Mining peripheral arterial disease cases from narrative clinical notes using natural language processing

Naveed Afzal, PhD,[a] Sunghwan Sohn, PhD,[a] Sara Abram, MD,[b] Christopher G. Scott, MS,[a] Rajeev Chaudhry, MBBS, MPH,[c] Hongfang Liu, PhD,[a] Iftikhar J. Kullo, MD,[b] *and* Adelaide M. Arruda-Olson, MD, PhD,[b] *Rochester, Minn*

ABSTRACT

Objective: Lower extremity peripheral arterial disease (PAD) is highly prevalent and affects millions of individuals worldwide. We developed a natural language processing (NLP) system for automated ascertainment of PAD cases from clinical narrative notes and compared the performance of the NLP algorithm with billing code algorithms, using ankle-brachial index test results as the gold standard.

Methods: We compared the performance of the NLP algorithm to (1) results of gold standard ankle-brachial index; (2) previously validated algorithms based on relevant International Classification of Diseases, Ninth Revision diagnostic codes (simple model); and (3) a combination of International Classification of Diseases, Ninth Revision codes with procedural codes (full model). A dataset of 1569 patients with PAD and controls was randomly divided into training ($n = 935$) and testing ($n = 634$) subsets.

Results: We iteratively refined the NLP algorithm in the training set including narrative note sections, note types, and service types, to maximize its accuracy. In the testing dataset, when compared with both simple and full models, the NLP algorithm had better accuracy (NLP, 91.8%; full model, 81.8%; simple model, 83%; $P < .001$), positive predictive value (NLP, 92.9%; full model, 74.3%; simple model, 79.9%; $P < .001$), and specificity (NLP, 92.5%; full model, 64.2%; simple model, 75.9%; $P < .001$).

Conclusions: A knowledge-driven NLP algorithm for automatic ascertainment of PAD cases from clinical notes had greater accuracy than billing code algorithms. Our findings highlight the potential of NLP tools for rapid and efficient ascertainment of PAD cases from electronic health records to facilitate clinical investigation and eventually improve care by clinical decision support. (J Vasc Surg 2017;65:1753-61.)

Peripheral arterial disease (PAD) is a chronic disease associated with high morbidity and mortality.[1-3] PAD affects at least 8.5 million people in the United States and in excess of 200 million people worldwide.[4] PAD is associated with increased risk for death, myocardial infarction, and stroke with annual risk for adverse cardiovascular events exceeding 5%.[1-3,5-7] Despite high prevalence and associated mortality, morbidity, and cost, PAD has received relatively little attention from clinical researchers, health systems, and government agencies.[2,8,9] The diagnosis of PAD is based on abnormal ankle-brachial index (ABI). However, not all PAD cases have ABI results available in their electronic health records (EHRs). In the absence of ABI results, time-consuming and laborious manual abstraction of narrative clinical notes is needed to ascertain PAD status.

Previously, we used billing code algorithms composed of PAD-related International Classification of Diseases, Ninth Revision (ICD-9) codes (simple model) or a combination of PAD-related ICD-9 codes with procedural codes (full model) to identify patients with PAD.[10] When applied to a community-based sample, these billing algorithms had limited performance.[10] In another prior study, we successfully developed and applied a natural language processing (NLP) algorithm to ascertain PAD status from radiology reports; however, radiology reports describe the results of radiology tests and

From the Department of Health Sciences Research,[a] Department of Cardiovascular Diseases,[b] Division of Primary Care Medicine, Knowledge Delivery Center and Center for Innovation,[c] Mayo Clinic.

Author conflict of interest: none.

Research reported in this publication was supported by the National Heart, Lung, and Blood Institute of the National Institutes of Health (award K01HL124045) and the NHGRI eMERGE (Electronic Records and Genomics) Network grants HG04599 and HG006379. This study was made possible using the resources of the Rochester Epidemiology Project supported by the National Institute on Aging of the National Institutes of Health (award R01AG034676) and the NLP framework established through the NIGMS award R01GM102283A1. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Correspondence: Adelaide M. Arruda-Olson, MD, PhD, Department of Cardiovascular Diseases, Mayo Clinic, 200 First St SW, Rochester, MN 55905 (e-mail: olson.adelaide@mayo.edu).

The editors and reviewers of this article have no relevant financial relationships to disclose per the JVS policy that requires reviewers to decline review of any manuscript for which they may have a conflict of interest.

0741-5214

http://dx.doi.org/10.1016/j.jvs.2016.11.031

do not contain the key components of the clinical notes such as impression, report, and plan of care.[11] To address these shortcomings, we tested the hypothesis that NLP of narrative clinical notes would improve accuracy of PAD ascertainment over billing code algorithms using ABI test results as the gold standard. In this study, we develop a NLP algorithm for automated ascertainment of PAD cases from clinical narrative notes and compare the performance of the NLP algorithm to billing code algorithms and gold standard ABI test results.

ARTICLE HIGHLIGHTS

- **Type of Research:** Retrospective case-control study
- **Take Home Message:** A knowledge-driven natural language processing algorithm for automatic ascertainment of peripheral arterial disease cases from clinical notes had greater accuracy than billing code algorithms.
- **Recommendation:** The authors recommend that natural language processing extraction algorithms continue to be explored to enhance the detection of peripheral arterial disease from large sets of clinical notes.

METHODS

Study setting and population. The study was conducted at Mayo Clinic, Rochester Minnesota and used the resources of the Rochester Epidemiology Project (REP) to assemble a community-based PAD case-control cohort from Olmsted County.[12] The REP consists of Mayo Clinic and the Mayo Clinic Hospitals, Olmsted Medical Center, and its affiliated hospitals. The REP is an integrated health information system that links medical records of all Olmsted County residents regardless of their ethnicity, socioeconomic, or insurance status.[12] In the present study, we applied this NLP algorithm to the Mayo clinical data warehouse. For this study, we obtained patient informed consent, and this study was approved by the institutional review boards of participating medical centers.

Gold standard. All patients from both datasets had undergone ABI testing in the Mayo noninvasive vascular laboratory using standardized protocols.[4] The ABI results were reported in pdf format and were not part of the narrative clinical notes. In brief, the systolic blood pressure was measured in each arm and dorsalis pedis and posterior tibial arteries bilaterally using a hand-held 8.3-MHz Doppler probe. The higher of the two-arm pressures and lower of the two-ankle pressures were used to calculate the ABI for each leg.[3] Normal ABI was defined as 1.0-1.3. PAD was defined as an ABI ≤0.9 at rest or 1 minute after exercise; or by the presence of poorly compressible arteries (ABI ≥1.40 or ankle systolic blood pressure >255 mm Hg).[4] These criteria were used to classify all subjects into case or control categories.

Dataset. The dataset consisted of 1569 patients (806 cases and 763 controls) (Fig 1). We randomly divided this dataset into two subsets: training and testing. The training dataset consisted of 935 patients and 300,364 clinical notes; there were 479 PAD cases (abnormal ABI) and 456 controls (normal ABI). The testing dataset comprised 634 patients, 212,047 clinical notes, and included 327 PAD cases and 307 controls.

Study design. We retrieved all clinical notes of the subjects participating in this study from the Mayo data warehouse created through June 2015. We applied the NLP algorithm to these retrieved clinical notes to ascertain PAD status as an output for each patient (Fig 2). We developed and conducted iterative refinement of an NLP algorithm in the training dataset. For subsequent validation, we applied the best version of the refined NLP algorithm to the testing dataset. For each dataset, we compared the performance of NLP algorithm with each billing code algorithm (simple model, full model) and NLP algorithm with the gold standard. The simple model was composed of PAD-related ICD-9 codes, whereas the full model was a combination of both PAD-related ICD-9 codes and procedural codes.[10]

NLP algorithm. The NLP algorithm was knowledge-driven and had two main components: text processing and patient classification (Fig 2). The text processing component found PAD-related concepts (the keywords listed in Table I) in the text using MedTagger, an open source clinical NLP pipeline that analyzed text and identified PAD-related medical concepts.[13] The NLP algorithm extracted PAD-related concepts from clinical notes and mapped them to the specific categories. For example, NLP algorithm identified a concept "lower extremity" from clinical notes and then mapped it to the category "Disease location III" (Table I). The NLP algorithm also checked assertion status of each concept that included certainty (ie, positive, negative, and possible), temporality (historical or current) along with experiencer (ie, associated with the patient or someone else). For example, if the NLP algorithm came across a sentence: "noninvasive studies are consistent with severe arterial occlusive disease of bilateral lower extremities," the system identifies the concepts "arterial occlusive disease" and "lower extremities" along with the corresponding assertion status (ie, arterial occlusive disease) is stated positively (certainty = positive), present (temporality = current), and associated with the patient (experiencer = patient). The patient classification

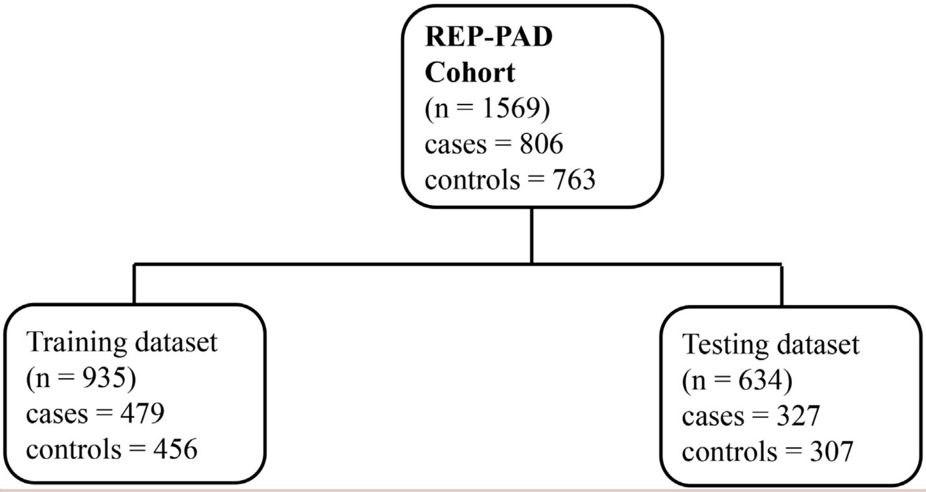

Fig 1. Dataset description. *PAD,* Peripheral arterial disease; *REP,* Rochester Epidemiology Project.

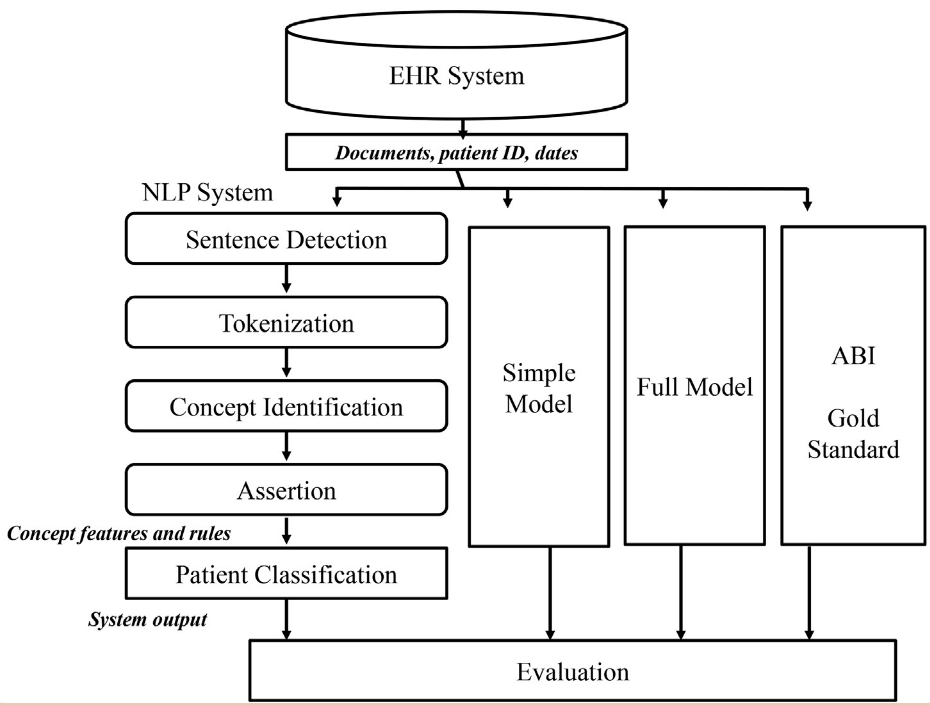

Fig 2. Study design. *ABI,* Ankle-brachial index; *EHR,* electronic health record; *ID,* identification; *NLP,* natural language processing.

component used a set of rules (described below) to classify the status of each individual.

Keywords were used to create comprehensive lists of appropriate concepts for ascertainment of PAD status. To identify the PAD-related keywords, a clinician comprehensively abstracted the narrative clinical notes of 20 PAD cases and 20 controls (without PAD). These notes were excluded for the next steps of NLP algorithm evaluation. Cardiovascular experts compiled the list of PAD-related concepts using clinical notes, which was expanded by addition of synonyms. The PAD-related concepts and the rules for patient classification were refined using an interactive process with analysis of false positives and false negatives in the training dataset. Fig 3 shows the PAD-related concept types and values of a sample clinical note. The right window shows the clinical note snippet that is processed by the NLP algorithm to populate annotations (PAD-related concepts were shown in yellow color in the right window) as they appear in the left window.

Table I. Peripheral arterial disease (PAD)-related keywords for ascertainment of PAD status

Confirmation keywords—disease location I
tibial/Iliac/femoral/popliteal; lle; rle; distal/infrarenal/abdominal aorta; aorto biiliac/bifermoral/iliac/femoral; calcaneal region; calx; hock/hockings; below/above knee; foot/feet; toe/toes; shin; anterior leg region; anterior part leg; plantar; heel; ankle; interdigital
Confirmation keywords—disease location II
tibial/Iliac/femoral/popliteal artery/arteries; sfa; dfa; cfa; distal/infrarenal/abdominal aorta/aorto (bi)iliac/aorto(bi)iliac/aorto(bi)-iliac; aorto-(bi)femoral
Confirmation keywords—disease location III
lower limb/limbs; lower extremity/extremities; leg/legs
Confirmation keywords—first diagnosis I
ncv (noncompressible vessels); nca (noncompressible arteries); pca (poorly compressible arteries); pcv (poorly compressible vessels); stiff vessels/arteries ischemia; positive abi/ankle brachial index/vascular labs/extremities study/arterial studies; thrombectomy; removal thrombus; thromboembolectomy; thrombosis/thrombose; embolectomy/embolectomies; arterial occlusive disease/occlusion/occluded; stenosis; peripheral arterial occlusive disease; peripheral arterial disease; arterial and venous occlusions; arterial occluded; arterial obstruction; block artery
Confirmation keywords—first diagnosis II
recanalization; angioplasty; pta (percutaneous transluminal angioplasty); stenting/stent; endarterectomy/endarterectomies
Confirmation keywords—first diagnosis III
revascularization; graft; bypass
Confirmation keywords—Second diagnosis
Amputation
Confirmation keywords—third diagnosis
claudication; lameness; leg pain walk; limp; calf/calve pain; ischemic ulcer; aso/arteriosclerosis obliterans; atherosclerotic disease; cramp; pain; discomfort
Exclusion keywords I
family history of; upper extremity/extremities; brachium; brachial region; forelimb; arm between shoulder elbow arm/arms; hand/hands; manus; brachial/axillary/celiac/coronary/cerebrovascular/renal/radial/ulnar/carotid/subclavian/innominate/mesenteric/brachio-cephalic artery/arteries; brachio-cephalic trunk; coronary arterial tree cerebrovascular disease; pseudoclaudication; pseudoclaudicatory pain; aaa (abdominal aortic aneurysm); spinal stenosis; venous thrombosis; thromboembolism; dvt (deep vein thrombosis); vein thrombosis thrombosis/femoral/popliteal/saphenous vein; aortomesenteric; superficial thrombophlebitis; thrombosis venous system; pseudoaneurysm; normal abi
Exclusion keywords II
vascular calcification; varicose veins
Exclusion keywords III
traumatic/trauma; injury wound; sarcoma; osteoma; diabetic foot; hammer toe
Exclusion keywords IV
lower extremity/extremities edema/cellulitis/venous system; carotid artery disease/spinal ischemia; iliac artery aneurysm; spinal/foraminal/lumbar/canal/cervical/carotid stenosis; femoral/popliteal/tibial vein/veins; carotid/cerebrovascular/renal/mesenteric arterial occlusive disease; carotid/renal endarterectomy; coronary/gastric/heart/carotid-to-axillary/sequential saphenous vein/saphenous vein harvest/myocardial infarction bypass; abdominal aortic aneurysm repair graft; capillary/deep pain thrombosis; carotid arteriosclerosis obliterans; renal allograft artery angioplasty

The following rules were used for PAD cases:

- One disease location keyword from disease location I + one diagnostic keyword from first diagnosis I within two sentences anchored by a diagnostic keyword in the same note.
- One disease location keyword from disease location II or disease location III + one diagnostic keyword from first diagnosis II within two sentences anchored by a diagnostic keyword in the same note.
- One disease location keyword from disease location III + one diagnostic keyword from first diagnosis III in the same note.

For controls (without PAD), the system used the following rules:

- If not satisfied, the PAD criteria described above or
- One exclusion keyword from exclusion I + one diagnostic keyword from first diagnosis I.
- One exclusion keyword from exclusion I or exclusion II + one diagnostic keyword from third diagnosis.
- One exclusion keyword from exclusion III + one diagnostic keyword from second diagnosis.
- One exclusion keyword from exclusion IV.

For each case of PAD, the NLP algorithm also provided the note type and index date (ie, the earliest date that

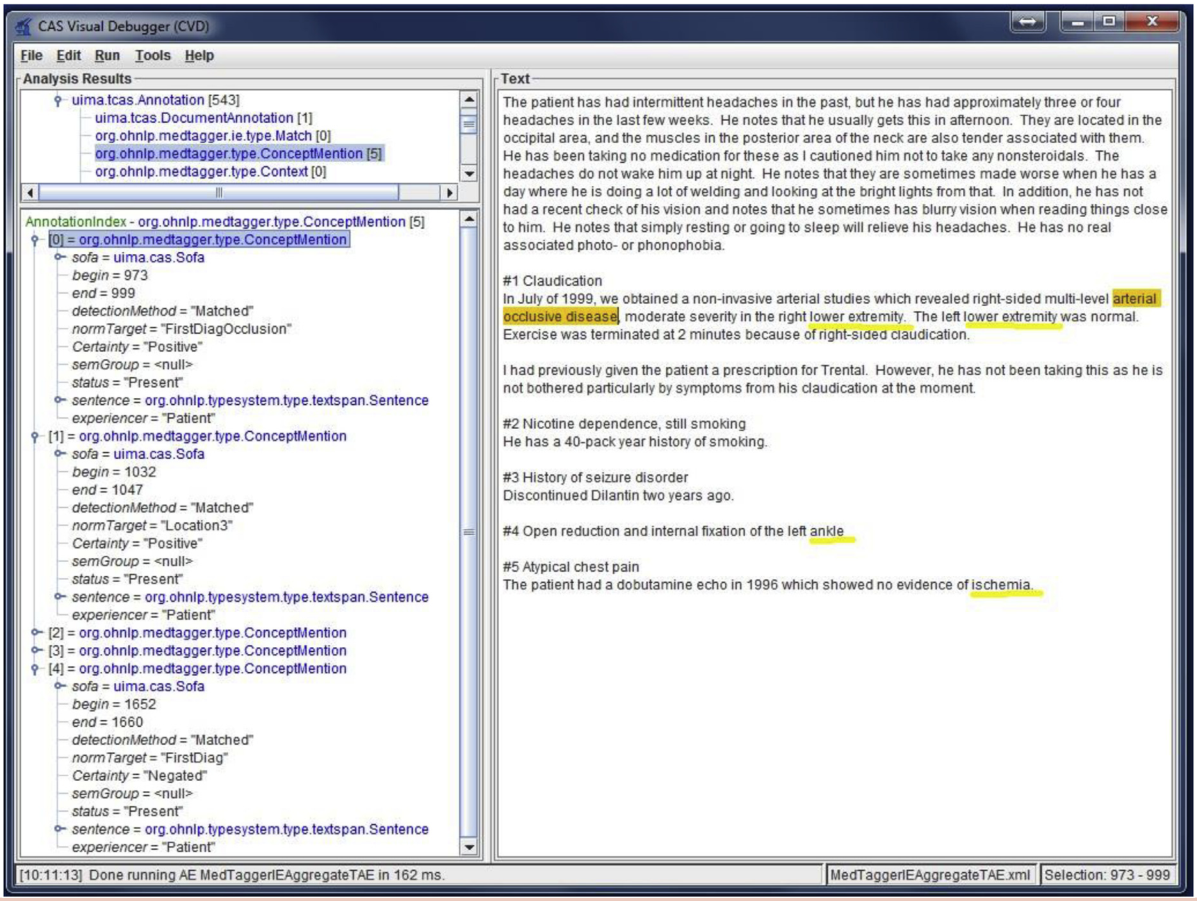

Fig 3. Peripheral arterial disease (PAD) concept visualization.

satisfied PAD conditions) along with evidence in the form of ± two sentences anchored by a diagnostic keyword that led the system to classify a patient as a PAD case.

Statistical analysis. Comparisons between algorithms were made using decision statistics calculated from 2 × 2 tables including positive predictive value (PPV), sensitivity, negative predictive value (NPV), and specificity compared with the gold standard ABI test results. These were calculated as follows: PPV = true positives/(true positives + false positives); sensitivity = true positives/(true positives + false negatives); NPV = true negatives/(true negatives + false negatives) and specificity = true negatives/(true negatives + false positives). Confidence intervals were estimated for each of these measures. Estimates of sensitivity, specificity, and overall accuracy between algorithms were compared using McNemar test. Generalized score statistics were used to compare PPV and NPV. Analyses were performed in SAS v 9.4 (SAS Institute, Cary, NC), and significance was set using a two-sided *P* value of <.05.

RESULTS

Interactive refinement of the NLP algorithm—training dataset. We initially included all clinical notes from patient encounters in the outpatient and inpatient settings, from internal medicine and internal medicine subspecialties, as well as from general surgery and surgical specialties. The clinical notes consist of multiple predefined sections (eg, history of present illness, past medical history, and impression/report/plan). During the iterative refinement of our NLP algorithm, we identified the note types, note sections, and service groups that led to the most false results; these were excluded from subsequent experiments and are listed in Appendix I.

Using this stepwise approach for the interactive refinement of the NLP algorithm, there was improvement of specificity, PPV, and accuracy of the system compared with the gold standard (Table II). Version E had the best performance and was subsequently applied to the testing dataset. The note types, note sections, and service groups included in version E are listed in Appendix II.

Comparison of NLP algorithm with billing code algorithms. Compared with billing code algorithms, the NLP algorithm had the highest accuracy in each of the datasets (Fig 4; Table III).

In the training dataset, the NLP algorithm showed high sensitivity, specificity, PPV, NPV, and accuracy. In the testing dataset, the NLP algorithm had better specificity, PPV, and accuracy when compared with both simple

Table II. Iterative refinement of natural language processing (NLP) system

	Training dataset				
	Sensitivity, % (95% CI)	Specificity, % (95% CI)	PPV, % (95% CI)	NPV, % (95% CI)	Accuracy, % (95% CI)
Version A: Keeping all note types and note sections	95.4 (93.5-97.3)	78.5 (74.7-82.3)	82.3 (79.2-85.5)	94.2 (91.8-96.6)	87.2 (85.0-89.3)
Version B: Excluding selected note types (Appendix I)	95.0 (93.0-96.9)	81.8 (78.3-85.3)	84.6 (81.5-87.6)	94.0 (91.6-96.3)	88.6 (86.5-90.6)
Version C: Excluding selected note sections (Appendix I)	91.6 (89.2-94.1)	87.9 (84.9-90.9)	88.9 (86.1-91.6)	90.9 (88.2-93.6)	89.8 (87.9-91.8)
Version E: Excluding selected note types, note sections and service groups (Appendix I)	91.4 (88.9-93.9)	94.1 (91.9-96.2)	94.2 (92.1-96.3)	91.3 (88.7-93.8)	92.7 (91.1-94.4)

CI, Confidence interval; *NPV,* negative predictive value; *PPV,* positive predictive value.

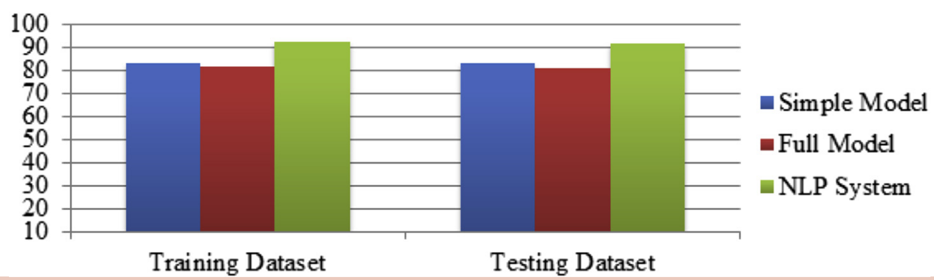

Fig 4. Accuracy of natural language processing (*NLP*) algorithm compared with billing code algorithms (simple model and full model) for ascertainment of peripheral arterial disease (PAD) status.

and full models (Table III). However, the NLP algorithm had similar sensitivity and NPV to the simple model, whereas the full model had higher sensitivity and NPV than the NLP algorithm.

DISCUSSION

The use of ICD-9 billing codes to ascertain phenotypes may have less than optimal accuracy.[14,15] We developed an NLP algorithm that was more accurate than billing code algorithms for identification of PAD cases from the EHRs.[10] We iteratively refined our NLP algorithm, in collaboration with clinician experts, and in a comprehensive stepwise interactive approach we identified the note sections, note types, and service types that generated the highest numbers of false results compared with the gold standard ABI results. For example, we found that more false results were generated from the note section "chief complaint" whereas less false results were generated from the note section "impression/report/plan." Importantly, in this section clinicians summarize the pertinent findings that support the plan of care, which is described in the same section.

Previously, the NLP algorithm was applied to the "chief complaint" section of patients who presented to an emergency department (ED).[16] The "chief complaint" indicated only the main reason for the evaluation, which may be ruled out during the visit (eg, patient had a normal ABI, and PAD was ruled-out). In contrast, our study included subjects evaluated in the inpatient or the outpatient settings, but we excluded notes from ED visits. In our institution, these notes from ED visits include combined narrative notes from multiple providers, and in the present study, the ED notes were a common reason for false results during the interactive refinement of our system. Others have applied NLP algorithms to hospital dismissal summaries, which summarize the hospital course.[17-19] In contrast, we validated an NLP algorithm applied to each of the progress notes that occurred during the course of a hospitalization. In addition, we also validated our system applied to outpatient clinical notes. The note types with the best performance, which were used in the final system all referred to a medical encounter. Notes that did not describe a medical encounter (eg, report of a phone conversation with a patient) were a reason for false results and were excluded from the final NLP algorithm.

Reasons for false positives—best NLP algorithm. We analyzed reasons for false results when our NLP algorithm was applied to narrative clinical notes (Table IV). We demonstrated that a reason for false positives includes notes in which clinicians suspected PAD and ordered the ABI, however, subsequent ABI results were normal and ruled out PAD. Another reason for false positives was the natural language complexity and ambiguity as the NLP algorithm was unable to recognize the correct experiencer of a disease (Table IV).

Table III. Results of natural language processing (*NLP*) algorithm compared with billing code algorithms (simple model and full model) for ascertainment of peripheral arterial disease (PAD) status

	Training dataset			Testing dataset				
	NLP, % (95% CI)	Full model, % (95% CI)	Simple model, % (95% CI)	NLP, % (95% CI)	Full model, % (95% CI)	Simple model, % (95% CI)	*P* value NLP vs full model	*P* value NLP vs simple model
Sensitivity	91.4 (88.9-93.9)	93.9 (91.8-96.1)	86.0 (82.9-89.1)	91.2 (88.1-94.2)	97.0 (95.1-98.8)	89.6 (86.3-92.9)	<.001	.45
Specificity	94.1 (91.9-96.2)	68.4 (64.2-72.7)	79.8 (76.1-83.5)	92.5 (89.6-95.5)	64.2 (58.8-69.5)	75.9 (71.1-80.7)	<.001	<.001
PPV	94.2 (92.1-96.3)	75.8 (72.3-79.2)	81.7 (78.4-85.1)	92.9 (90.0-95.7)	74.3 (70.2-78.4)	79.9 (75.8-84.0)	<.001	<.001
NPV	91.3 (88.7-93.8)	91.5 (88.5-94.5)	84.5 (81.0-87.9)	90.7 (87.5-93.9)	95.2 (92.2-98.1)	87.3 (83.3-91.3)	.01	.10
Accuracy	92.7 (91.1-94.4)	81.5 (79.0-84.0)	83.0 (80.6-85.4)	91.8 (89.7-93.9)	81.1 (78.1-84.1)	83.0 (80.1-85.9)	<.001	<.001

CI, Confidence interval; *NPV,* negative predictive value; *PPV,* positive predictive value.

Table IV. Reasons for false positives and false negatives in the best natural language processing (NLP) system

Category	Example
False positives	
Suspected PAD	"…he does have palpable dorsalis pedis pulses of the *feet* bilaterally with a difficult to palpate right posterior tibial pulse. Noninvasive arterial studies of the *lower extremities* will be performed to evaluate the severity and extent of *peripheral arterial disease*."
Ambiguity and complexity of natural language	"…She lives with her husband who is status post renal transplant in 1999. He struggles with diabetes and *arteriosclerosis* in both *lower extremities*."
False negatives	
Absence of location and/or diagnostic keywords	"…she clearly has *peripheral arterial disease*, but the left is worse than the right…"
ABI results not reported in clinical notes because of recently developed acute health problem	"Diminished pedal pulses: We will obtain noninvasive vascular studies of the lower extremities (September 12, XXXX)". ABI testing completed and reported on September 18, XXXX as mildly abnormal but not mentioned in the subsequent clinical notes. On the same day, the patient had new symptoms, was diagnosed with new atrial fibrillation, and underwent comprehensive cardiovascular evaluation for assessment of this acute condition.
Typographic errors	Instead of the correct term "ABI" clinical note snippet contains a typographic error "ADI": "…patient also has an appointment with the cardiologist today for further evaluation of his abnormal *ADI*."

ABI, Ankle-brachial index; *PAD,* peripheral arterial disease.

Reasons for false negatives. The absence of location and/or diagnostic keywords within ± two sentences window was a frequent reason for false negatives. Another reason for false negatives was the absence of comments in clinical notes regarding a recently conducted ABI test that showed abnormal results. However, this happened most often in cases when the patient developed acute health problem on the same day as the ABI report. The other reasons for false negatives were typographic errors.

Strengths and limitations. The present study has important strengths. First, the Mayo Clinic data warehouse archives comprehensive narrative clinical notes from both inpatient and outpatient encounters. Second, we had available to us the Mayo vascular laboratory dataset that archives all results/reports of noninvasive lower extremity arterial testing performed in the Mayo accredited vascular laboratory. Third, the NLP algorithm is independent of billing codes. The NLP algorithm uses keywords (Table I) and rules that are independent of EHR systems; hence, the system can be implemented for any other EHR systems. Fourth, a collaborative effort of a multidisciplinary team of investigators including clinicians, computer scientists, and biostatisticians was fundamental for the development of the NLP algorithm described herein. A limitation of this study is that data were retrieved from the data warehouse of a single academic medical center.

In future studies, we will apply and validate this NLP algorithm to identify PAD cases in other healthcare systems. Subsequently, we will deploy the refined NLP algorithm to Mayo Clinic EHR for automated identification of PAD cases at the point-of-care, with linkage to clinical decision support that will include reminders for risk modification strategies for patients with PAD as follows: antiplatelet therapy, statins therapy, antihypertensive therapy as well as smoking cessation.

CONCLUSIONS

In this study, we described a knowledge-driven NLP algorithm that ascertains PAD cases from clinical notes with higher accuracy compared with billing code algorithms; this system will support big data clinical studies with potential for translation to patient care. The presence of such a system could enhance capabilities to conduct PAD research on a large scale with potential favorable impact on public health and eventually improve care by clinical decision support.

We thank Kent Bailey, PhD, Bradley R. Lewis, MS, and Carin Smith for statistical analysis, Jared Robb for data collection, and Tamie Tiedemann for secretarial support.

AUTHOR CONTRIBUTIONS

Conception and design: NA, SS, SA, CS, HL, IK, AA-O
Analysis and interpretation: NA, SS, CS, HL, IK, AA-O
Data collection: NA, SS, SA, CS, AA-O
Writing the article: NA, SS, SA, CS, RC, HL, IK, AA-O
Critical revision of the article: SS, RC, HL, IK, AA-O
Final approval of the article: NA, SS, SA, CS, RC, HL, IK, AA-O
Statistical analysis: NA, CS, HL, IK, AA-O
Obtained funding: IK, AA-O
Overall responsibility: AA-O

REFERENCES

1. Criqui MH, Denenberg JO, Langer RD, Fronek A. The epidemiology of peripheral arterial disease: importance of identifying the population at risk. Vasc Med 1997;2:221-6.
2. Hirsch AT, Criqui MH, Treat-Jacobson D, Regensteiner JG, Creager MA, Olin JW, et al. Peripheral arterial disease detection, awareness, and treatment in primary care. JAMA 2001;286:1317-24.
3. Hirsch AT, Haskal ZJ, Hertzer NR, Bakal CW, Creager MA, Halperin JL, et al. ACC/AHA 2005 Guidelines for the Management of Patients With Peripheral Arterial Disease (Lower Extremity, Renal, Mesenteric, and Abdominal Aortic): A Collaborative Report from the American Association for Vascular Surgery/Society for Vascular Surgery,* Society for Cardiovascular Angiography and Interventions, Society for Vascular Medicine and Biology, Society of Interventional Radiology, and the ACC/AHA Task Force on Practice Guidelines (Writing Committee to Develop Guidelines for the Management of Patients With Peripheral Arterial Disease). J Am Coll Cardiol 2006;47:e1-192.
4. Kullo IJ, Rooke TW. Peripheral artery disease. N Engl J Med 2016;374:861-71.
5. Murabito JM, D'Agostino RB, Silbershatz H, Wilson PW. Intermittent claudication a risk profile from the Framingham Heart Study. Circulation 1997;96:44-9.
6. Saw J, Bhatt DL, Moliterno DJ, Brener SJ, Steinhubl SR, Lincoff AM, et al. The influence of peripheral arterial disease on outcomes: a pooled analysis of mortality in eight large randomized percutaneous coronary intervention trials. J Am Coll Cardiol 2006;48:1567-72.
7. Olin JW, Allie DE, Belkin M, Bonow RO, Casey DE, Creager MA, et al. ACCF/AHA/ACR/SCAI/SIR/SVM/SVN/SVS 2010 Performance Measures for Adults With Peripheral Artery Disease: A Report of the American College of Cardiology Foundation/American Heart Association Task Force on Performance Measures, the American College of Radiology, the Society for Cardiac Angiography and Interventions, the Society for Interventional Radiology, the Society for Vascular Medicine, the Society for Vascular Nursing, and the Society for Vascular Surgery (Writing Committee to Develop Clinical Performance Measures for Peripheral Artery Disease) Developed in Collaboration With the American Association of Cardiovascular and Pulmonary Rehabilitation; the American Diabetes Association; the Society for Atherosclerosis Imaging and Prevention; the Society for Cardiovascular Magnetic Resonance; the Society of Cardiovascular Computed Tomography; and the PAD Coalition Endorsed by the American Academy of Podiatric Practice Management. J Am Coll Cardiol 2010;56:2147-81.
8. Hirsch AT, Allison MA, Gomes AS, Corriere MA, Duval S, Ershow AG, et al. A call to action: women and peripheral artery disease a scientific statement from the American Heart Association. Circulation 2012;125:1449-72.
9. Hirsch AT. Treatment of peripheral arterial disease-extending "intervention" to "therapeutic choice". N Engl J Med 2006;354:1944-7.
10. Fan J, Arruda-Olson AM, Leibson CL, Smith C, Liu G, Bailey KR, et al. Billing code algorithms to identify cases of peripheral artery disease from administrative data. J Am Med Inform Assoc 2013;20:e349-54.
11. Savova GK, Fan J, Ye Z, Murphy SP, Zheng J, Chute CG, et al. Discovering peripheral arterial disease cases from radiology notes using natural language processing. AMIA Annu Symp Proc 2010;2010:722-6.
12. St Sauver JL, Grossardt BR, Yawn BP, Melton LJ, Pankratz JJ, Brue SM, et al. Data resource profile: the Rochester Epidemiology Project (REP) medical records-linkage system. Int J Epidemiol 2012;41:1614-24.
13. Liu H, Bielinski SJ, Sohn S, Murphy S, Wagholikar KB, Jonnalagadda SR, et al. An information extraction framework for cohort identification using electronic health records. AMIA Jt Summits Transl Sci Proc 2013;2013:149-53.
14. Birman-Deych E, Waterman AD, Yan Y, Nilasena DS, Radford MJ, Gage BF. Accuracy of ICD-9-CM codes for identifying cardiovascular and stroke risk factors. Med Care 2005;43:480-5.
15. Schmiedeskamp M, Harpe S, Polk R, Oinonen M, Pakyz A. Use of International Classification of Diseases, Ninth Revision Clinical Modification Codes and medication use data to identify nosocomial *Clostridium difficile* infection. Infect Control Hosp Epidemiol 2009;30:1070-6.
16. Chapman WW, Christensen LM, Wagner MM, Haug PJ, Ivanov O, Dowling JN, et al. Classifying free-text triage chief complaints into syndromic categories with natural language processing. Artif Intell Med 2005;33:31-40.

17. 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 1999:256-60.
18. 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;7:593-604.
19. Hripcsak G, Austin JH, Alderson PO, Friedman C. Use of natural language processing to translate clinical information from a database of 889,921 chest radiographic reports. Radiology 2002;224:157-63.

Submitted Aug 10, 2016; accepted Nov 14, 2016.

Appendix I. Excluded: Note types, note sections, and service groups

Note types	Note sections	Service groups
Miscellaneous	Chief complaint	Orthopedic
Test-oriented miscellaneous	History of present Illness	Podiatry
Dismissal summary	Family history	Endocrinology
Therapy	System reviews	Emergency medicine
Emergency medicine Hospital admission note Visit	Anticipated problems and interventions	Allergy
Hospital admission note	Informed consent	Dermatology
Emergency medicine visit	Patient education	Sports medicine
	Physical examination	Spine center
		Work rehabilitation
		Plastic surgery
		Nursing home
		Social services
		Addiction

Appendix II. Included: Note types, note sections, and service groups

Note types	Note sections	Service groups
Consult	Impression/report/plan	Primary care
Subsequent visit	Diagnosis	Hospital internal medicine
Patient progress	Principal/primary diagnosis	General medicine
Supervisory	Secondary diagnoses	Family medicine
Limited examination	Past medical/surgical history	Critical care
Specialty evaluation	Ongoing care	Urgent care
Multisystem evaluation	Immunizations	Cardiology
Injection	Key findings/test results	Vascular
Educational visit	Preprocedure information	Pulmonary
Hospital service transfer	Postprocedure information	Oncology
	Vital signs	Nephrology
	Current medications	Neurology
	Revision history	Pathology
	Special instructions	Gastroenterology
	Advance directives	Vascular wound care
	Discharge activity	Vascular surgery
	Final pathology diagnosis	Cardiac surgery