Rapid Identification of Pneumonias in BioSense Data Using Radiology Text Reports

Armen Asatryan, MD, MPH,[1] Haobo Ma, MD, MS,[1] Roseanne English, BS,[2] Jerome Tokars, MD, MPH[2]

[1]Science Applications International Corporation, [2]Centers for Disease Control and Prevention; Atlanta, GA

Objective
To study the performance of a simple keyword search of radiology reports for identifying pneumonia.

Background
BioSense currently receives demographic and chief complaint data from more than 360 hospitals and text radiology reports from 36 hospitals. Detection of pneumonia is an important as several Category A bioterrorism diseases as well as avian influenza can manifest as pneumonia. Radiology text reports are often received within 1-2 days and may provide a faster way to identify pneumonia than coded diagnoses.

Methods
We studied 67,714 reports of chest X-rays performed for 42,510 hospital visits (which corresponds to 34,883 patients) and the resulting final ICD-9 diagnostic codes from 13 medical facilities during February 2006 to January 2007. All patient classes (inpatient, outpatient, and emergency) were studied. We designed a keyword-based text parsing application using SAS software to detect the following terms: airspace, consolidation, density, infiltrate, opacity, and pneumonia or pneumonitis. The program accounted for negations and double negations. To validate the text-parsing algorithm, we selected 300 radiology records containing one or more pneumonia keywords either as a positive or negative (e.g., "no pneumonia was found") assertion and 100 records without these keywords. These 400 radiographs were evaluated by both a physician (used as the gold standard) and the keyword-based text parsing program. We used a logistic regression model to evaluate the independent association of the keywords with a final diagnosis of pneumonia (ICD-9 codes 480-486).

Results
Physician review of the 400-record sample classified 77 records as having the keyword "pneumonia," 63 as having other studied keywords, and 260 as having none of these. In comparison, the text parsing method showed 98.5% sensitivity and 98.6% specificity for identifying the keywords. The 34,883 patients in the full dataset had the following demographic characteristics: male to female ratio 1.1, ages ranged from newborn (less than 1 day old) to 105 years old with a mean of 55 and a mode of 81 years old. Of the 67,714 radiology reports of these patients, 8,631 (12.8%) occurred in patients with a ICD-9 final diagnosis of pneumonia; text parsing showed the following frequencies of keywords: opacity in 22.1% of reports, infiltrate in 15.0%, density in 8.2%, pneumonia in 6.0%, consolidation in 5.5%, and airspace disease in 1.1% (some reports had ≥2 keywords). A logistic regression model showed that 5 keywords were independently associated with a final diagnosis of pneumonia: infiltrate, odds ratio (OR) 3.3; pneumonia, OR 3.0; airspace, OR 2.7; consolidation, OR 2.7; and opacity OR 2.0 (density was not associated, OR=1.2). One or more of these 5 keywords was found in 71.7% of reports of patients with a pneumonia diagnosis vs. 30.8% of patients without a pneumonia diagnosis (sensitivity 72%, specificity 69%, kappa=.23).

Discussion
Our results show that a keyword-based program can successfully parse radiology reports to identify pneumonia-related keywords. Five keywords (infiltrate, pneumonia, airspace, consolidation, and opacity) were independently associated with a final diagnosis of pneumonia. An index based on radiology reports with these 5 keywords has reasonable sensitivity compared with final diagnoses but the kappa statistic shows poor agreement between the two. Explanations for this may include the inadequacy of coded diagnoses and that physicians may clinically diagnose pneumonia in patients with a negative radiology report. Therefore, either a radiography report or an ICD-9 coded diagnosis may indicate pneumonia, but neither is definitive. Further research is necessary to evaluate radiology text keywords using a formal case definition applied with chart review as the gold standard, and to evaluate the value of natural language processors in finding keywords and implementing more sophisticated concepts such as degrees of certainty. These studies may help to determine the benefit of radiology reports in biosurveillance systems such as BioSense.

Contact: Armen Asatryan, MD, MPH
E-mail: armen.asatryan@saic.com