Study examines use of a natural language processing tool for electronic health records in assessing colonoscopy quality

June 21 2012

A new study shows that natural language processing programs can "read" dictated reports and provide information to allow measurement of colonoscopy quality in an inexpensive, automated and efficient manner. The quality variation observed in the study within a single academic hospital system reinforces the need for routine quality measurement. The study appears in the June issue of GIE: *Gastrointestinal Endoscopy*, the monthly peer-reviewed scientific journal of the American Society for Gastrointestinal Endoscopy (ASGE).

Gastroenterology specialty societies have advocated that providers routinely assess their performance on colonoscopy quality measures. Such routine measurement has been hampered by the costs and time required to manually review colonoscopy and pathology reports. Natural language processing (NLP) is a field of computer science in which programs are trained to extract relevant information from text reports in an automated fashion.

"Routine measurement is not taking place, primarily because of the inconvenience and expense. Measuring adenoma detection rates and other quality measures typically requires manual review of colonoscopy and pathology reports. To address the difficulty in measuring physician quality, we developed the first NLP–based computer software application for measuring performance on colonoscopy quality indicators," said study lead author Ateev Mehrotra, MD, MPH,

University of Pittsburgh, School of Medicine. "Our study highlights the potential for NLP to evaluate performance on colonoscopy quality measures in an inexpensive and automated manner. This type of routine quality measurement can be the foundation for efforts to improve colonoscopy quality."

Colonoscopy is a cost-effective and common method of screening for colorectal cancer. However, colonoscopy may be imperfect in screening because, among other reasons, physicians miss adenomas, the precursors to colorectal cancer. There is great variation among physicians in the proportion of colonoscopies in which an adenoma is found as well as variations in other aspects of colonoscopy quality. This has led gastroenterology specialty societies to call for physicians to regularly monitor their performance on colonoscopy quality measures so that care can be improved.

Methods

The researchers' objective was to demonstrate the potential applications for and the efficiency of NLP-based colonoscopy quality measurement. In a cross-sectional study design, they used a previously validated NLP program to analyze colonoscopy reports and associated pathology notes. The resulting data were used to generate provider performance on colonoscopy quality measures. Nine hospitals in the University of Pittsburgh Medical Center health care system participated in the study. The study sample consisted of 24,157 colonoscopy reports and associated pathology reports from 2008 to 2009. Main outcome measurements were provider performance on seven quality measures: American Society of Anesthesiologsts (ASA) classification indicated; informed consent documented; quality of bowel preparation described; cecal landmarks noted; adenoma detection; withdrawal time documented; and biopsy taken for chronic diarrhea.

Results

Performance on some colonoscopy quality measures was poor, while others were at benchmark levels, and there was a wide range of performance. Across hospitals, the adequacy of preparation was noted overall in only 45.7 percent of procedures (range 14.6 percent-86.1 percent across nine hospitals), cecal landmarks were documented in 62.7 percent of procedures (range 11.6 percent-90.0 percent), and the adenoma detection rate was 25.2 percent (range 14.9 percent-33.9 percent).

The researchers concluded that the study results highlight the potential of NLP to measure performance on colonoscopy quality measures. The NLP tool efficiently analyzed a large sample of colonoscopy reports. They stated that the findings demonstrate that there is clear variation in performance, even within a highly regarded academic health care system. Across the nine hospitals, there was almost a threefold variation in the adenoma detection rate. The variation in performance for other quality measures among physicians was even greater in some cases.

The researchers noted several very important study limitations. Because it is limited to a single hospital system and there is variation in the manner that physicians record colonoscopy reports, it is likely that the NLP tool would need to be adapted to the reporting style and language used by other physicians to achieve comparable performance in another setting. Consistent with previous studies, they did not adjust provider scores for differences in patient population. This type of risk adjustment will have to be considered, especially if beginning to profile physicians who subspecialize and who treat patients with clearly different adenoma detection rates (such as patients with inflammatory bowel disease or younger patients).

In an accompanying editorial, John C. Deutsch, MD, Department of

Gastroenterology and Cancer Center, Essentia Health Systems, Duluth, Minn., stated, "I commend the authors for their efforts at trying to develop a method for extracting quality measures from endoscopy and pathology reports. Using a computer to retrieve information based on language and histology seems like a valuable endeavor."

Provided by American Society for Gastrointestinal Endoscopy