Assessing the accuracy of cardiovascular data in electronic health records

August 4 2017, by Anna Williams

In a new study published in the journal *Circulation*, Northwestern Medicine investigators identified similarities and differences between cardiovascular data pulled from electronic health records (EHRs) and data collected in a traditional cohort study.

The findings illustrate both the potential benefits and limitations of using clinical EHR data for epidemiological research, especially as EHRs become standard in healthcare systems and are increasingly leveraged for major research programs, such as the "All of Us" precision medicine initiative.

"Electronic health records are going to be a valuable resource for research as we go forward, but no one has really looked at how reliable and valid the data that is being included is," said senior author Norrina Allen, PhD, MPH, assistant professor of Preventive Medicine in the Division of Epidemiology. "This study is exciting because it's the first time that we've been able to look at the accuracy of EHRs, and examine how we can leverage their strengths."

Faraz Ahmad, MD, MS, a fellow in advanced heart failure and transplant cardiology, was the first author of the study.

To explore the validity of the two sources, the investigators examined the health data of individual patients who were included in both HealthLNK—an electronic database of close to 3 million Chicago area residents from six large health systems—and the Multi-Ethnic Study of

Atherosclerosis (MESA), a traditional cohort study.

Because the MESA data was collected specifically for the purposes of cardiology research, and followed a rigorous protocol, the investigators could employ it as a "gold standard" with which to compare the clinical EHR data.

The investigators specifically compared measures of patient demographics, cardiovascular risk factors—such as blood pressure and BMI—and cardiovascular events.

They discovered that the EHRs had a significant amount of missing data for basic demographic characteristics, such as ethnicity, as compared to MESA.

After excluding the missing data, however, the investigators found good agreement between HealthLNK and MESA on gender, race/ethnicity, age and BMI measurements. They discovered differences in the two databases when it came to individuals' blood pressure measures, prevalence of risk factors and cardiovascular events.

"We have to understand the limitations of EHRs and take advantage of its strengths in order to increase our understanding of heart disease," Allen explained. "We now know how good the data are and we know what type of research we can use it for, and we can use that to ask questions of the data. This allows us to expand our research into electronic health records."

The authors also note the need for continued research to understand the quality of data from electronic records—especially across different health systems and environments—and how to use diverse data sources to improve detection and classification.

"Our findings suggest that combining traditional methods of epidemiological research with electronic health record data may lead to a more complete picture of an individual's health over time," said Ahmad, who also conducted the research under faculty mentor and co-author Abel Kho, MD, MS, associate professor of Medicine in the Division of General Internal Medicine and Geriatrics and of Preventive Medicine in the Division of Health and Biomedical Informatics, and director of the Center for Health Information Partnerships (CHiP). "Our next step is to expand the linkage between traditional cohorts and electronic data research networks to other sites across the country."

More information: Faraz S. Ahmad et al. Validity of Cardiovascular Data From Electronic Sources:The Multi-Ethnic Study of Atherosclerosis and HealthLNK, *Circulation* (2017). DOI: 10.1161/CIRCULATIONAHA.117.027436

Provided by Northwestern University