873. Using natural language processing on electronic medical notes to detect the presence of an indwelling urinary catheter

Adi Gundlapalli, MD, PhD, MS[1]; Guy Divita[2,3]; Tyler Forbush[3]; Andrew Redd PhD[3]; Marjorie Carter, MSPH[3]; Ashley J. Gendrett, BS, MPH[4]; Kalpana Gupta, MD, MPH[5]; Ying Suo, MS[2]; B.S. Begum Durgahee, MS[3]; Sarah Krein, PhD, RN[6]; Michael Rubin, MD, PhD[2]; Anne Sales, PhD, RN[7]; Matthew Samore, MD[8]; Barbara W. Trautner, MD, PhD[9]; [1]University of Utah School of Medicine and Salt Lake City VA Health Care System, Salt Lake City, UT; [2]Internal Medicine, University of Utah School of Medicine, Salt Lake City, UT; [3]Ideas Center, VA Salt Lake City Health Care System, Salt Lake City, UT; [4]Houston Center for Innovations in Quality, Effectiveness, and Safety, Michael E. Debakey Veterans Affairs Medical Center, Houston, TX; [5]Department of Medicine/ Boston University School of Medicine, Boston, MA; [6]Department of Internal Medicine, Division of General Medicine, University of Michigan Medical School, Ann Arbor, MI; [7]Division of Nursing Business and Health Systems, University of Michigan, School of Nursing, Ann Arbor, MI; [8]University of Utah School of Medicine, Division of Epidemiology, Salt Lake City, UT; [9]Section of Infectious Diseases, Department of Medicine, Baylor College of Medicine, Houston, TX

Session: 111. Device-Associated HAIs
Friday, October 10, 2014: 12:30 PM

Background. We set out to develop a natural language processing (NLP) algorithm to extract concepts related to indwelling urinary catheters from electronic medical notes.

Methods. A concept lexicon was developed based on domain knowledge, prior expertise and review of medical notes. Concepts were classified as evidence of either catheter presence or catheter absence. A reference standard set of 1595 randomly selected documents from inpatient admissions were annotated by human reviewers to identify all positively and negatively asserted concepts. An NLP algorithm was tuned using 100 documents from the set. Novel lexicon semantics including evidence of catheter absence and inherently negated terms were used. Electronic medical record note titles with the highest hit rate for concepts were identified. The NLP algorithm was then tested on a set of 1495 documents to determine agreement between NLP and human reference standard, sensitivity and positive predictive value (PPV).

Results. The overall cohort included 5,589 unique patients with 77,938 hospital days from two VA hospitals over a one-year period. The lexicon contained 590 concepts for catheter presence (e.g., Foley catheter was placed) and 18 for evidence of absence (e.g., Patient has bathroom privileges). Iterative review of NLP outputs on the training set included false positive analyses and fine-tuning of the algorithm. Overall, nurse's notes were the most frequent inpatient note titles; these also yielded the highest number of concepts with respect to urinary catheters. The overall agreement between the NLP and reference standard was 71%. With 348 instances of 'evidence of catheter presence' the system found 246 for a sensitivity of 87%. With 84 false positive concepts associated with catheter presence, the PPV was 59%. For 'evidence of catheter absence', the agreement was 72% (450 instances), sensitivity was 77% and PPV was 68%.

Conclusion. We have shown that it is possible to identify the presence of an indwelling urinary catheter from the free text of electronic medical notes. Further refinement and scaling-up of NLP algorithms to large document sets is ongoing. This is the first key step in developing protocols to assist humans in large-scale review of patient charts for CAUTI.