Development and Validation of a Natural Language Processing Tool to Identify Patients Treated for Pneumonia across VA Emergency Departments

B. E. Jones[1,2]   B. R. South[3]   Y. Shao[3]   C.C. Lu[4]   J. Leng[4]   B. C. Sauer[1,4]   A. V. Gundlapalli[1,3,5,6]
M. H. Samore[1,3]   Q. Zeng[1,7]

[1] IDEAS Center of Innovation, VA Salt Lake City Health Care System, Salt Lake City, Utah, United States
[2] Division of Pulmonary and Critical Care Medicine, University of Utah School of Medicine, Salt Lake City, Utah, United States
[3] Department of Biomedical Informatics, University of Utah School of Medicine, Salt Lake City, Utah, United States
[4] Division of Epidemiology, University of Utah School of Medicine, Salt Lake City, Utah, United States
[5] Department of Internal Medicine Immunology, University of Utah School of Medicine, Salt Lake City, Utah, United States
[6] Department of Internal Medicine, VA Salt Lake City Health Care System, Salt Lake City, Utah, United States
[7] Clinical Research and Leadership, George Washington University School of Medicine and Health Sciences, Washington, District of Columbia, United States

Address for correspondence B. E. Jones, MD, MSc, IDEAS Center of Innovation, VA Salt Lake City Health Care System, 500 Foothill Drive Building 2, Salt Lake City, UT 84148-0002, United States (e-mail: barbara.jones@hsc.utah.edu).

Abstract

Background   Identifying pneumonia using diagnosis codes alone may be insufficient for research on clinical decision making. Natural language processing (NLP) may enable the inclusion of cases missed by diagnosis codes.

Objectives   This article (1) develops a NLP tool that identifies the clinical assertion of pneumonia from physician emergency department (ED) notes, and (2) compares classification methods using diagnosis codes versus NLP against a gold standard of manual chart review to identify patients initially treated for pneumonia.

Methods   Among a national population of ED visits occurring between 2006 and 2012 across the Veterans Affairs health system, we extracted 811 physician documents containing search terms for pneumonia for training, and 100 random documents for validation. Two reviewers annotated span- and document-level classifications of the clinical assertion of pneumonia. An NLP tool using a support vector machine was trained on the enriched documents. We extracted diagnosis codes assigned in the ED and upon hospital discharge and calculated performance characteristics for diagnosis codes, NLP, and NLP plus diagnosis codes against manual review in training and validation sets.

Results   Among the training documents, 51% contained clinical assertions of pneumonia; in the validation set, 9% were classified with pneumonia, of which 100% contained pneumonia search terms. After enriching with search terms, the NLP system alone demonstrated a recall/sensitivity of 0.72 (training) and 0.55 (validation), and a precision/positive predictive value (PPV) of 0.89 (training) and 0.71 (validation). ED-assigned

Keywords
► pneumonia
► natural language processing
► decision-making
► diagnosis
► surveillance

diagnostic codes demonstrated lower recall/sensitivity (0.48 and 0.44) but higher precision/PPV (0.95 in training, 1.0 in validation); the NLP system identified more "possible-treated" cases than diagnostic coding. An approach combining NLP and ED-assigned diagnostic coding classification achieved the best performance (sensitivity 0.89 and PPV 0.80).

Conclusion System-wide application of NLP to clinical text can increase capture of initial diagnostic hypotheses, an important inclusion when studying diagnosis and clinical decision-making under uncertainty.

Background and Significance

Pneumonia is the leading cause of death from infectious disease in the United States, with over 50,000 deaths[1] and 1 million hospitalizations annually.[2] Patients commonly present to the emergency department (ED) with respiratory symptoms, physiologic signs of infection, and an abnormal chest X-ray; however, the presenting signs and symptoms are often subjective and nonspecific, with substantial overlap with other diseases.[3] Pneumonia is a syndrome, with no true gold standard for diagnosis. Physicians may initially have a suspicion of pneumonia based upon initial clinical information, which can be either supported or weakened by diagnostic studies (laboratories, chest imaging). Identifying the underlying pathogen causing pneumonia is often delayed or elusive, with fewer than half of all cases of pneumonia resulting in a microbiologic confirmation of disease, even with aggressive testing.[4] The diagnosis of pneumonia thus carries with it substantial uncertainty, which makes it an excellent clinical scenario in which to study medical decision making.

Large epidemiologic studies of pneumonia typically use diagnostic coding to identify their study populations.[5] Principal diagnosis codes are entered manually by coders who review the entire episode of care, including physician notes, workup during the hospitalization, laboratory results, and discharge summary, to identify the chief reason for admission at the end of the entire encounter. ED patients in the Veterans Affairs (VA) system who are hospitalized are assigned an initial International Classification of Disease-9th Edition (ICD-9) code for the ED encounter, and a final hospital ICD-9 code upon discharge. However, the diagnosis is often uncertain or incomplete by the time the physician completes his/her management in the ED, as many tests ordered to inform the principal diagnosis, such as microbiologic cultures, do not result immediately. Many patients initially managed in the ED for suspected pneumonia are found to have a different final diagnosis, and thus cases with the most diagnostic uncertainty may be lost by the traditional approach.

Capturing the diagnostic hypothesis at the time of the initial presentation is possible through an analysis of the text in clinical documents. Unlike diagnostic coding, ED physician notes are typically completed during or shortly after the ED visit, and thus may better reflect working diagnostic hypotheses at the point of care, especially for hospitalized patients. Natural language processing (NLP) is a method of automated data extraction that converts unstructured free text to structured, encoded data by applying probabilistic or rule-based algorithms to combinations of terms.[6] It has previously been applied to radiology reports for radiographic concepts of pneumonia,[7,8] and resulting tools have been used for both surveillance[9,10] and provider cognitive support.[11] NLP has been applied to pneumonia discharge summaries to automate severity of illness estimation,[12] and text searching strategies have been proposed to identify concepts of pneumonia within the plan section of clinician documents.[13] However, no NLP system has been developed to identify patients with assertion of pneumonia from the physician's initial clinical document.

Objectives

Our goal was to develop a novel approach to more accurately identify patients initially treated for pneumonia in the ED using text data. The aims of our study were to:

1. Develop and validate an NLP tool that identifies clinical assertions of pneumonia from physician ED notes.
2. Compare ICD-9 codes, NLP, and combined classification approaches against manual chart review for the identification of clinical assertions of pneumonia within the text of physician ED notes.

Methods

Study Population

All data were extracted and analyzed using the Veterans Informatics and Computing Infrastructure (VINCI).[14] We identified all visits to the U.S. Veterans Affairs (VA) EDs throughout the United States during the years 2006 to 2012 that had chest imaging (computed tomography [CT] scan or chest X-ray) obtained within 24 hours of the visit time and at least one clinical document generated within 24 hours from the visit time with a standard note title consistent with an ED note or an addendum. From this corpus of clinical documents, we selected two samples (►**Fig. 1**). First, we selected an enriched sample of 1,000 documents that contained the search term "pneumonia," misspellings, and synonyms (►**Supplementary Material A**, available in the online version). The rationale for selecting an enriched data set versus a random sample was to enable our algorithms to maximize discrimination, as the proportion of all patients with suspected pneumonia among all ED visits with chest imaging would be expected to be less than 10%. Of the 1,000 documents selected, we excluded all documents that
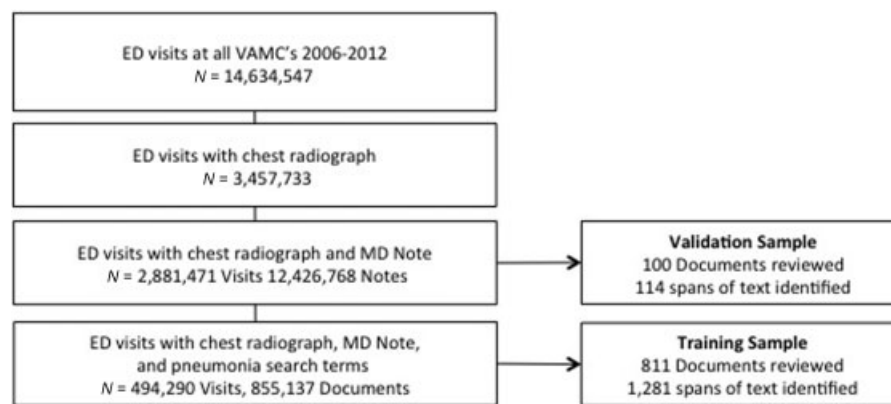
Fig. 1 Study population.

were not signed by a physician or that lacked the structure of a clinical note. Second, we selected a random sample of 100 clinically relevant physician notes for validation not enriched by search terms, to estimate the performance of different case identification strategies among the entire population.

Definition of Clinical Assertion of Pneumonia

We classified each document for evidence of clinical assertion of pneumonia, defined as whether or not the diagnosis of pneumonia was suspected by the authoring clinician at the time of the encounter. Using an annotation tool developed to simplify annotation tasks,[15] the documents were first preannotated with the preliminary list of pneumonia search terms. Two human reviewers (B.J. and B.S.) then read 300 documents in three batches of 100, with three adjudication sessions after each batch. We selected and annotated 1,000 clinical documents from the enriched sample. In each document, reviewers highlighted one or more text segments, or spans, and classified them into one of the four categories below. Reviewers then classified the entire document. The reviewer classified each span and document into one of four categories:

A. Certain pneumonia—provider proceeded with treatment for pneumonia with no discussion of other alternative diagnoses.
B. Possibly pneumonia—provider proceeded with treatment for pneumonia but mentioned the diagnosis was possible or probable, or listed additional possible diagnoses.
C. Possibly pneumonia—pneumonia was mentioned, but provider asserted the suspicion was low, and they did not proceed with treatment for pneumonia.
D. Certainly not pneumonia—provider did not mention pneumonia at all, or indicated that pneumonia was ruled out by diagnostic studies.

During the adjudication process for each document batch, disagreements between reviewers were discussed, and we developed an annotation guideline that documented our final definition of each classification (►**Supplementary Material B**, available in the online version). By the third batch of 100 documents, a document-level classification interrater agreement kappa of 88.3% was achieved, and the remainder of the documents were annotated independently

by B.J. Both reviewers annotated the validation independently, and disagreements were adjudicated.

NLP Tool Development

To develop an NLP classification tool that could be reliably compared with ICD-9 coding strategies, we developed a classification that could distinguish A and B (assertions of certain or possible pneumonia that was treated) from C and D (certainly not pneumonia or possible, but not treated). The NLP classification was a combination of two levels of classifications: one at the span level and the other at the document level. A study on the annotation data showed that for more than 98% documents the document-level classification agreed with the classification of the last span in that document; this is consistent with the structure of a physician note, in which the assessment is organized at the end of the document. Thus, we took the category of the last span-level classification (predicted by the span-level classifier) as the category of the whole document.

The span-level classifier was built using the support vector machine (SVM).[16,17] We started by tokenizing the spans and extracting both one-gram and two-grams as features. This yielded a total of 15,928 features. Then, we performed a feature selection to select the features with high discriminative power. The feature selection process was done as follows. First of all, we discarded features that occurred in only one span. Next, for each of the remaining features, we calculated the prevalence of category A and B in the spans with the feature. Theoretically, if the feature has little association with the category (i.e., little discriminative power), then this prevalence should be close to 50%, the prevalence of category A and B in all the spans. Therefore, we selected the features such that the prevalence of A and B was either below 10% or above 70%. This yielded a total of 1,775 features.

The spans were then converted into binary vectors using the 1,775 selected features as inputs for SVM, an approach previously used with some success to identify assertion classification in the i2b2 corpora.[18] We chose the SVM with radial basis function kernel, which made it capable of classifying sets with nonlinear boundaries.[19] There were two parameters associated with this kernel: C and gamma. By experimenting with C values from 1 to 10 with step 1, and gamma values from 0.01 to 0.1 with step 0.01, we found that when C = 6 and gamma = 0.05,

the SVM performed with the best accuracy. Therefore, these values were chosen for the final model. Because we found nearly 100% consistency between the last span and the document-level classification, and this was clinically consistent with the expected location of the most relevant information contained in a clinical note, we trained the system to classify each document based upon the value of last span from each document. The parameter values were validated over the 10-fold cross-validations process; the data segments used were all the 10 folds of data. The final model was the one trained on the whole training data over a single pass.

Diagnostic Coding of Pneumonia

For each ED visit that generated the clinical documents reviewed, we identified diagnostic coding for pneumonia coding using the primary/principal ICD-9 code form pneumonia (481–486), or a secondary ICD-9 code for pneumonia and a primary/principal ICD-9 code for sepsis (038.0, 038.11, 038.12, 038.x, 995.91, 995.92, 785.52) or respiratory failure (518.81–84, 799.1), consistent with previous studies.[5] In the VA system, hospitalized patients receive an ICD-9 code both at the time of the ED visit as well as at the end of the hospital encounter. Therefore, of those ED visits that resulted in a hospital admission, we identified both the ED-assigned ICD-9 codes and discharge ICD-9 code.

Analysis

To test the span classification in the training set, we conducted a 10-fold cross-validation, partitioning the data randomly into training and test sets to internally validate the model's predictions. The documents were divided into 10-folds of equal size. Then, we held out one fold at one time, trained the SVM on the spans of the documents from the remaining 9 folds, applied the trained model to the last span of each document from the held-out fold, and took that as the classification of the documents. This process was conducted 10 times, and each time a different fold was held out so that each fold was used for testing exactly once. The classification on the held-out folds from the 10 times training and testing was then combined to calculate the recall/sensitivity, specificity, precision/positive predictive value (PPV) and negative predictive value (NPV). As the system was originally trained to identify clinical suspicion at the span level, we calculated the accuracy of SVM-assigned suspicion for pneumonia of each span against manual review using a 10-fold internal validation.

For each document, we compared the accuracy of NLP classification, ED-assigned ICD-9 coding, and hospital-assigned ICD-9 coding against the gold standard of manual review for clinical assertion of certain or possible/treated pneumonia. Because ED-assigned ICD-9 codes are collected very close to the end of the ED encounter in the VA system, and we noted the PPV of ED-assigned ICD-9 coding to be high, we also tested the accuracy of a combined NLP plus ED-assigned ICD-9 coding approach, which identified a case treated for pneumonia if it was classified as positive by either NLP or ED-assigned ICD-9 code. For the validation set, we calculated the proportion of positive cases identified with pneumonia search terms and tested the performance of the NLP after enrichment with search terms, since the NLP system was

designed to be processed on those documents. We calculated the recall/sensitivity, specificity, precision/PPV, and NPV for each approach in the training and validation sets using two-by-two tables. All statistical analyses were performed using Stata 14 MP (StataCorp. 2015. *Stata Statistical Software: Release 14.* StataCorp LP, College Station, Texas, United States).

Results

We identified 14,634,547 visits to 94 VA EDs throughout the United States, of which 3,457,733 had chest imaging obtained within 24 hours of the visit time. A total of 2,881,471 of these visits had at least one clinical document generated within 24 hours from the visit time that had a standard note title consistent with an ED note or an addendum; this generated 12,426,768 notes (4.3 documents per visit).

Among the 1,000 documents retrieved based upon the pneumonia search terms selected, 189 were excluded due to lack of relevant information or evidence of physician authorship, resulting in 811 represented complete ED notes authored by a physician (►**Fig. 1**). Two hundred thirty-two (29%) cases were identified as certain pneumonia, 180 (22%) possible-treated, 78 (10%) possible-not treated, and 321 (40%) as certainly not pneumonia. The documents contained a total of 1,281 spans of text from which the NLP system was trained.

Among 171 documents retrieved for validation, 71 were excluded due to lack of relevant information, resulting in 100 notes. Nine cases were identified as being treated for pneumonia. Among these, 100% contained pneumonia search terms. The documents in the validation contained 114 spans of text.

Span-level classification demonstrated high accuracy against manual review in the training set (►**Table 1**). Features indicating high likelihood of a positive class included antibiotic terms and modifiers of the diagnosis (such as "community," "acquired," or "cap"); features identified in negative classification included mention of medical history, negation, and alternative diagnoses (i.e., "uri," "edema"). ►**Table 2** illustrates the performance characteristics of the classification approaches against the human review in both training and validation (complete contingency tables are available in ►**Supplementary Material C**, available in the online version). Both methods of ICD-9 coding demonstrated higher positive predicted value but lower recall/sensitivity against human review for clinical suspicion of certain or possible-treated pneumonia. The NLP system demonstrated a recall/sensitivity of 0.72 and precision/PPV of 0.94 in the training set (►**Fig. 2**), although performance decreased in the validation (0.56, 0.71). The combined approach of the NLP classification plus ED-assigned ICD-9 codes—that is, classifying any NLP- or ED-ICD-

Table 1 NLP span-level performance using 10-fold cross-validation

Specificity	0.94
PPV	0.89
NPV	0.82

Abbreviations: NLP, natural language processing; NPV, negative predictive value; PPV, positive predictive value.
Note: $N = 1,281$ spans of text within 811 documents from 811 ED visits.

Table 2 Accuracy of ED-assigned ICD-9, discharge ICD-9, NLP, and combined NLP + ED-assigned ICD-9 codes to identify initial clinical assertion of pneumonia

Approach		Recall/Sensitivity	Specificity	Precision/Positive predictive value	Negative predictive value
ED-assigned ICD-9	Training	0.48	0.97	0.95	0.64
	Validation	0.44	1.0	1.0	0.95
Hospital discharge ICD-9	Training	0.45	0.92	0.89	0.54
	Validation	0.67	0.99	0.86	0.97
NLP	Training	0.72	0.95	0.94	0.77
	Validation	0.56	0.98	0.71	0.96
NLP plus ED-assigned ICD-9	Training	0.88	0.93	0.92	0.88
	Validation	0.89	0.98	0.80	0.99

Abbreviations: ED, emergency department; ICD, International Classification of Disease; NLP, natural language processing.
Note: Performance characteristics are reported against a reference standard of manual chart review. Positive assertion of pneumonia is defined as "certain" or "possible-treated" pneumonia.

9 code-positive document as positive for clinical assertion of pneumonia—demonstrated superior accuracy (sensitivity 0.89, PPV 0.80 in the validation).

►**Fig. 3** shows the composition of assertions of pneumonia in cases identified by each approach among the training documents. The ED-assigned ICD-9 approach identified the majority of certain pneumonia cases; however, it failed to identify a large number of possible-treated cases. NLP classification and combined NLP + ED-assigned ICD-9 approaches identified a substantially greater number of possible pneumonia cases, without increasing the number of false-positive cases.

Discussion

Our study describes the development and validation of an NLP tool to identify cases with a clinical assertion of pneumonia
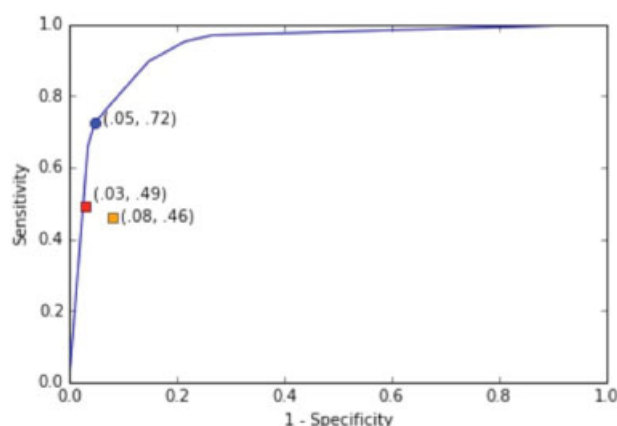

Fig. 2 Signal detection plot of International Classification of Disease (ICD)-based classification, natural language processing (NLP) classification, and receiver operating characteristic (ROC) curve of the NLP classification. The red square indicates the emergency department (ED)-assigned ICD classification, and the orange square indicate the hospital assigned ICD classification. Overall receiver operator characteristics area under the curve of the NLP was 0.935, with 95% confidence interval = [0.917, 0.953]. The blue dot is the NLP classification (specificity = 95%) calibrated to match the ICD-based classification on specificity.

using clinical documents authored by clinicians treating ED patients. When compared with the traditional approach to identifying cohorts through diagnostic coding, our NLP system identified a larger number of patients initially treated with pneumonia, in particular cases in which the diagnosis was less certain. A combined approach with ED-assigned diagnostic coding and NLP classification demonstrated high sensitivity and PPV. Using NLP classification can thus enhance research on decision making for pneumonia by including a more clinically relevant initial sample of patients.

A challenge to pneumonia research is that the diagnosis of pneumonia carries a substantial degree of uncertainty, which is difficult to capture using traditional approaches to cohort selection. The diagnosis of pneumonia is thus a dynamic process throughout a patient's encounter with medical care, and often providers must proceed with treatment of pneumonia plus several other possible causes of the patient's presentation. Our study found that ICD-9 coding identified cases with clinical suspicion for pneumonia with high PPV but low sensitivity compared with human review, failing to capture nearly half of all of the cases that were treated for suspected pneumonia in our enriched sample. This finding is consistent with previous studies reporting low sensitivities in diagnostic coding.[20–22] Hospital discharge ICD-9 coding, which is the most commonly used approach to identify research subjects, occurs at the end of the patient's encounter, when there is less diagnostic uncertainty than the initial encounter. In addition to concerns for inconsistencies in diagnostic coding practices, the use of ICD-9 coding, while possibly adequate to study the disease process of pneumonia, fails to include a large number of clinical scenarios, especially those with greater uncertainty in the diagnosis. We calibrated our NLP classification tool to maintain a specificity of 0.95 in the training set; with this, the NLP captured a greater number of cases, many of which were treated as possible pneumonia based upon the initial clinical document reviewed.

Capturing patients with diagnostic uncertainty is crucial to understanding the process of decision making in medicine. A recent report by the Institute of Medicine[23] highlighted
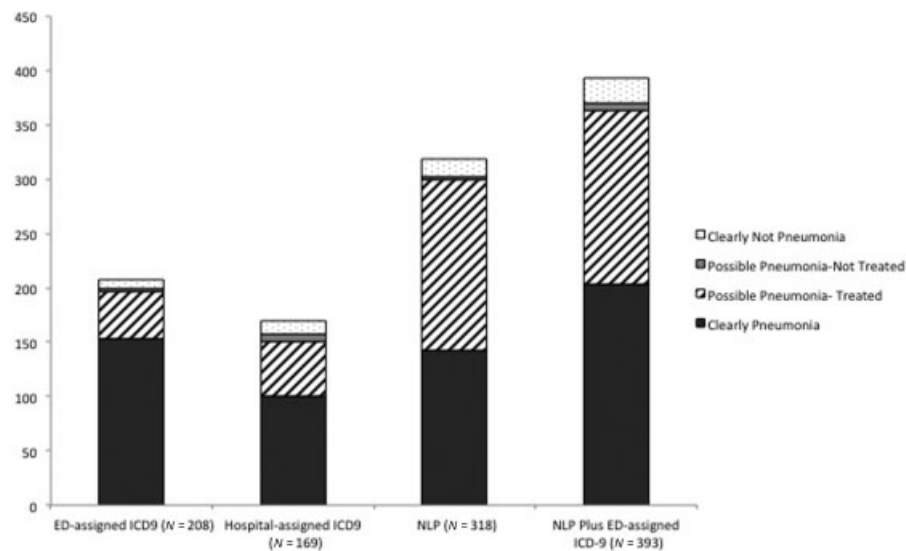
Fig. 3 Composition of assertions of pneumonia among cases identified with each cohort selection approach in the training set.

diagnosis errors in medicine as a major challenge to delivering quality health care, calling for more research and information technology efforts toward understanding and improving the decision-making process surrounding diagnosis. The rich data infrastructure available through the evolving electronic health record provides new opportunities to study and support decision making throughout the process of a patient's encounter. To better understand the process of diagnosis and management of pneumonia, and to develop tools that support more accurate diagnosis while acknowledging that treatment must often occur under uncertainty, we must start with a group of patients that reflect the reality of this uncertainty. NLP provides a new way to identify diagnostic hypotheses at the patient's initial presentation, to examine the evolution of diagnoses throughout a patient's encounter.

We recognize several limitations to our study. We trained the NLP on a small, enriched sample of documents using search terms, and the performance characteristics of the NLP classification approach did degrade in the validation set. In a qualitative error analysis of false positives from the validation, the most common features were receipt of antibiotics for another diagnosis or a nonspecific term sometimes associated with pneumonia, such as "sepsis"; in the false negatives, possible diagnoses that mentioned "r/o" ("rule out") appeared misinterpreted by the NLP as negative cases. We will apply these findings to improve the classifier with future iterations. Despite these errors, the combined approach that uses both ICD-9 codes and NLP classification appeared promising. Additional training of the NLP system on a larger corpus of documents should be expected to improve its performance and is the subject of future work. Similarly, the validation set was also small, and thus the characteristics reported may not exactly reflect those of a random sample of ED visits with chest imaging. However, we found the prevalence of search terms among the study population to be 17%, near the 22% found in the validation set, so the prevalence of patients treated with pneumonia in the larger population is likely near 10%. The human reviewers selected the spans against which the SVM

was trained and tested, and thus a span selection approach must still be developed in order for the NLP system to be automated. However, 100% of the spans containing evidence for clinical assertion for pneumonia contained pneumonia search terms; thus, we anticipate that a span selection tool that extracts spans surrounding the pneumonia search terms will be feasible. Accurately selecting documents that contain information to reduce processing time is also a challenge, as we found many documents that lacked relevant information. We have since refined our document selection strategy to identify the most clinically relevant documents for each visit, to optimize the NLP processing of our national data set.

Our study suggests that NLP is a feasible and promising technology that can extract concepts of clinical suspicion of pneumonia throughout a patient's encounter with medical care. An approach that combined NLP with ED-assigned diagnosis codes improved performance beyond either method alone. Further work is needed to identify additional opportunities for NLP applications to better understand and support medical decision making.

Conclusion

In a study designed to develop an NLP tool, we found that a large proportion of patients initially treated for pneumonia were not captured by the traditional cohort selection techniques that use ICD-9 coding. We developed and validated an NLP tool to identify clinical suspicion for pneumonia at the initial presentation that demonstrated high accuracy when combined with diagnostic coding. The NLP for pneumonia provides a new way to identify a large population of cases of pneumonia for both research of decision making and development of informatics tools.

Clinical Relevance Statement

The initial diagnosis of pneumonia can be uncertain. We found that traditional approaches to identifying pneumonia cases in population studies failed to capture a large number

of cases initially treated for pneumonia, especially those cases with possible pneumonia. NLP identified more cases initially treated with pneumonia, which is critical when studying decision making for pneumonia and designing decision support. A combined approach of NLP plus ED-assigned ICD-9 codes demonstrated the highest accuracy.

Multiple Choice Question

Which of the following important advantages to using natural language processing (NLP) over diagnostic coding to define patients treated for pneumonia were identified by the study?

a. NLP was more sensitive, capturing a larger number of patients treated for pneumonia.
b. NLP was more specific, excluding a greater number of patients who were not treated for pneumonia.
c. NLP captured a larger number of patients who were treated for "possible pneumonia," thus including patients with the most diagnostic uncertainty.
d. a and c
e. All of the above.

Correct Answer: The correct answer is d, option a and c.

References

1 Centers for Disease Control and Prevention, National Center for Health Statistics. Underlying Cause of Death 1999–2015 on CDC WONDER Online Database, released December, 2016. Data are from the Multiple Cause of Death Files, 1999–2015, as compiled from data provided by the 57 vital statistics jurisdictions through the Vital Statistics Cooperative Program. Available at: http://wonder.cdc.gov/ucd-icd10.html. Accessed October 23, 2017
2 Ramirez JA, Wiemken TL, Peyrani P, et al; University of Louisville Pneumonia Study Group. Adults hospitalized with pneumonia in the United States: incidence, epidemiology, and mortality. Clin Infect Dis 2017;65(11):1806–1812
3 Mandell LA, Wunderink RG, Anzueto A, et al; Infectious Diseases Society of America; American Thoracic Society. Infectious Diseases Society of America/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. Clin Infect Dis 2007;44(Suppl 2):S27–S72
4 Jain S, Self WH, Wunderink RG, et al; CDC EPIC Study Team. Community-acquired pneumonia requiring hospitalization among U.S. adults. N Engl J Med 2015;373(05):415–427
5 Ruhnke GW, Coca-Perraillon M, Kitch BT, Cutler DM. Trends in mortality and medical spending in patients hospitalized for community-acquired pneumonia: 1993-2005. Med Care 2010; 48(12):1111–1116
6 Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: an introduction. J Am Med Inform Assoc 2011;18(05):544–551
7 Dublin S, Baldwin E, Walker RL, et al. Natural language processing to identify pneumonia from radiology reports. Pharmacoepidemiol Drug Saf 2013;22(08):834–841
8 Fiszman M, Chapman WW, Aronsky D, Evans RS, Haug PJ. Automatic detection of acute bacterial pneumonia from chest X-ray reports. J Am Med Inform Assoc 2000;7(06):593–604
9 Haas JP, Mendonça EA, Ross B, Friedman C, Larson E. Use of computerized surveillance to detect nosocomial pneumonia in neonatal intensive care unit patients. Am J Infect Control 2005;33(08):439–443
10 Liu V, Clark MP, Mendoza M, et al. Automated identification of pneumonia in chest radiograph reports in critically ill patients. BMC Med Inform Decis Mak 2013;13:90
11 Dean NC, Jones BE, Ferraro JP, Vines CG, Haug PJ. Performance and utilization of an emergency department electronic screening tool for pneumonia. JAMA Intern Med 2013;173(08):699–701
12 Friedman C, Knirsch C, Shagina L, Hripcsak G. Automating a severity score guideline for community-acquired pneumonia employing medical language processing of discharge summaries. Proc AMIA Symp 1999:256–260
13 DeLisle S, Kim B, Deepak J, et al. Using the electronic medical record to identify community-acquired pneumonia: toward a replicable automated strategy. PLoS One 2013;8(08):e70944
14 VA Informatics and Computing Infrastructure. Available at:https://www.hsrd.research.va.gov/for_researchers/vinci/. Accessed January 28, 2018
15 Leng JSS, Gundlapalli A, South B. The Extensible Human Oracle Suite of Tools (eHOST) for Annotation of Clinical Narratives. American Medical Informatics Association Spring Congress; 2010. May 25, 2010
16 Cortes CaVV. Support-vector networks. Mach Learn 1995;20(03):273
17 Hearst MADS, Osman E. Support vector machines. IEEE Intel Sys Appl 1998;13:18–28
18 Uzuner Ö, South BR, Shen S, DuVall SL. 2010 i2b2/VA challenge on concepts, assertions, and relations in clinical text. J Am Med Inform Assoc 2011;18(05):552–556
19 Lin C-J, Chang C-C. LIBSVM: a library for support vector machines. ACM Trans Intell Syst Technol 2011;2(27):1–27
20 Aronsky D, Haug PJ, Lagor C, Dean NC. Accuracy of administrative data for identifying patients with pneumonia. Am J Med Qual 2005;20(06):319–328
21 van de Garde EM, Oosterheert JJ, Bonten M, Kaplan RC, Leufkens HG. International classification of diseases codes showed modest sensitivity for detecting community-acquired pneumonia. J Clin Epidemiol 2007;60(08):834–838
22 Whittle J, Fine MJ, Joyce DZ, et al. Community-acquired pneumonia: can it be defined with claims data? Am J Med Qual 1997;12(04):187–193
23 Committee on Diagnostic Error in Health Care, Board on Health Care Services, Institute of Medicine, The National Academies of Sciences, Engineering, and Medicine. In: Balogh EP, Miller BT, Ball JR, eds. Improving Diagnosis in Health Care. Washington, DC: National Academies Press (US); 2015