A Natural Language Processing System to Identify Lumbar Spine Imaging Findings Related to Low Back Pain from Radiology Reports

Wei Ling Katherine Tan, Saeed Hassanzapour, PhD,1 Patrick Heagerty, PhD,2 Sean Rundell, DPT, PhD,3 Pradeep Suri, MD, MPH,4 Kathryn James, PA-C, MPH,5 David Carrell, PhD,6 Curtis Langlotz, MD, PhD,7 David Kallmes, MD,8 John Comstock, BA, James, K.T. et al.9 Lumbar Imaging with Reporting of Epidemiology (LIRE)—protocol for a pragmatic cluster randomized trial. Contemporary clinical trials. 2015, 45, pp.157-163.

Methods

Study Design and Population Studied

• Population studied: Lumbar Imaging with Reporting of Epidemiology (LIRE) study—pragmatic randomized clinical trial. 4 US health systems.
  - Adult patients (≥ 18 y/o) whose primary care provider ordered x-ray or MR of the lumbar spine.
  - Study design: Reference standard of N=871 radiology text reports dated between October 2013 and September 2016, stratified by study site and imaging modality (x-ray or MR); each report annotated for the presence / absence of 26 findings.

Analysis

• Inter-rater agreement (Figure 3):
  - Data: Subset of reference standard (N=800).
  - Metric: Cohen’s kappa for each annotator pair.
  - NLP algorithm evaluation (Figure 4):
    - Data: 20% of reference-standard for testing (N=174).
    - Metrics: Sensitivity, Specificity, Area Under the Receiver Operating Characteristic Curve (AUC) for each finding.

Conclusions

• The described 26 radiological findings related to LBP have substantial agreement from medical experts, and accurately identified by NLP as benchmarked by reference-standard annotations.
• Machine-learned models provided substantial gains in model sensitivity with similar specificity, compared to rule-based models.
• Our results suggest that NLP algorithms and predictions can be integrated into large Electronic Medical Records (EMR) databases to identify patients with certain radiological findings related to LBP for clinical and research purposes.