MEDINFO 2019: Health and Wellbeing e-Networks for All
L. Ohno-Machado and B. Séroussi (Eds.)

doi:10.3233/SHTI190527

Understanding Urgency in Radiology Reporting: Identifying Associations Between Clinical Findings in Radiology Reports and Their Prompt Communication to Referring Physicians

Xing Meng[a†], Michael V. Heinz[b†], Craig H. Ganoe[c], Ryan T. Sieberg[d], Yvonne Y. Cheung[d], Saeed Hassanpour[a,c,e]

[a] *Department of Computer Science, Dartmouth College, Hanover, New Hampshire, USA*
[b] *Department of Psychiatry, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, USA*
[c] *Department of Biomedical Data Science, Dartmouth College, Hanover, New Hampshire, USA*
[d] *Department of Radiology, Dartmouth-Hitchcock Medical Center, Lebanon, New Hampshire, USA*
[e] *Department of Epidemiology, Dartmouth College, Hanover, New Hampshire, USA*
[†] *These authors contributed equally to this work*

Abstract

In this study, we aim to develop an automatic pipeline to identify clinical findings in the unstructured text of radiology reports that necessitate communications between radiologists and referring physicians. Our approach identified 20 distinct clinical concepts and highlighted statistically significant concepts with strong associations to cases that require prompt communication.

Keywords:

Natural language processing, communication, radiology.

Introduction

It is well understood in clinical practice that lapses in communication, either due to delays or a lack of communication altogether, increase the likelihood of adverse patient outcomes. In radiology, high case volumes and the expectation of timely clinical interpretation present challenges in identifying and communicating cases that require urgent management on the part of the referring physician. In this study, we developed a pipeline for the identification of critical clinical concepts that are the most likely to appear in patient cases requiring prompt communication between a radiologist and a referring physician.

Related Work

There are many examples of previous uses of automatic information extraction from medical records, particularly with radiology report text analysis [1]. MetaMap has been developed and utilized to map radiology reports and other clinical note texts to concepts in the Unified Medical Language System (UMLS) Metathesaurus [2]. Apache clinical Text Analysis and Knowledge Extraction System (cTAKES) is a system to analyze clinical notes and annotate UMLS concepts in free text [3], which is utilized in this project and will be further discussed in the following sections. Of note, this study is built on our previous work to use machine learning to identify urgent cases for prompt communication in radiology reporting [4].

Methods

We developed a natural language processing (NLP) pipeline to extract clinical findings from a corpus of free-text radiology reports at our institution that required prompt communication between radiologists and referring physicians. We then compared these clinical findings to the extracted findings from regular radiology reports. We used a statistical test of significance to identify critical findings in these reports that are associated with prompt communication. The overview of our pipeline is shown in Figure 1.

Dataset

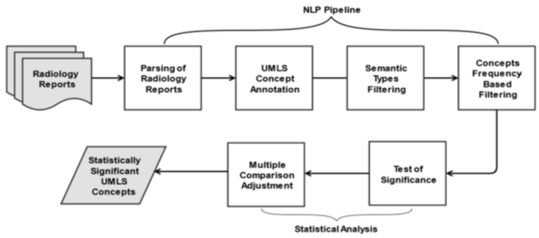

Figure 1– An Overview of Our Framework to Identify Statistically Significant Clinical Findings Requiring Prompt Communication with Referring Physicians Based on Radiology Reports

We took a retrospective approach in this research, starting with 1,460 radiology reports extracted from electronic medical records within the years of 2010 – 2018 from Dartmouth-Hitchcock Medical Center (DHMC). These reports came from a mixture of three imaging modalities: (1) computed tomography (CT); (2) magnetic resonance imaging (MRI); and (3) X-ray. Two radiologists manually annotated each of these reports, adding a binary label, either "this report requires prompt communication to referring physician" or "this report does not require prompt communication." In order to minimize the errors, reports that received the same label from both radiologists were included in the analysis and those that did not were excluded. Our final data set included 1,389 radiology reports. 1,057 of these were interpreted by both radiologists as requiring urgent communication to a referring physician, while 332 were interpreted as not requiring urgent communication.

NLP Pipeline

Parsing of Radiology Reports

We focused only on analyzing the text that appears in the "impression" sections of radiology reports. This is because the impression section contains information about diagnosis and follow-up recommendations.

UMLS Concept Annotation

We applied Apache clinical Text Analysis and Knowledge Extraction System (cTAKES), which identifies nouns and noun phrases in clinical text that correspond to UMLS concepts. Using cTAKES, we identified the concepts in radiology reports that are classified in 5 high level classes based on a subset of UMLS semantic types: (1) disorders/diseases, (2) signs/symptoms, (3) procedures, (4) anatomy, and (5) drugs. cTAKES also provides preferred names from UMLS for extracted concepts to facilitate a standard representation. In addition, this pipeline incorporates NegEx [5] to identify the concepts that are used in a negative context in the radiology report. We mapped all the medical terms and phrases that appeared in the impression section to UMLS concepts, excluding those that were negated.

Semantic Types Filtering

The identified UMLS concepts were aggregated according to the five high level semantic type classes. Using the full list of concepts identified by cTAKES, we found that the aggregated concepts helped to narrow down a list of potential concepts associated with the need for prompt communication with referring physicians. After consultation with the domain-expert clinicians involved in our study, we focused solely on UMLS concepts belonging to the diseases/disorders class in our analysis from the five aforementioned UMLS classes returned by cTAKES.

Concepts Identified by Frequency-Based Filtering

The identified UMLS concepts found in the reports were also filtered by frequency, including those that were only found in 2% or greater of reports. This frequency-based filtering reduced the noise in our analysis and directed our focus to reasonably common findings in radiology reports that encompass the majority of cases requiring prompt communication. The concept frequencies were further analyzed in the following statistical analysis steps to differentiate reports requiring prompt communication from those not requiring prompt communication.

Statistical Analysis

We used Fisher's exact test to find statistically significant associations between identified UMLS concepts and the radiology reports that require prompt communication. The concepts with a P-value < 0.05 are considered strongly associated with necessary prompt communication with referring physicians.

We addressed the multiple comparisons problem in evaluating the significance of extracted UMLS concepts. The false discovery rate provides a complementary measure to positive predictive value, which indicates the probability of a positive test result being accurate. In order to identify truly significant UMLS concepts while still maintaining a low false positive rate, we adjusted the P-values calculated by Fisher's exact test. To adjust the P-values, we utilized the Benjamini-Hochberg method, which is a powerful method to address the multiple comparisons problem through false discovery rate.

Results

The described NLP pipeline identified 580 UMLS concepts. Table 1 shows the extracted concepts with adjusted P-values

Table 1– Frequency of Selected UMLS Concepts in Radiology Reports that Require Prompt Communication
P = Occurrences with Prompt Communication
N = Occurrences without Prompt Communication

UMLS Concept	*Semantic Type*	*P*	*N*	*Odds Ratio*	*P-value*	*Adjusted P-value*
Pneumothorax	Disease or Syndrome	33	0	∞	2.25E-04	4.19E-02
Malignant Neoplasms	Neoplastic Process	73	3	8.05	2.66E-06	7.72E-04
Fracture	Injury or Poisoning	127	7	6.27	3.34E-09	1.94E-06
Nodule	Acquired Abnormality	107	13	2.73	2.89E-04	4.19E-02

less than 0.5 from radiology reports that are associated with reports requiring prompt communication. The most commonly identified UMLS concepts in radiology reports that required prompt communication with referring physicians are shown in Table 2.

Conclusions

We developed an automatic pipeline to identify clinical concepts that are statistically significant for radiology reports requiring prompt communication with referring physicians. However, our imbalanced dataset containing urgent and non-urgent cases for prompt communication is a potential limitation of this study. We plan to address this limitation by extending our dataset in our future work.

Table 2– Most Frequency UMLS Concepts in Impression Sections of Radiology Reports that Require Prompt Communication

P = Positive for Occurrences，N = Negative for Occurrences

UMLS Concept	*Semantic Type*	*P*	*N*
Fracture	Injury or Poisoning	127	7
Nodule	Acquired Abnormality	107	13
Malignant Neoplasms	Neoplastic Process	73	3
Communicable Diseases	Disease or Syndrome	62	6
Neoplasms	Neoplastic Process	53	8
Pleural effusion disorder	Disease or Syndrome	50	5
Traumatic injury	Injury or Poisoning	49	5
Cyst	Disease or Syndrome	43	6
Disorder of skeletal system	Disease or Syndrome	42	16
Protrusion	Anatomical Abnormality	39	15
Intervertebral Disc Degeneration	Disease or Syndrome	39	11
Pneumothorax	Disease or Syndrome	33	0
Abscess	Disease or Syndrome	29	2
Pneumonia	Disease or Syndrome	28	3
Squamous intraepithelial lesion	Neoplastic Process	26	13
Laceration	Injury or Poisoning	23	14
Arthropathy	Disease or Syndrome	22	4

Acknowledgements

The authors would like to thank William Keller, MD, Lamar Moss, and Sophie Montgomery for their help in this project.

References

[1] S. Hassanpour, and C.P. Langlotz, Unsupervised topic modeling in a large free text radiology report repository, J. Digit Imaging, 29, 1 (2016) 59-62.
[2] A.R. Aronson, Effective mapping of biomedical text to the UMLS Metathesaurus: The MetaMap program, Proc. AMIA Symp. (2001) 17.
[3] G.K. Savova, J.J. Masanz, P.V. Ogren, J. Zheng, S. Sohn, K.C. Kipper-Schuler, and C.G. Chute, Mayo clinical Text Analysis and Knowledge Extraction System (cTAKES): Architecture, component evaluation and applications, J. Am. Med. Informatics Assoc. 17 (2010) 507–513.
[4] X. Meng, C.H. Ganoe, R.T. Sieberg, Y.Y. Cheung, and S. Hassanpour, Assisting radiologists with reporting urgent findings to referring physicians: A machine learning approach to identify cases for prompt communication, J. Biomed. Inform. (2019) In press.
[5] W.W. Chapman, W. Bridewell, P. Hanbury, G.F. Cooper, and B.G. Buchanan, A simple algorithm for identifying negated findings and diseases in discharge summaries, J. Biomed. Inform. 34 (2001) 301–310.

Address for correspondence
Saeed Hassanpour, PhD
One Medical Center Drive, HB 7261, Lebanon, NH 03756, USA
Fax: (603) 650-1966 Email: Saeed.Hassanpour@dartmouth.edu