Natural Language Processing for the Identification of Surgical Site Infections in Orthopaedics

Caroline P. Thirukumaran, MBBS, MHA, PhD, Anis Zaman, MA, Paul T. Rubery, MD, Casey Calabria, CIC, BSN-RN, MBA, Yue Li, PhD, Benjamin F. Ricciardi, MD, Wajeeh R. Bakhsh, MD, and Henry Kautz, PhD

Investigation performed at the University of Rochester, Rochester, New York

Background: The identification of surgical site infections for infection surveillance in hospitals depends on the manual abstraction of medical records and, for research purposes, depends mainly on the use of administrative or claims data. The objective of this study was to determine whether automating the abstraction process with natural language processing (NLP)-based models that analyze the free-text notes of the medical record can identify surgical site infections with predictive abilities that match the manual abstraction process and that surpass surgical site infection identification from administrative data.

Methods: We used surgical site infection surveillance data compiled by the infection prevention team to identify surgical site infections among patients undergoing orthopaedic surgical procedures at a tertiary care academic medical center from 2011 to 2017. We compiled a list of keywords suggestive of surgical site infections, and we used NLP to identify occurrences of these keywords and their grammatical variants in the free-text notes of the medical record. The key outcome was a binary indicator of whether a surgical site infection occurred. We estimated 7 incremental multivariable logistic regression models using a combination of administrative and NLP-derived variables. We split the analytic cohort into training (80%) and testing data sets (20%), and we used a tenfold cross-validation approach. The main analytic cohort included 172 surgical site infection cases and 200 controls that were repeatedly and randomly selected from a pool of 1,407 controls.

Results: For Model 1 (variables from administrative data only), the sensitivity was 68% and the positive predictive value was 70%; for Model 4 (with NLP 5-grams [distinct sequences of 5 contiguous words] from the medical record), the sensitivity was 97% and the positive predictive value was 97%; and for Model 7 (a combination of Models 1 and 4), the sensitivity was 97% and the positive predictive value was 97%. Thus, NLP-based models identified 97% of surgical site infections identified by manual abstraction with high precision and 43% more surgical site infections compared with models that used administrative data only.

Conclusions: Models that used NLP keywords achieved predictive abilities that were comparable with the manual abstraction process and were superior to models that used administrative data only. NLP has the potential to automate and aid accurate surgical site infection identification and, thus, play an important role in their prevention.

Clinical Relevance: This study examines NLP's potential to automate the identification of surgical site infections. This automation can potentially aid the prevention and early identification of these surgical complications, thereby reducing their adverse clinical and economic impact.

Surgical site infections are frequently occurring and are the most expensive of all hospital-acquired infections[1-3]. In orthopaedics, the mean surgical site infection rates range from 0.5% to 3% following hip and knee replacement surgical procedures[4-6] and from 0.2% to 7.2% following spine surgical procedures[7-9], and the cost of treating a surgical site infection can be as high as $65,000[9]. Surgical site infections are also among the most common causes of readmissions following joint replacement surgical procedures[10] and are included as quality metrics in several payment reforms such as the U.S.

Comprehensive Care for Joint Replacement model[11] and the U.S. Hospital Readmissions Reduction Program[12].

Successful surgical site infection prevention requires elaborate surveillance mechanisms that identify signs of impending surgical site infections, expeditiously treat early surgical site infections, and generate knowledge from these events to inform future prevention efforts[13]. In several institutions across the United States, the gold standard for surgical site infection identification is the manual abstraction of the patient's medical record by trained infection prevention personnel[1]. Although this practice may be suitable for single institutions[14], large-scale research to inform surgical site infection prevention strategies generally depends on the use of administrative data and diagnostic codes, with attendant limitations of both sensitivity and specificity[15]. Thus, there are important opportunities in automating the identification of surgical site infection from the medical record and potentially avoiding the limitations of both manual abstraction (e.g., timeliness, reproducibility, and scalability) and administrative data.

One approach is the use of natural language processing (NLP) to analyze the free text of a patient's medical record. NLP is a subfield of artificial intelligence and includes a set of tools that can be used to analyze the unstructured free text in a patient's medical record to encode concepts and to derive meaning[16,17]. This is conducted using rule-based approaches, which require a lexicon of words indicative of surgical site infections as well as machine learning-based approaches that have greater dependencies on the computer for extracting meaningful patterns[18]. In clinical practice, surgical site infections are identified using guidelines specified by the voluntary U.S. National Healthcare Safety Network (NHSN) program[19]. These guidelines include the level of involvement of the skin and subcutaneous tissue, the presence of purulent drainage, and the reporting of the signs and symptoms. Frequently, these observations are documented as free text in the patient's medical record. Thus, in the absence of discrete and codified systems for capturing these conversations, the capabilities of NLP make it a suitable approach for analyzing this information.

Few NLP studies have demonstrated the advantages of NLP-based models for surgical site infection identification in a mix of medical and surgical patients[17,20,21]. However, whether similar methods can be reliably used in orthopaedics is an open question. The objective of our study was to construct rule-based NLP algorithms for retrospectively identifying patients who developed surgical site infections following orthopaedic surgical procedures. We hypothesized that the predictive accuracy of the NLP-based models would be comparable with results from the gold-standard manual surgical site infection abstraction process and would be considerably superior to models based solely on administrative data.

Materials and Methods
Data Sources and Cohort
We used the gold-standard surgical site infection surveillance data for orthopaedic patients from January 2011 through June 2017 at Strong Memorial Hospital, the major teaching hospital of the University of Rochester, New York. This data set was generated by the infection prevention team as a part of their surveillance efforts, which are based on the NHSN guidelines[22]. It included patient, surgery, and surgical site infection details from orthopaedic patients who developed ≥1 surgical site infection (hereafter called cases) in the 90 days following the surgical procedures. Notably, at Strong Memorial Hospital, the infection prevention team tracks the occurrence of surgical site infections following every surgical procedure, and, hence, the data are not limited to the surgical procedures monitored by the NHSN. We included 172 cases and linked these data to the patients' administrative and medical record data. The medical record data included free-text notes such as progress notes, discharge summaries, history and physical examination notes, and telephone encounter notes. Because surgical site infection-relevant language is most likely to be found in notes written by the clinical team, we limited the notes to those written by physicians, residents, advanced practice providers, and registered nurses and to notes written within 10 days before and 14 days after the detection of a surgical site infection.

We obtained administrative and medical record data for 1,407 patients who underwent orthopaedic surgical procedures at Strong Memorial Hospital from 2015 to 2016, did not develop surgical site infections in the 90-day postoperative period (hereafter called controls), and had combinations of diagnosis and procedures similar to cases. The distribution of demographic characteristics and clinical comorbidities of these controls were similar to those of controls from previous years (Appendix Table 1). To ensure that there was no class imbalance[23], we repeatedly and randomly sampled 200 patients from the controls at various steps in the analysis, resulting in analytic cohorts of 372 patients at each instance.

The University of Rochester's Research Subject Review Board approved the study protocol. Data management and analysis were performed using Stata/MP 15[24] (StataCorp) and Python 3.0[25] on Unix.

Outcome
The key outcome was a binary indicator of any surgical site infection (superficial, deep, or organ or space) within 90 days of the primary surgical procedure as recorded in the surgical site infection surveillance data set.

Linguistic Variables and Features
We cleaned the free-text notes by detecting and managing inconsistencies such as the removal of stop words (i.e., words such as "a," "an," "the," and "to" that occur frequently but do not contribute in formulating the main context or the true meaning of a sentence) and filtering out non-English words. We reviewed the clinical literature and used the clinical expertise of our team to compile a list of >40 clinical keywords that are most frequently used by clinicians to describe surgical site infections (Table I). Next, we created an exhaustive list of stemmed or lemmatized versions of the keywords to capture

TABLE I List of Keywords Used for Describing Surgical Site Infections and Classifying Notes*
Abscess
Acinetobacter
Antibiotic
Apnea
Bradycardia
Candida
Clostridium difficile
Cough
Culture
Dehiscence
Delayed healing
Drainage
Dysuria
E. coli (Escherichia coli)
Emesis
Enterobacter
Enterococcus
Erythema
Escherichia
Fever
Heat
Hypothermia
Incision drained
Incision opened
Infection
Inflammation
Klebsiella
Lethargy
Material
MRSA (methicillin-resistant Staphylococcus aureus)
Nausea
Organism
Pain
Pathogen
Pseudomonas
Purulent
Redness
Serous
Sinus tract
SSI (surgical site infection)
Staphylococcus
Swelling
Tenderness
Vomiting
Wound

*The exact keywords and the stemmed or lemmatized versions of these keywords were used to generate linguistic variables and features. This list was compiled using the NHSN's guidelines for identification of surgical site infections and from the input of the clinical team members.

the variable grammatical forms of these keywords. For example, for the keyword "infection," the stemmed version was "infecti," and the lemmatized version was "infection." We also created 5-grams (distinct sequences of 5 contiguous words) of each note, and we examined the presence of the keywords and their variants in these 5-grams. We also captured instances when the clinical team documented the absence of a keyword. For this, we used web-based resources to create an extensive list of negations that were treated as prefixes for the keywords to capture linguistic cues that were suggestive of a patient not developing a surgical site infection[26-28]. Examples of these negations include "no purulent drainage" and "rule out infection."

During the notes analysis, the occurrence of a keyword, its stemmed or lemmatized version, or its 5-grams were expressed as counts for the variable or feature. This was further refined such that none of the negation prefixes were present in the 4 words preceding a keyword and that the negation prefix and the keyword did not appear in the note. Notably, because a mention of the keyword "pain" can occur among both cases and controls, we used the 5-gram strategy to identify persistence of pain among the 2 groups. Controls were also determined by the absence of other linguistic patterns suggestive of surgical site infections.

Covariates

Some models controlled for important covariates from administrative data such as demographic characteristics (age, sex, and race), clinical comorbidities (identified using the Elixhauser comorbidity algorithm)[29], the year of the surgical procedure, and the Clinical Classification Software[30] diagnosis categories to control for the clinical heterogeneity among orthopaedic subspecialties.

Analysis

We examined the distribution of demographic and clinical characteristics in the cohort. We examined the distribution of surgical site infections in surgical subspecialties among the cases from the main analytic cohort (n = 172) and the cases from an external validation cohort (n = 36) that were obtained from another tertiary care hospital affiliated with Strong Memorial Hospital. We constructed a series of sequential and incremental logistic regression models to determine whether models with NLP-derived features and variables could reasonably predict surgical site infections. The models included predictors from administrative data (described in the covariates section) (Model 1), NLP keywords (Model 2), stemmed or lemmatized versions of keywords (Model 3), and 5-grams (Model 4). We further combined Model 1 with Model 2 to create Model 5, Model 1 with Model 3 to create Model 6, and Model 1 with Model 4 to create Model 7.

For training and validation, we split the analytic cohort (composed of 172 cases and 200 repeatedly and randomly selected controls) into training (80%) and testing data sets (20%). Next, we used tenfold cross-validation to randomly split the data set into 10 groups, from which 9 groups were used

NATURAL LANGUAGE PROCESSING FOR THE IDENTIFICATION OF
SURGICAL SITE INFECTIONS IN ORTHOPAEDICS

TABLE II Characteristics of Cases and Controls in the Cohort				
	Cases* (N = 172)	Controls† (N = 1,407)	Total (N = 1,579)	P Value‡
Age§ (yr)	44.7 ± 22.1	45.9 ± 20.5	45.8 ± 20.7	0.6
Female sex#	71 (41%)	689 (49%)	760 (48%)	0.1
Race#				0.1
White	142 (83%)	1,231 (87%)	1,373 (87%)	
Black	24 (14%)	124 (9%)	148 (9%)	
Other	6 (3%)	52 (4%)	58 (4%)	
No. of comorbidities**	1.5 (0 to 3)	0 (0 to 1)	0 (0 to 1)	<0.001

*These are the patients who developed surgical site infections in the 90 days following the surgical procedure. †These are the patients who did not develop surgical site infections in the 90 days following the surgical procedure. ‡P values were determined with chi-square tests for categorical variables and with Kruskal-Wallis tests for continuous variables. §The values are given as the mean and the standard deviation. #The values are given as the number of patients, with the percentage in parentheses. **The values are given as the median, with the interquartile range in parentheses. Comorbidities were identified using the Elixhauser comorbidity algorithm.

for training the models, and 1 group was used to test the models. We externally validated the findings by reestimating the 5-gram model (Model 4) to ascertain whether the main analysis findings were robust.

We report the positive predictive value (precision) and sensitivity (recall) for each model because these aid in identifying cases with high precision; these records can subsequently be reviewed by infection prevention personnel. We also focus our discussion on the F1 score, as it is the harmonic mean of the positive predictive value and sensitivity, and the C-statistic for model discrimination.

Results

Characteristics of Cases and Controls

The mean age (and standard deviation) of the overall cohort was 45.8 ± 20.7 years, 48% (n = 760) were female, and 87% (n = 1,373) were white (Table II). Of the 208 surgical site infection cases (172 from the main analytic cohort and 36 from the external validation cohort), 72% (n = 150) met deep or organ or space surgical site infection criteria and 26% (n = 53) met superficial surgical site infection criteria. Twenty percent (n = 42) of the cases had undergone spine surgical procedures, 18% (n = 37) were trauma cases, and 17% (n = 36) underwent joint replacement surgical procedures (Table III).

Distribution of Keywords Among Cases and Controls

Figure 1 presents the heat maps for the frequency distributions of the keywords among the cases and controls. Notably, among the cases, keywords such as "wound," "pain," "infection," and "fever" were the more frequently occurring words. Among the controls, although "pain" was the more frequently occurring keyword, we did not note mentions of other keywords.

Performance Parameters of Models Examining the Variation in Surgical Site Infections

For Model 1 (variables from administrative data only), the sensitivity was 68% and the positive predictive value was 70%; for Model 4 (with NLP 5-grams from the medical record),

the sensitivity was 97% and the positive predictive value was 97%; and for Model 7 (a combination of Models 1 and 4), the sensitivity was 97% and the positive predictive value was 97%. Thus, NLP-based models identified 97% of surgical site infections identified by manual abstraction with high precision and 43% more surgical site infections compared with models that used administrative data only.

Model 1 had the lowest F1 score (66%) and C-statistic (78% ± 7%) (Table IV). The F1 scores and C-statistic for Models 2 to 4 were comparable with those from Models 5 to 7 and were far superior to those from Model 1; the F1 score was 93% for Model 2, 95% for Model 3, and 97% for Model 4, and the C-statistic was 95% ± 4% for Model 2, 97% ± 4% for Model 3, and 96% ± 3% for Model 4. Figure 2 presents the C-statistic curves for Model 1, which had the lowest C-statistic, and Model

TABLE III Distribution of Surgical Subspecialties Among the Surgical Site Infection Cases from the Main Analytic Cohort and the Cohort Used for External Validation (N = 208)*	
Orthopaedic Subspecialty	No. of Cases†
Spine	42 (20%)
Trauma	37 (18%)
Reconstructive surgery‡	36 (17%)
Oncology	26 (13%)
Foot and ankle	20 (10%)
Upper extremity	18 (9%)
Sports	16 (8%)
Other	13 (6%)

*The 172 cases for the main analysis were obtained from Strong Memorial Hospital. †The values are given as the number of cases, with the percentage in parentheses. ‡The 36 cases for reconstructive surgery were obtained from another hospital affiliated with Strong Memorial Hospital and were used for external validation of the prediction models.
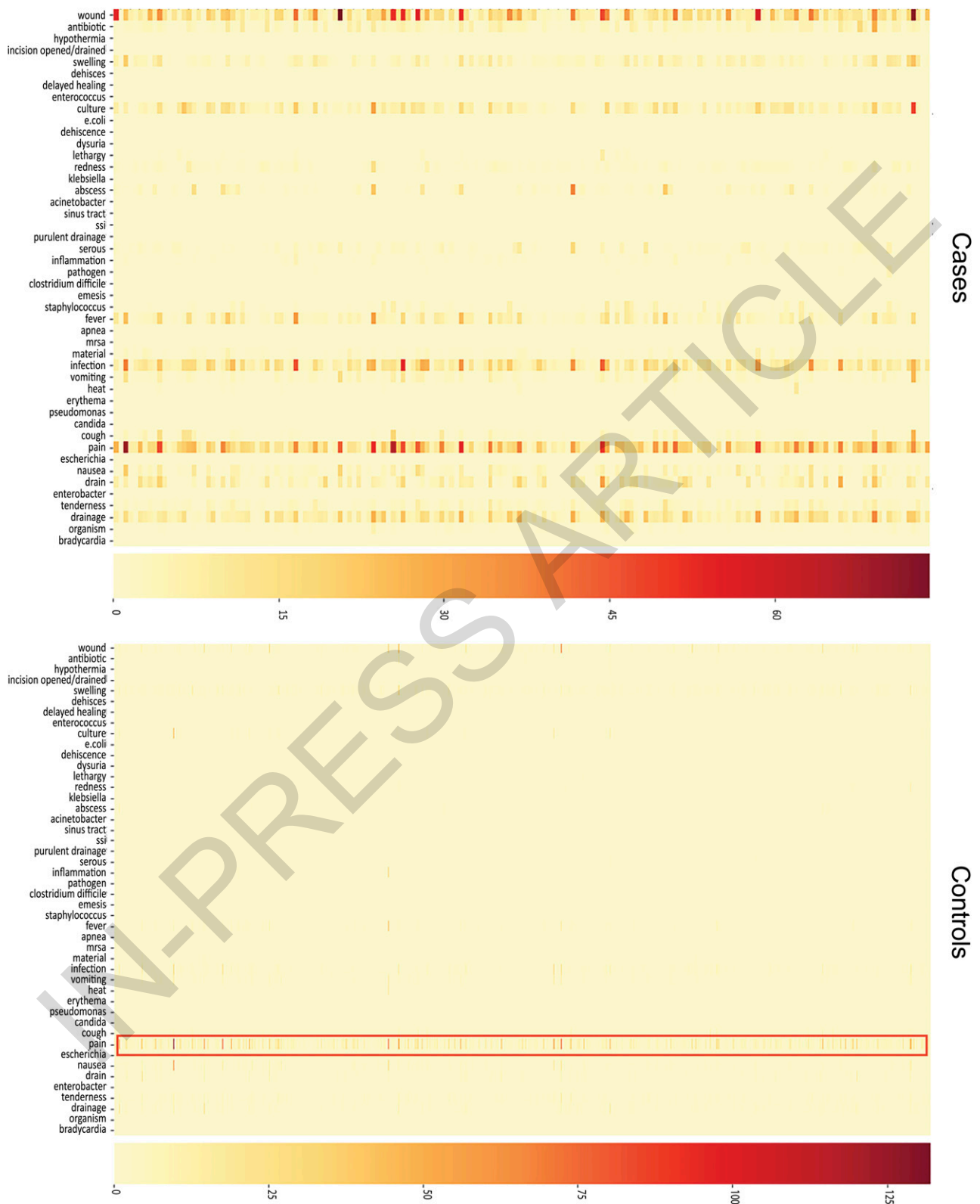
Fig. 1

Heat maps for the distribution of keywords among cases and controls. The heat map pattern for controls demonstrates presence of the keyword "pain" (highlighted with the red box) and its grammatical variants, and the absence of other keyword patterns noted in cases. E. coli = *Escherichia coli,* mrsa = methicillin-resistant *Staphylococcus aureus,* and ssi = surgical site infection.

TABLE IV Model Performance Parameters for the Logistic Regression Models				
	Positive Predictive Value*	Sensitivity†	F1 Score‡	C-Statistic§
Model 1 (variables from administrative data)	70%	68%	66%	78% ± 7%
Model 2 (NLP keywords only)	93%	93%	93%	95% ± 4%
Model 3 (stemmed and lemmatized versions of NLP keywords only)	95%	95%	95%	97% ± 4%
Model 4 (5-grams only)	97%	97%	97%	96% ± 3%
Model 5 (Model 1 and Model 2)	92%	92%	92%	97% ± 2%
Model 6 (Model 1 and Model 3)	96%	96%	96%	98% ± 2%
Model 7 (Model 1 and Model 4)	97%	97%	97%	96% ± 3%

*This is the number of true positives divided by the sum of the number of true and false positives. †This is the number of true positives divided by the sum of the number of true positives and false negatives. ‡This is the harmonic mean of the positive predictive value and sensitivity and is calculated as $\frac{2 \times Positive\ Predictive\ Value \times Sensitivity}{(Positive\ Predictive\ Value + Sensitivity)}$. §This is the ability of the model to discriminate between cases (patients with surgical site infection) and controls (patients without surgical site infection) and is represented using the ROC curve.

7. Notably, during external validation, Model 4 (which included the 5-grams) had a sensitivity of 94%.

Discussion

The objective of our study was to compare the predictive capabilities of rule-based NLP models to predict surgical site infections with those of manual abstraction of surgical site infections from clinical records, and with models using variables from administrative data only. We demonstrated that the NLP models achieved high F1 scores and C-statistics that were comparable with the manual abstraction process and far superior to administrative data models. The addition of the administrative variables to these models only marginally increased the predictive abilities.

The manual process of surveillance for surgical site infection prevention can be time-consuming and labor-intensive and may produce variable results[31-33]. With the increasing availability of electronic medical record data, the use of electronic surveillance systems for the identification of hospital-acquired infections is on the rise. However, these automated systems, either commercial or institution-specific, have mostly used the discrete fields in the medical record, such as microbiology results or antimicrobial drug administration records, to reliably identify surgical site infections[32,34-37]. These studies have shown sensitivities in the range of 60% to 97% and positive predictive values in the range of 33% to 97%, demonstrating considerable variation in the predictive accuracies of these models[32,33].
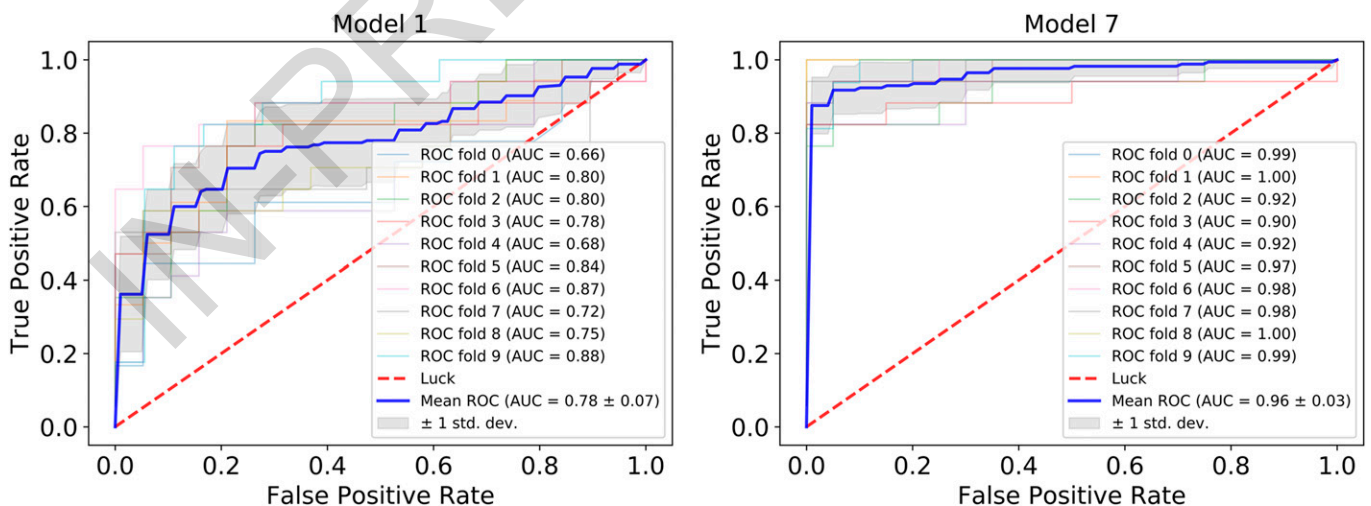

Fig. 2

Receiver operating characteristic (ROC) curves for key analytic models. ROC curves for Model 1 include variables from the administrative data only, and ROC curves for Model 7 include a combination of administrative data and 5-grams of NLP-based features and variables. These curves compare the ability of the 2 models to discriminate between cases and controls (patients who did and did not develop surgical site infections). AUC = area under the curve, and std dev = standard deviation.

The use of NLP for surgical site infection identification is highly relevant, yet under-investigated, given the clinical guidelines that are used to identify surgical site infections. Many of the mentions of surgical site infection signs and symptoms are captured in the free text documented by the surgeons and other members of the clinical team. Moreover, 97% of surgical site infections occur after a patient's discharge from the hospital[10]. The starts of many of these post-discharge events are telephone encounters or other communications between the patient and the caregiving team that are not captured as discrete fields in the clinical records, thereby highlighting the importance of the use of NLP for identifying surgical site infections. Our study demonstrates that the use of NLP for identifying surgical site infections is feasible and produces results that are comparable with those from the manual abstraction process. The high sensitivity and positive predictive value of these models demonstrate that there were few false negatives and that patients flagged as cases by the NLP algorithms had a high likelihood of having surgical site infections. Furthermore, the high C-statistics of the NLP-based models represent their excellent ability to differentiate between cases and controls. The relatively unimportant changes in the performance parameters of the NLP models with the addition of the administrative variables provide evidence that the use of an elaborate lexicon of keywords and their variants that codifies clinical conversations may suffice for predicting surgical site infections. A key strength of our approach is that because the NHSN criteria for identifying surgical site infections remain fairly consistent across surgical specialties, our lexicon of keywords could be used for non-orthopaedic surgical procedures.

Few other studies have used varying NLP approaches for the surveillance of postoperative infections. FitzHenry et al. used U.S. Veterans Affairs Surgical Quality Improvement Program data from 1999 to 2006 and demonstrated that models using a combination of structured data, codified SNOMED CT (Systematized Nomenclature of Medicine Clinical Terms) concepts, and word phrases achieved sensitivities of about 77% for the prediction of wound infections[20]. Murff et al. used a similar approach to compare the predictive accuracy of NLP-based models with administrative data-based Patient Safety Indicators for identifying postoperative complications (excluding surgical site infections). They found model sensitivities ranging from 59% for pulmonary embolism or deep vein thrombosis to 91% for myocardial infarction[17]. Chapman et al. used rule-based NLP with radiographic reports to identify intra-abdominal surgical site infections among patients who underwent a gastrointestinal surgical procedure and achieved a positive predictive value of 91% and sensitivity of 89%[21]. Although these studies examined large populations without accounting for the clinical heterogeneity between surgical specialties[20] or they focused on specific clinical reports such as radiology reports[21] or discharge summaries[38], they highlighted the benefits of using NLP over other traditional methods for the identification of postoperative complications.

The successful use of NLP for surgical site infection identification has several practice and policy implications. The identification of most true positive cases by NLP-based models can serve as a valuable complement to manual surveillance, thereby reducing the time and effort that infection prevention personnel devote to surveillance activities[31-33]. This is especially valuable with the onset of payment reforms that have introduced rewards and penalties for avoiding or failing to avoid surgical site infections and related readmissions in orthopaedics. Furthermore, by integrating validated NLP algorithms with the electronic health record system, the advantages of these models can be extended to other complications such as sepsis, to other clinical specialties, and to other settings such as the outpatient setting. The use of valid NLP models may limit the need to hire and train dedicated manpower for manually abstracting records to generate research databases whose analysis informs the design of preventive interventions.

Our study had limitations. First, the analytic models used data from a single academic medical center. Hence, the generalizability of the findings is limited. Nevertheless, our findings provide proof of concept and support an innovative approach to identify surgical site infections, which other medical centers could adopt for achieving efficiencies and supporting the ongoing surgical site infection surveillance process. Second, the common practice of copying and pasting clinical text in progress notes is an important concern. Although our analysis used the free text of medical records, the findings demonstrate that the predictive accuracy of the NLP models was comparable with the manual abstraction models, which should allay this concern[39]. Third, the analytic models were designed to identify whether or not a surgical site infection occurred. In this study, we did not design algorithms that would differentiate between the types of surgical site infections (superficial, deep, or organ or space). Future work that focuses on the use of NLP for differentiating the types of surgical site infections will be valuable.

In conclusion, our study demonstrated the feasibility and validity of using rule-based NLP models with the free text of medical records to retrospectively identify surgical site infections among patients who underwent an orthopaedic surgical procedure. The use of an NLP-based automated system to support ongoing manual efforts can improve the efficiency, effectiveness, timeliness, and scalability of surgical site infection surveillance programs in hospitals.

Appendix

(eA) Supporting material provided by the authors is posted with the online version of this article as a data supplement at jbjs.org (http://links.lww.com/XXXXXXX). ■

Caroline P. Thirukumaran, MBBS, MHA, PhD[1]
Anis Zaman, MA[1]
Paul T. Rubery, MD[1]
Casey Calabria, CIC, BSN-RN, MBA[1]
Yue Li, PhD[1]

Benjamin F. Ricciardi, MD[1]
Wajeeh R. Bakhsh, MD[1]
Henry Kautz, PhD[1]

[1]University of Rochester, Rochester, New York

Email address for C.P. Thirukumaran: caroline_thirukumaran@
urmc.rochester.edu

ORCID iD for C.P. Thirukumaran: 0000-0003-3316-0951
ORCID iD for A. Zaman: 0000-0003-1061-6928
ORCID iD for P.T. Rubery: 0000-0003-2576-3242
ORCID iD for C. Calabria: 0000-0003-0640-1529
ORCID iD for Y. Li: 0000-0001-7592-0299
ORCID iD for B.F. Ricciardi: 0000-0001-9384-4840
ORCID iD for W.R. Bakhsh: 0000-0002-0284-9994
ORCID iD for H. Kautz: 0000-0001-5219-2970

References

1. Anderson DJ, Podgorny K, Berríos-Torres SI, Bratzler DW, Dellinger EP, Greene L, Nyquist AC, Saiman L, Yokoe DS, Maragakis LL, Kaye KS. Strategies to prevent surgical site infections in acute care hospitals: 2014 update. Infect Control Hosp Epidemiol. 2014 Jun;35(6):605-27.

2. Zimlichman E, Henderson D, Tamir O, Franz C, Song P, Yamin CK, Keohane C, Denham CR, Bates DW. Health care-associated infections: a meta-analysis of costs and financial impact on the US health care system. JAMA Intern Med. 2013 Dec 9-23;173(22):2039-46.

3. Magill SS, Edwards JR, Bamberg W, Beldavs ZG, Dumyati G, Kainer MA, Lynfield R, Maloney M, McAllister-Hollod L, Nadle J, Ray SM, Thompson DL, Wilson LE, Fridkin SK; Emerging Infections Program Healthcare-Associated Infections and Antimicrobial Use Prevalence Survey Team. Multistate point-prevalence survey of health care-associated infections. N Engl J Med. 2014 Mar 27;370(13):1198-208.

4. Greene L, Mills R, Moss R, Sposato K, Vignari M. Guide to the elimination of orthopedic surgical site infections. APIC Guide. 2010.

5. Cram P, Lu X, Kates SL, Singh JA, Li Y, Wolf BR. Total knee arthroplasty volume, utilization, and outcomes among Medicare beneficiaries, 1991-2010. JAMA. 2012 Sep 26;308(12):1227-36.

6. Kurtz SM, Ong KL, Lau E, Bozic KJ, Berry D, Parvizi J. Prosthetic joint infection risk after TKA in the Medicare population. Clin Orthop Relat Res. 2010 Jan;468(1):52-6. Epub 2009 Aug 8.

7. Chahoud J, Kanafani Z, Kanj SS. Surgical site infections following spine surgery: eliminating the controversies in the diagnosis. Front Med (Lausanne). 2014 Mar 24; 1:7.

8. Koutsoumbelis S, Hughes AP, Girardi FP, Cammisa FP Jr, Finerty EA, Nguyen JT, Gausden E, Sama AA. Risk factors for postoperative infection following posterior lumbar instrumented arthrodesis. J Bone Joint Surg Am. 2011 Sep 7;93(17): 1627-33.

9. Schwarz EM, Parvizi J, Gehrke T, Aiyer A, Battenberg A, Brown SA, Callaghan JJ, Citak M, Egol K, Garrigues GE, Ghert M, Goswami K, Green A, Hammound S, Kates SL, McLaren AC, Mont MA, Namdari S, Obremskey WT, O'Toole R, Raikin S, Restrepo C, Ricciardi B, Saeed K, Sanchez-Sotelo J, Shohat N, Tan T, Thirukumaran CP, Winters B. 2018 International Consensus Meeting on Musculoskeletal Infection: research priorities from the general assembly questions. J Orthop Res. 2019 May; 37(5):997-1006. Epub 2019 Apr 25.

10. Merkow RP, Ju MH, Chung JW, Hall BL, Cohen ME, Williams MV, Tsai TC, Ko CY, Bilimoria KY. Underlying reasons associated with hospital readmission following surgery in the United States. JAMA. 2015 Feb 3;313(5):483-95.

11. U.S. Centers for Medicare and Medicaid Services. Comprehensive Care for Joint Replacement Model. 2019. https://innovation.cms.gov/initiatives/CJR. Accessed 2019 Aug 14.

12. U.S. Centers for Medicare and Medicaid Services. Hospital Readmissions Reduction Program. https://www.cms.gov/medicare/medicare-fee-for-service-payment/acuteinpatientpps/readmissions-reduction-program.html. Accessed 2019 Aug 14.

13. Lee TB, Montgomery OG, Marx J, Olmsted RN, Scheckler WE; Association for Professionals in Infection Control and Epidemiology. Recommended practices for surveillance: Association for Professionals in Infection Control and Epidemiology (APIC), Inc. Am J Infect Control. 2007 Sep;35(7):427-40.

14. van Mourik MS, Troelstra A, van Solinge WW, Moons KG, Bonten MJ. Automated surveillance for healthcare-associated infections: opportunities for improvement. Clin Infect Dis. 2013 Jul;57(1):85-93. Epub 2013 Mar 26.

15. van Mourik MS, van Duijn PJ, Moons KG, Bonten MJ, Lee GM. Accuracy of administrative data for surveillance of healthcare-associated infections: a systematic review. BMJ Open. 2015 Aug 27;5(8):e008424.

16. Nadkarni PM, Ohno-Machado L, Chapman WW. Natural language processing: an introduction. J Am Med Inform Assoc. 2011 Sep-Oct;18(5):544-51.

17. Murff HJ, FitzHenry F, Matheny ME, Gentry N, Kotter KL, Crimin K, Dittus RS, Rosen AK, Elkin PL, Brown SH, Speroff T. Automated identification of postoperative complications within an electronic medical record using natural language processing. JAMA. 2011 Aug 24;306(8):848-55.

18. Taggart M, Chapman WW, Steinberg BA, Ruckel S, Pregenzer-Wenzler A, Du Y, Ferraro J, Bucher BT, Lloyd-Jones DM, Rondina MT, Shah RU. Comparison of 2 natural language processing methods for identification of bleeding among critically ill patients. JAMA Netw Open. 2018 Oct 5;1(6):e183451.

19. U.S. Centers for Disease Control and Prevention. National Healthcare Safety Network patient safety component manual. 2019 Jan. https://www.cdc.gov/nhsn/pdfs/pscmanual/pcsmanual_current.pdf. Accessed 2019 Aug 14.

20. FitzHenry F, Murff HJ, Matheny ME, Gentry N, Fielstein EM, Brown SH, Reeves RM, Aronsky D, Elkin PL, Messina VP, Speroff T. Exploring the frontier of electronic health record surveillance: the case of postoperative complications. Med Care. 2013 Jun;51(6):509-16.

21. Chapman AB, Mowery DL, Swords DS, Chapman WW, Bucher BT. Detecting evidence of intra-abdominal surgical site infections from radiology reports using natural language processing. AMIA Annu Symp Proc. 2018 Apr 16;2017:515-24.

22. U.S. Centers for Disease Control and Prevention. National Healthcare Safety Network. 2017. http://www.cdc.gov/nhsn/. Accessed 2019 Aug 14.

23. Sammut C, Webb GI. Encyclopedia of machine learning. New York: Springer Science & Business Media; 2011.

24. StataCorp. Stata Statistical Software: Release 15. StataCorp LP; 2018.

25. Van Rossum G, Drake FL. Python 3 reference manual. CreateSpace; 2009.

26. Code G. negex - NegExTerms.wiki. https://code.google.com/archive/p/negex/wikis/NegExTerms.wiki. Accessed 2019 Aug 14.

27. Chapman WW, Hillert D, Velupillai S, Kvist M, Skeppstedt M, Chapman BE, Conway M, Tharp M, Mowery DL, Deleger L. Extending the NegEx lexicon for multiple languages. Stud Health Technol Inform. 2013;192:677-81.

28. Chapman WW, Bridewell W, Hanbury P, Cooper GF, Buchanan BG. A simple algorithm for identifying negated findings and diseases in discharge summaries. J Biomed Inform. 2001 Oct;34(5):301-10.

29. Quan H, Sundararajan V, Halfon P, Fong A, Burnand B, Luthi JC, Saunders LD, Beck CA, Feasby TE, Ghali WA. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care. 2005 Nov;43(11):1130-9.

30. U.S. Agency for Healthcare Research and Quality. HCUP: Clinical Classification Software for ICD-9-CM. 2012 Jan. http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccsfactsheet.jsp. Accessed 2019 Aug 14.

31. Russo PL, Shaban RZ, Macbeth D, Carter A, Mitchell BG. Impact of electronic healthcare-associated infection surveillance software on infection prevention resources: a systematic review of the literature. J Hosp Infect. 2018 May;99(1):1-7. Epub 2017 Sep 8.

32. de Bruin JS, Seeling W, Schuh C. Data use and effectiveness in electronic surveillance of healthcare associated infections in the 21st century: a systematic review. J Am Med Inform Assoc. 2014 Sep-Oct;21(5):942-51. Epub 2014 Jan 13.

33. Freeman R, Moore LS, García Álvarez L, Charlett A, Holmes A. Advances in electronic surveillance for healthcare-associated infections in the 21st century: a systematic review. J Hosp Infect. 2013 Jun;84(2):106-19. Epub 2013 May 4.

34. Sips ME, Bonten MJM, van Mourik MSM. Semiautomated surveillance of deep surgical site infections after primary total hip or knee arthroplasty. Infect Control Hosp Epidemiol. 2017 Jun;38(6):732-5. Epub 2017 Apr 3.

35. Bolon MK, Hooper D, Stevenson KB, Greenbaum M, Olsen MA, Herwaldt L, Noskin GA, Fraser VJ, Climo M, Khan Y, Vostok J, Yokoe DS; Centers for Disease Control and Prevention Epicenters Program. Improved surveillance for surgical site infections after orthopedic implantation procedures: extending applications for automated data. Clin Infect Dis. 2009 May 1;48(9):1223-9.

36. Inacio MC, Paxton EW, Chen Y, Harris J, Eck E, Barnes S, Namba RS, Ake CF. Leveraging electronic medical records for surveillance of surgical site infection in a total joint replacement population. Infect Control Hosp Epidemiol. 2011 Apr;32(4): 351-9.

37. Halpin H, Shortell SM, Milstein A, Vanneman M. Hospital adoption of automated surveillance technology and the implementation of infection prevention and control programs. Am J Infect Control. 2011 May;39(4):270-6.

38. Melton GB, Hripcsak G. Automated detection of adverse events using natural language processing of discharge summaries. J Am Med Inform Assoc. 2005 Jul-Aug;12(4):448-57. Epub 2005 Mar 31.

39. Sheehy AM, Weissburg DJ, Dean SM. The role of copy-and-paste in the hospital electronic health record. JAMA Intern Med. 2014 Aug;174(8):1217-8.