columbia University and his colleagues from Microsoft.

"The results highlight the promise of using Web search logs as a new direction for screening for pancreatic carcinoma," the authors wrote.

Privacy and oversight

But concerns have been raised about protecting individuals' privacy and the oversight of online health data research.

"People have a very different sense of privacy around their medical data," said Elizabeth Buchanan, PhD, an ethicist at the University of Wisconsin-Stout in Menomonie, Wisconsin.

They may also have different expectations for privacy depending on whether they are posting health information on a social media site or whether they are conducting a search, Buchanan said. Most companies' terms-of-use policies outline that user data will be logged and possibly used for research or other purposes, including commercial ones, Buchanan said.

"We should be aware that third party apps are collecting, repackaging, and repurposing our data whether it is posted in a public space or if it is something we consider more private like a search query," she said.

It's important to be aware of how this data might be used in ways that are beneficial, for example, for disease surveillance or for patient outreach, while also understanding the ways that composite online health data might be used to identify an individual or even used to harm them, Buchanan said.

"The promise of personalized medicine and predictive analytics is that it can help," she said. "But we want to be careful of the larger more dangerous uses of these kinds of data," she said.

Santillana said his group protects the privacy of searchers' health information by using aggregate data and trying to ensure that individuals can't be re-identified through the data.

"We do our best to maintain the anonymity of the population we are trying to help," he said.

Other potentially promising uses that track an individual's search behavior may trigger greater public concerns about privacy, Santillana said. For example, what if insurance companies got access to the information and used it to refuse to sell the person insurance?

"If a patient is identified as likely to get a diagnosis