Researchers use EHRs to identify cancer symptom clusters

October 22 2019

Kun Huang, Ph.D. of Regenstrief Institute and Indiana University School of Medicine and Xiao Luo, Ph.D.of IUPUI conducted a study using electronic health records to identify cancer symptom clusters. The work earned the Best Paper Award at the 10th Association for Computing Machinery (ACM) Conference on Bioinformatics, Computational Biology and Health Informatics Credit: Regenstrief Institute

Patients with chronic diseases such as breast cancer or colorectal cancer

often experience fatigue, pain, depression and other symptoms which can lead to distress and functional impairment if left untreated. With the ultimate goal of helping clinicians manage and treat symptoms that negatively affect health and quality of life, researchers from the Regenstrief Institute and IUPUI have devised and tested novel methodologies to extract data on symptoms from electronic health records (EHRs) and have successfully investigated associations between symptom clusters and disease.

"Our novel methods can be generalized beyond breast and colorectal cancer to analyze symptom clusters of other chronic diseases where symptom management and treatment is critical," said Regenstrief Institute investigator Kun Huang, Ph.D., senior author of the study and an internationally recognized leader in translational bioinformatics. "Identifying and understanding symptom clusters—which symptoms tend to go together—fatigue and depression, for example—and when these symptoms occur during the course of treatment—provides critical information to a patient's care team, especially as we look forward to precision health and try to find the right treatment for the right patient at the right time.

"This information is also important to researchers as they explore whether biological reasons behind these co-occurrences exist and, if so, determine why," Dr. Huang said. In addition to his appointment as a Regenstrief Institute research scientist, he is assistant dean for data sciences, Precision Health Initiative professor of genomics data sciences and professor of medicine at Indiana School of Medicine.

"Identifying Symptom Clusters in Breast Cancer and Colorectal Cancer Patients using EHR Data," won the Best Paper Award at the 10th Association for Computing Machinery (ACM) Conference on Bioinformatics, Computational Biology and Health Informatics and is published in the conference proceedings. This prestigious prize is

presented for a paper that represents groundbreaking research. Through the Best Paper Award, the ACM highlights theoretical and practical innovations likely to shape the future of computing.

Detecting and tracking symptom severity and frequency in an EHR are not easy. Language often varies. For example, fatigue may be recorded in an EHR as listless, exhausted, tired, no energy or worn out. Depression may be indicated as sad, unhappy, hopeless or discouraged. To mitigate the terminology problem, the research team designed methodology to enable the analysis of free text clinical notes in EHRs in addition to extracting information from structured entries in the medical record.

"EHR data has not been extensively used to understand patient-reported symptoms for individuals with chronic diseases," said study author Xiao Luo, Ph.D., assistant professor of computer and information technology in IUPUI's School of Engineering and Technology. "Utilizing EHR data obtained from the Indiana Network for Patient Care, we developed a framework that employs components of data mining, NLP (natural language processing) and machine learning to explore clinical information accumulated throughout the course of these diseases," she said. Dr. Luo, who designed the study's methodology, specializes in predictive data analytics and intelligent data integration.

The researchers looked at symptoms during two time periods—the first year of chemotherapy and the 48 to 54 months following chemotherapy. Age, smoking history and whether the individual had diabetes were also considered. Symptom clusters were identified based on severity and combination of symptoms.

Using algorithms they developed, the researchers found that the symptom clusters were not the same for breast cancer patients and colorectal cancer patients. They also noted varying symptom clusters at

different time frames following chemotherapy. Breast cancer patients had slightly more symptoms than colorectal cancer patients during the first year after chemotherapy. Colorectal cancer patients had slightly more depression 48 to 54 months after chemotherapy. The researchers also were able to identify linkages. For example, if a high-stage colorectal cancer patient had no fatigue symptom, he or she was also unlikely to experience depression.

Cancer patients present with multiple, co-occurring symptoms related to the disease and its treatment. "Typically clinicians identify, measure and treat patient's symptoms such as—fatigue, depression, anxiety—individually, rather than as co-occurring" said study author Susan Storey, Ph.D., R.N., assistant professor at IU School of Nursing and a member of the Indiana University Melvin and Bren Simon Cancer Center and the Diabetes Translational Research Center. "This new methodology sets us firmly on the path to better understand symptoms, the way they cluster and how they are related to disease outcomes and quality of life. As we expand this work it will have significant implications for patients and the clinicians who treat them because it will help us understand why and how these symptom clusters interact with each other in order to develop strategies to counter them." Dr. Storey has 25 years of experience as an oncology certified advance practice nurse.

Provided by Regenstrief Institute