HHS Public Access
Author manuscript
*Med Care*. Author manuscript; available in PMC 2019 September 01.

Published in final edited form as:
*Med Care*. 2018 September ; 56(9): e54–e60. doi:10.1097/MLR.0000000000000831.

# Improved Identification of Venous Thromboembolism From Electronic Medical Records Using a Novel Information Extraction Software Platform

**Raymund B. Dantes, MD, MPH**[*], **Shuai Zheng, PhD**[†], **James J. Lu, PhD**[‡], **Michele G. Beckman, MPH**[§], **Asha Krishnaswamy, MSc**[‖], **Lisa C. Richardson, MD**[¶], **Sheri Chernetsky-Tejedor, MD**[*,†], and **Fusheng Wang, PhD**[†]

[*]Departments of Medicine, Division of Hospital Medicine, Atlanta GA

[†]Biomedical Informatics, Emory University School of Medicine, Atlanta GA

[‡]Department of Mathematics and Computer Science, Emory University, Atlanta GA

[§]Division of Blood Disorders, National Center on Birth Defects and Developmental Disabilities, Centers for Disease Control and Prevention, Atlanta GA

[‖]Division of Heart Disease and Stroke Prevention, National Center for Chronic Disease Prevention and Health Promotion, Atlanta GA

[¶]Division of Cancer Prevention and Control, National Center for Chronic Disease Prevention and Health Promotion, Centers for Disease Control and Prevention, Atlanta GA

## Abstract

**Introduction—**The United States federally mandated reporting of venous thromboembolism (VTE), defined by Agency for Healthcare Research & Quality Patient Safety Indicator 12 (AHRQ PSI-12), is based on administrative data, the accuracy of which has not been consistently demonstrated. We used IDEAL-X, a novel information extraction software system, to identify VTE from electronic medical records and evaluated its accuracy.

**Methods—**Medical records for 13,248 patients admitted to an orthopedic specialty hospital from 2009 to 2014 were reviewed. Patient encounters were defined as a hospital admission where both surgery (of the spine, hip, or knee) and a radiology diagnostic study that could detect VTE was performed. Radiology reports were both manually reviewed by a physician and analyzed by IDEAL-X.

**Results—**Among 2083 radiology reports, IDEAL-X correctly identified 176/181 VTE events, achieving a sensitivity of 97.2% [95% confidence interval (CI), 93.7%–99.1%] and specificity of

Reprints: Raymund B. Dantes, MD, MPH, Division of Hospital Medicine, Emory University Hospital, 1364 Clifton Road, NE, P.O. Box M-7, Atlanta, GA 30322. raymund.dantes@emoryhealthcare.org.

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the CDC. The interpretation and reporting of these data are the responsibility of the authors and in no way should be seen as an official policy of the US government.

An abstract of this study was Presented at the American Medical Informatics Association 2015 Annual Symposium, San Francisco, CA, November 14–18, 2015.

The authors declare no conflict of interest.

99.3% (95% CI, 98.9%–99.7%) when compared with manual review. Among 422 surgical encounters with diagnostic radiographic studies for VTE, IDEAL-X correctly identified 41 of 42 VTE events, achieving a sensitivity of 97.6% (95% CI, 87.4%–99.6%) and specificity of 99.8% (95% CI, 98.7%–100.0%). The performance surpassed that of AHRQ PSI-12, which had a sensitivity of 92.9% (95% CI, 80.5%–98.4%) and specificity of 92.9% (95% CI, 89.8%–95.3%), though only the difference in specificity was statistically significant ($P < 0.01$).

**Conclusion—**IDEAL-X, a novel information extraction software system, identified VTE from radiology reports with high accuracy, with specificity surpassing AHRQ PSI-12. IDEAL-X could potentially improve detection and surveillance of many medical conditions from free text of electronic medical records.

### Keywords

venous thromboembolism; natural language processing; machine learning; quality improvement

---

Venous thromboembolism (VTE), a blood clot that forms within a venous blood vessel, may form spontaneously or as a result of surgery, trauma, immobilization, pregnancy, cancer, and many other conditions. Deep vein thrombosis (DVT) is a VTE that forms in the large veins of the lower leg, thigh, and arms. A DVT may break off and travel through the venous system into the lungs, causing a pulmonary embolism (PE), a blood clot disrupting blood flow through the lungs, which can be fatal. In the United States, there are an estimated 350,000–900,000 VTEs per year, resulting in ~100,000 deaths per year.[1]

Among persons receiving major orthopedic surgery [including total knee replacement (TKR), total hip replacement (THR), or hip fracture surgery (HFS)], approximately 40%–60% may develop a distal DVT without prophylactic therapy.[2] Modern VTE prophylaxis, which consists of pharmacologic therapy or mechanical compression devices, can reduce the incidence of symptomatic VTE to approximately 0.5%–1% during the hospitalization.[3] Because postoperative VTE is common, potentially morbid, and preventable, tracking and preventing VTE has become a common metric for measuring the quality of health care delivery.

Perioperative PE or DVT have been designated by the US Government Agency for Healthcare Research and Quality (AHRQ) as a component of Patient Safety Indicator 12 (PSI-12). Starting in fiscal year 2015, the Centers for Medicare & Medicaid Services began reducing hospital payments to hospitals that perform poorly in the Hospital-acquired Conditions and Value-based Purchasing programs. PSI-12 (as part of the PSI-90 composite measure) is a component of both pay-for-performance systems. All PSIs are quantified based on administrative data [International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9 CM), Diagnosis Related Group (DRG), and/or Medicare Severity DRG (MS-DRG)] and the software is readily available for healthcare organizations. About half of events identified by an older version of PSI-12 that did not exclude preexisting VTE were found to be true postoperative DVT or PE events.[4] Studies of a more recent version of PSI-12 that excluded preexisting VTE showed improved performance for detecting postoperative VTE, but results were notably higher in academic hospitals (99%

positive predictive value), compared with a convenient sample of unidentified hospitals (81% positive predictive value).[5]

While automated information extraction (IE) from electronic medical records has high potential for improving VTE identification, existing tools either require large, manually annotated datasets for training, or considerable linguistic expertise for constructing domain specific rules. We present IDEAL-X,[6] a novel IE system designed for lay users without linguistics or programming background, and in collaboration with the Centers for Disease Control and Prevention evaluate its accuracy for identifying VTE diagnosis directly from the free text of radiology reports in electronic medical records.

## METHODS

### Human Subjects Consideration

This public health surveillance activity was reviewed by human subject personnel at the Emory University Institutional Review Board and was determined to be nonhuman subjects research and therefore did not require patient informed consent to review medical records.

### Data Source

Full text of radiology reports and clinical data were retrospectively extracted from the electronic medical records (Cerner Corp, Kansas City, MO) of patients admitted to Emory University Orthopedic and Spine Hospital from February 1, 2009 to December 9, 2014. This timeframe excludes patients and data collected after the implementation of ICD-10 coding at this institution. Cases were defined as patient encounters that included a hospital admission, surgery (of the spine, hip, or knee), and any radiology diagnostic study that could detect VTE. A physician (R.B.D.) manually reviewed each radiology report for diagnosis of a DVT or PE, which served as the gold standard comparison for this study. This review was blided to the results of the IDEAL-X determinations.

### Case Characteristics

Demographic and clinical characteristics of each case in the above timeframe were extracted and analyzed. These included patient age, sex, length of hospitalization, and evidence of pharmacologic and/or mechanical VTE prophylaxis during the hospital stay. Differences in characteristics among cases in each category were analyzed using the $\chi^2$ test or Fisher exact test (where appropriate) for categorical variables and the Wilcoxon rank sum test for continuous variables.

### IDEAL-X

IDEAL-X[6] is a novel software product with an easy-to-use interface that allows users without linguistics or programming background to extract desired information from unstructured text. Software demonstrations are available at www.idealx.net.

In this study, IDEAL-X was used to analyze the same radiology report using 2 different modes.

For controlled vocabulary mode, the user specified the terminology (eg, thrombosis, clot) and contextual information (such as relevant and irrelevant report sections) to be extracted before running the software. In this study, all radiology reports from year 2013 were then used as a development dataset to calibrate the software to the appropriate terminology and contextual information.

In online machine learning mode, all terminology and contextual information was learned incrementally by the software. The software interface and workflow allow a lay user to browse an input text document and identify the variables of interest. From a small (typically 4–5) set of documents, manual extraction provides the initial training to the data extraction engine of IDEAL-X. For each subsequent document, the user can review and revise incorrect answers automatically filled by the data extraction engine.[6] The user feedback updates and further improves the data extraction model. Once the user determines that the data extraction model has achieved sufficient precision, the remaining text documents can be automatically annotated in batch.

Since this study was considered exploratory, no test positivity cutoffs or sample size was predetermined.

### Radiology Report-level Analysis

IDEAL-X first analyzed every eligible radiology report in our healthcare system performed on patients admitted to Emory University Orthopedic and Spine Hospital from February 1, 2009 to December 9, 2014. IDEAL-X performance was measured in both controlled vocabulary and online machine learning modes, and benchmarked against manual physician review, in order to assess the accuracy of IDEAL-X on a task-by-task basis.

### Patient Encounter-level Analysis and AHRQ PSI-12 Comparison

To demonstrate patient encounter-level accuracy (as multiple radiology studies could be obtained during and admission) and to facilitate a direct comparison with AHRQ PSI-12, IDEAL-X analyzed radiology reports only completed during the patient encounter related to orthopedic surgery. An AHRQ PSI-12 event was defined using the 2014 definition (version 4.5a),[7] and based on ICD-9 codes (and present on admission indicators) previously assigned to the patient encounter by coders employed by the hospital. Of note, to satisfy the exclusion criteria of PSI-12 for this part of the analysis: (i) all patients with prior VTE by ICD-9 code were excluded; (ii) all patients at this hospital were assumed not to be pregnant or in puerperium, as those patients would have been transferred to a tertiary care hospital with obstetric support; and (iii) all patients were assumed not to have had an inferior vena cava filter placed during the hospitalization before orthopedic surgery as interventional radiology services were not available at this hospital. Both IDEAL-X and PSI-12 VTE determinations were benchmarked against manual physician review.

## RESULTS

Medical records for 13,248 patients admitted to Emory University Orthopedic and Spine Hospital from February 1, 2009 to December 9, 2014 were reviewed (Fig. 1). Among these patients, 677 diagnostic studies for VTE were performed across 546 patient admissions.

Across the entire timeframe, including other inpatient and outpatient encounters within our health care system, there were a total of 2551 diagnostic studies that could detect VTE among these 13,248 patients.

## Case Characteristics

The majority of cases under investigation for VTE underwent spine surgery [299 (55%), Table 1], with nearly equal proportions of THR [111 (20%)] and TKR [101 (18%)]; relatively few patients underwent HFS alone [35 (6%)]. One patient underwent both HFS and a THR, and this patient is categorized as a THR patient for subsequent analysis.

Most patients were female [359 (66%)], and the median age among all surgery types was either 65 or 66, except hip fracture patients, who were significantly older (median age: 78 y, $P > 0.001$). The median hospitalization duration was 6 days [interquartile range (IQR), 4–10 d]. THR and TKR patients had significantly shorter median hospitalization duration [5 d (IQR, 4–7 d) for each] compared with spine and HFS patients (median: 7 (IQR, 5–12) d and 7 (IQR, 6–12) d, respectively].

All patients had documented mechanical prophylaxis within 24 hours of completing surgery. Additionally, all THR, HFS, and TKR patients received pharmacologic VTE prophylaxis postoperatively, whereas only 23% of spine surgery patients received pharmacologic VTE prophylaxis. Low molecular weight heparin was the most frequently used pharmacologic agent, used by 77%–83% of patients undergoing THR, HFS, or TKR.

Among cases under investigation for VTE, 64% underwent ultrasound of a lower extremity, 4% underwent ultrasound of an upper extremity, and 50% underwent computerized tomographic scan of the chest with intravenous contrast, and a single patient underwent magnetic resonance imaging of the chest. No other radiology tests diagnostic of VTE were performed in this population.

Within this institution, documentation of a symptom is required to order any radiology test. Several common symptom choices are given, in addition to a free text response. Among patients undergoing a diagnostic study for DVT, symptoms of “edema” and “pain in limb” were indicated among 50% of radiologic studies. Among patients undergoing a diagnostic study for PE, “shortness of breath” was indicated in 55% of radiologic studies, followed by “chest pain” among 13% of radiologic studies.

Among all surgical patients with a radiologic study for VTE ordered during their hospitalization, 49 (9%) were diagnosed with a VTE (Table 1). The proportion of spine surgery patients with VTE was significantly higher than other surgery types combined (12% vs. 6%, $P = 0.016$). This difference was the result of a higher rate of PE among spine surgery patients compared with other surgery types combined (8% vs. 3%, $P = 0.014$).

## Radiology Report-level Analysis

**Controlled Vocabulary Mode—**IDEAL-X was calibrated in controlled vocabulary mode using the development dataset of 468 radiology reports in 2013 (the most recent complete calendar year of data). Among the remaining 2083 radiology reports in the testing dataset

(2009–2012, 2014), a DVT was detected by manual (physician) review in 112 of 1153 (9.7%) upper or lower extremity ultrasonography studies, and a PE was detected in 69 of 930 (7.4%) of computed tomography or magnetic resonance imaging of the chest.

When looking for the presence of any VTE (DVT and PE), IDEAL-X in controlled vocabulary mode correctly identified a VTE in 176 of 181 radiology reports, corresponding to a sensitivity of 97.2% [95% confidence interval (CI), 93.7%–99.1%, Table 2]. IDEAL-X incorrectly categorized 12 reports as describing a VTE when none was present (Appendix), corresponding with a specificity of 99.3% (95% CI, 98.9%–99.7%). Performance of IDEAL-X in controlled vocabulary mode was similar when looking for either DVT or PE (Table 2).

**Online Learning Mode—**IDEAL-X performance in controlled vocabulary mode was superior to online machine learning mode, which achieved a lower overall sensitivity of 92% (95% CI, 88.3%–96.1%) and similar 99% specificity (95% CI, 98.5%–99.4%), and required ~50% of reports to be processed before achieving > 95% sensitivity and specificity (Fig. 2).

### Patient Encounter-level Analysis and PSI-12 Comparison

Performance of IDEAL-X was also tested using only radiology reports completed during the patient encounter containing the orthopedic surgery, allowing a direct comparison to AHRQ PSI-12 coding from the same encounter. IDEAL-X in controlled vocabulary mode correctly identified 41 of 42 VTE events, achieving a sensitivity of 97.6% (95% CI, 87.4%–99.6%) and specificity of 99.8% (95% CI, 98.7%– 100.0%) (Table 3). The performance surpasses that of AHRQ PSI-12, which had an overall sensitivity of 92.9% (95% CI, 80.5%–98.4%) and specificity of 92.9% (95% CI, 89.8%–95.3%), though only the difference in specificity was statistically significant ($P < 0.01$).

## DISCUSSION

IDEAL-X, a novel IE software system, was capable of correctly identifying VTE from the free text of radiology reports with very high sensitivity and specificity, surpassing the overall performance of VTE identification based on AHRQ PSI-12 among a large population of orthopedic surgery patients. Clinical quality metrics sourced directly from clinical records may have increased validity compared to those from administrative data sources, but have typically been too labor-intensive to routinely obtain. This study highlited several performance characteristics of IDEAL-X that may make it a useful tool for surveillance of VTE. We found that creating a customized controlled vocabulary simplified the deployment process, and was better suited for detecting relatively rare events, including VTE (overall incidence <1% at this hospital). Online machine learning enables a user to provide IDEAL-X with real-time feedback to continually improve the accuracy of extraction of subsequent reports. Though initial calibration of IDEAL-X may be no faster than manual annotation, a user’s role may evolve from annotator to reviewer as IDEAL-X becomes more intelligent and accurate. Once sufficient accuracy is achieved, the system can extract remaining data automatically in batches, achieving considerable time-savings. IDEAL-X’s convenient workflow requires no linguistic expertise from the user, and can be easily adapted to different clinical applications to improve detection and surveillance of medical conditions.

Other studies have evaluated the use of automated IE or natural language processing (NLP) to detect VTE but have not achieved a similarly high sensitivity. Murff et al[8] detected VTE with a sensitivity of 59% and specificity of 91% using an NLP approach that applied the Systematized Nomenclature of Medicine-Clinical Terms (SNOMED) ontonology to broadly search narrative reports (including radiology reports) and other clinical and administrative medical records in the Veterans Administration. Our study was likely able to achieve a higher sensitivity due to the ability to customize and train IDEAL-X directly from the dataset in use. Rochefort et al[9] previously achieved a ~80% sensitivity and ~99% specificity for detecting DVT and PE using a “bag-of-words” model, which is derived from approaches typically used in information retrieval, to broadly screen for key words and some modifiers. Our study was able to achieve higher sensitivity for detecting VTE using an IE based approach, possibly because of the increased ability of IDEAL-X to both (1) identify the likely locations of relevant sentences and keywords, thus disregarding irrelevant contextual information; and (2) better account for context and grammar, thus better simulating the way a human would read text.

The AHRQ Patient Safety Indicators, introduced in 2003, were initially introduced to facilitate internal quality improvement. Beginning in fiscal year 2015, Centers for Medicare & Medicaid Services began including PSI-12 (as part of the PSI-90 composite measure) in both the Hospital-acquired Conditions and the Value-based Purchasing programs, each program has the potential to reduce a hospital’s Medicare reimbursement by at least 1%. Disease surveillance that is based on administrative data relies on the accuracy of documentation by medical providers and/or the ability of medical coders to locate and correctly and uniformly interpret free text of medical records. While our study confirmed that administrative/billing coding could be reasonably accurate in identifying VTE, IDEAL-X could identify VTE with higher specificity. Furthermore, the machine-learning approach of IDEAL-X suggests results could improve further, as our analysis of the 12 false-positive classifications showed that 10 of these errors could have been avoided with additional vocabulary learned during the study. While there are many technical barriers to wide implementation of a NLP-based surveillance for VTE (eg, integration with electronic medical record systems, standardization), this study demonstrates that such a process could lead to more accurate disease surveillance, resulting in improved estimates of VTE burden in healthcare and easier and more efficient tracking and monitoring of VTE occurrence and VTE prevention activities. In addition, electronic surveillance could facilitate prevention efforts by providing improved quality measurements and potentially more equitable and effective pay-for-performance programs.

This study had several limitations. This study focused on an orthopedic surgery population with a historically high risk of VTE in a single orthopedic specialty hospital setting, therefore results will need to be generalized across other hospitals and among patients with other types of surgery. While IDEAL-X outperformed PSI-12 in identifying VTE at our institution, coding practices likely vary across hospital systems and, therefore, PSI-12 accuracy may also vary. In addition, the recent transition to ICD-10 coding has resulted in updated PSI-12 definitions, whose accuracy will need to be established. IDEAL-X also currently requires medical records to exist in an electronic text format and cannot interpret scanned images of reports or handwritten text without additional software tools. This study also analyzed radiology reports from a single institution, with a limited number of radiologists, and therefore variations in word choice and syntax among radiologists across institutions could not be assessed. Finally, our study used physician review of radiology records as the gold standard for detection of VTE, however this does not account for VTE events where no diagnostic study was ordered, and thus this method likely underestimates the total burden of VTE.

In conclusion, IDEAL-X identified VTE from free text of radiology reports with very high sensitivity and specificity, surpassing the specificity AHRQ PSI-12. Further research is needed to generalize these findings across different health care settings and diverse patient populations. IDEAL-X's convenient workflow requires no linguistic or programming expertise from the user, and has potential for adoption in different clinical applications to improve detection and surveillance of many medical conditions.

## Acknowledgments

Funding was provided to Emory University via a contract from the Centers for Disease Control and Prevention (CDC) as part of a CDC Innovation Fund award.

## APPENDIX

### TABLE A1

Review of 12 False-positive Radiology Reports Where IDEAL-X in Controlled Vocabulary Mode Identified a Venous Thromboembolism When None was Present on Manual Review

| Frequency | False-positive Reason | Example |
|---|---|---|
| 6* | Description of radiology study limitations | "respiratory motion limits evaluation for pulmonary embolism," "small emboli in the segmental and subsegmental arteries can be missed" |
| 3* | Description of superficial vein | "superficial venous thrombosis of the left basilic vein" |
| 1* | Recommendation for additional testing | "nuclear study can be performed as clinically indicated for evaluation of pulmonary emboli" |
| 1 | Contradictory/ambiguous descriptions | "findings can be seen with … sequela of chronic pulmonary embolism, although there is no direct evidence of chronic pulmonary embolism" |
| 1 | Radiology report syntax typo | "1. no pulmonary embolism" (no space between "1."+"no") |

*Ten false-positives could be averted in the future by adding these learned terminology to existing controlled vocabulary in IDEAL-X.

## References

1. Office of the Surgeon General (US), National Heart L, and Blood Institute (US). The Surgeon General's Call to Action to Prevent Deep Vein Thrombosis and Pulmonary Embolism. Rockville, MD: Office of the Surgeon General (US); 2008.
2. Geerts WH, Pineo GF, Heit JA, et al. Prevention of venous thromboembolism: the Seventh ACCP Conference on Antithrombotic and Thrombolytic Therapy. Chest. 2004; 126(suppl):338S–400S. [PubMed: 15383478]
3. Januel JM, Chen G, Ruffieux C, et al. Symptomatic in-hospital deep vein thrombosis and pulmonary embolism following hip and knee arthroplasty among patients receiving recommended prophylaxis: a systematic review. JAMA. 2012; 307:294–303. [PubMed: 22253396]

4. White RH, Sadeghi B, Tancredi DJ, et al. How valid is the ICD-9-CM based AHRQ patient safety indicator for postoperative venous thromboembolism? Med Care. 2009; 47:1237–1243. [PubMed: 19786907]
5. Sadeghi B, White RH, Maynard G, et al. Improved coding of postoperative deep vein thrombosis and pulmonary embolism in administrative data (AHRQ Patient Safety Indicator 12) after introduction of new ICD-9-CM diagnosis codes. Med Care. 2015; 53:e37–e40. [PubMed: 23552433]
6. Zheng S, Lu JJ, Ghasemzadeh N, et al. Effective information extraction framework for heterogeneous clinical reports using online machine learning and controlled vocabularies. JMIR Med Inform. 2017; 5:e12. [PubMed: 28487265]
7. Agency for Healthcare Research & Quality. Perioperative pulmonary embolism or deep vein thrombosis rate: technical specifications, Version 4.5a. 2014. Available at: www.qualityindicators.ahrq.gov. Accessed December 1, 2015
8. Murff HJ, FitzHenry F, Matheny ME, et al. Automated identification of postoperative complications within an electronic medical record using language processing. JAMA. 2011; 306:848–855. [PubMed: 21862746]
9. Rochefort CM, Verma AD, Eguale T, et al. A novel method of adverse event detection can accurately identify venous thromboembolisms (VTEs) from narrative electronic health record data. J Am Med Inform Assoc. 2015; 22:155–165. [PubMed: 25332356]

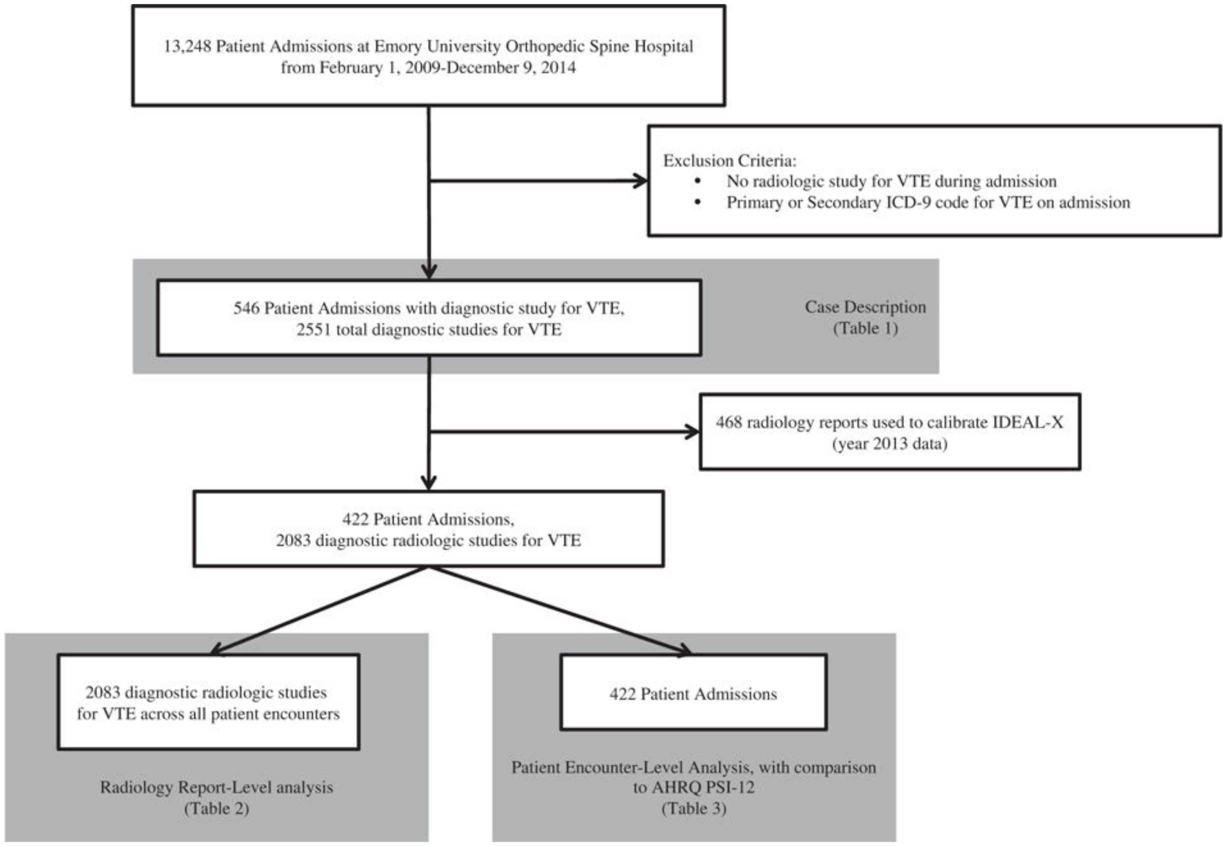


**FIGURE 1.**

Study population. AHRQ PSI-12 indicates Agency for Healthcare Research & Quality Patient Safety Indicator 12; ICD-9, International Classification of Diseases, Ninth Revision; VTE, venous thromboembolism.

Author Manuscript

Author Manuscript

Author Manuscript

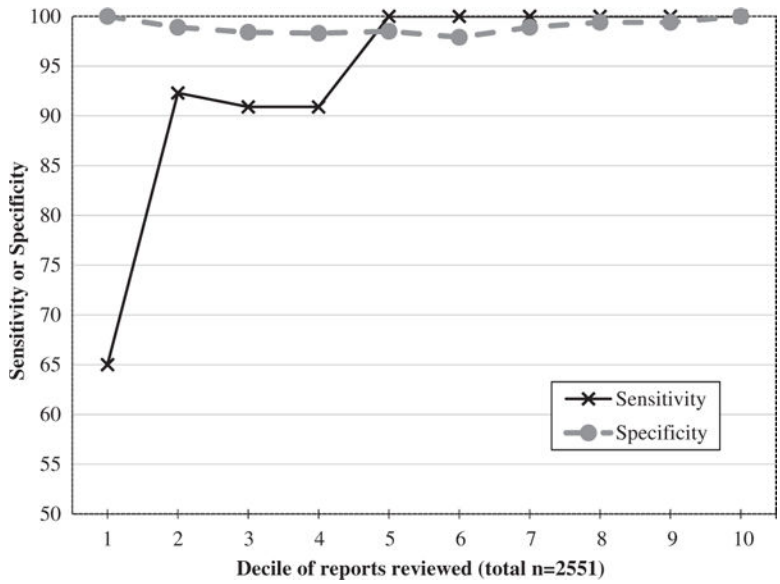


**FIGURE 2.**

Changes in sensitivity and specificity of IDEAL-X in online machine learning mode by decile of total diagnostic studies for venous thromboembolism (n = 2551).

# TABLE 1

Characteristics for Encounters for Hip, Knee, and Spinal Surgeries at Emory Orthopedic and Spine Hospital With Radiographic Study for VTE, February 1, 2009 to December 9, 2014

| | Surgery Type [n (%)] | | | | | |
|---|---|---|---|---|---|---|
| **Characteristic** | **Spine** | **Total Hip Replacement*** | **Hip Fracture Only*** | **Total Knee Replacement** | **Total [n (%)]** | ***P*** |
| Total patient encounters | 299 (55) | 111 (20) | 35 (6) | 101 (18) | 546 (100) | |
| Sex | | | | | | |
| Male | 101 (34) | 36 (32) | 13 (37) | 37 (37) | 187 (34) | 0.902 |
| Female | 198 (66) | 75 (68) | 22 (63) | 64 (63) | 359 (66) | |
| Age [median (interquartile range)] (y) | 65 (54–71) | 66 (56–73) | 78 (68–86) | 66 (62–71) | 66 (57–73) | < 0.001 |
| Race | | | | | | |
| White | 207 (69) | 67 (60) | 30 (86) | 55 (54) | 359 (66) | 0.002 |
| Black | 77 (26) | 39 (35) | 2 (6) | 37 (37) | 155 (28) | 0.001 |
| Other | 10 (3) | 1 (1) | 0 | 5 (5) | 16 (3) | 0.285 |
| Unknown | 5 (2) | 4 (4) | 3 (9) | 4 (4) | 16 (3) | 0.073 |
| Length of hospitalization [median (IQR)] (d) | 7 (5–12) | 5 (4–7) | 7 (6–12) | 5 (4–7) | 6 (4–10) | < 0.001 |
| VTE prophylaxis, pharmacologic or mechanical | 299 (100) | 111 (100) | 35 (100) | 101 (100) | 546 (100) | NA |
| Received any pharmacologic prophylaxis after surgery | 68 (23) | 111 (100) | 35 (100) | 100 (99) | 314 (58) | < 0.001 |
| Unfractionated heparin | 31 (10) | 19 (17) | 9 (26) | 7 (7) | 66 (12) | 0.007 |
| Low molecular weight heparin | 35 (12) | 90 (81) | 29 (83) | 78 (77) | 232 (42) | < 0.001 |
| Direct thrombin or factor Xa inhibitor | 2 (1) | 20 (18) | 2 (6) | 24 (24) | 48 (9) | < 0.001 |
| Warfarin | 23 (8) | 16 (14) | 3 (9) | 9 (9) | 51 (9) | 0.223 |
| Aspirin monotherapy | 16 (5) | 14 (13) | 4 (11) | 9 (9) | 43 (8) | 0.058 |
| Mechanical prophylaxis | 299 (100) | 111 (100) | 35 (100) | 101 (100) | 546 (100) | NA |
| Radiographic diagnostic testing (first test per category) | | | | | | |
| Ultrasound of lower extremity | 192 (64) | 76 (68) | 15 (43) | 66 (75) | 349 (64) | 0.050 |
| Ultrasound of upper extremity | 17 (6) | 2 (2) | 4 (11) | 1 (1) | 24 (4) | 0.020 |
| CT of chest with contrast | 158 (53) | 51 (46) | 20 (57) | 46 (46) | 275 (50) | 0.364 |
| MRI of chest | 1 (0) | 0 | 0 | 0 | 1 (0) | 0.843 |
| Symptoms indicated when radiology study ordered | | | | | | |
| DVT (ultrasound of upper or lower extremity) | | | | | | |

| Characteristic | Surgery Type [n (%)] Spine | Total Hip Replacement* | Hip Fracture Only* | Total Knee Replacement | Total [n (%)] | *P* |
|---|---|---|---|---|---|---|
| Any symptom | 209 (100) | 78 (100) | 19 (100) | 67 (100) | 373 (100) | N/A |
| Edema | 51 (24) | 26 (33) | 3 (16) | 27 (40) | 107 (29) | 0.049 |
| Pain in limb | 47 (22) | 14 (18) | 1 (5) | 18 (27) | 80 (21) | 0.235 |
| Other | 111 (53) | 38 (49) | 15 (79) | 22 (33) | 186 (50) | — |
| PE (CT PE protocol or MRI chest) | | | | | | |
| Any symptom | 159 (100) | 51 (100) | 20 (100) | 46 (100) | 276 (100) | NA |
| Shortness of breath | 85 (53) | 28 (55) | 9 (45) | 29 (63) | 151 (55) | 0.552 |
| Chest pain | 22 (14) | 9 (18) | 2 (10) | 3 (7) | 36 (13) | 0.401 |
| Other | 52 (33) | 14 (27) | 9 (45) | 14 (30) | 89 (32) | — |
| Postoperative VTE, by physician review of radiology reports | 35 (12) | 8 (7) | 2 (6) | 4 (4) | 49 (9) | 0.089** |
| DVT, lower extremity | 7 (2) | 3 (3) | 0 | 4 (4) | 14 (3) | 0.704 |
| DVT, upper extremity | 5 (2) | 0 | 0 | 0 | 5 (1) | 0.445 |
| PE | 23 (8) | 5 (5) | 2 (6) | 0 | 30 (5) | 0.009*** |

* One patient had both hip fracture repair and total hip replacement was categorized under total hip replacement.

CT indicates computerized tomography; DVT, deep vein thrombosis; IQR, interquartile range; MRI, magnetic resonance imaging; NA, not available; PE, pulmonary embolism; VTE, venous thromboembolism.

** $P = 0.016$ when comparing spine surgery to all others combined.

*** $P = 0.014$ when comparing spine surgery to all others combined.

## TABLE 2

Performance of IDEAL-X System in Controlled Vocabulary Mode, Analyzing Total Radiology Reports

| Event | Radiology Report Types | Total Reports | Positive Reports By Manual Review | Positive Reports By IDEAL-X | Measure | IDEAL-X Performance (95% CI) |
|---|---|---|---|---|---|---|
| DVT | Ultrasonography of upper or lower extremity | 1153 | 112 | 109 | Sensitivity<br>Specificity | 97.3% (92.4%–99.4%)<br>99.4% (98.7%–99.8%) |
| PE | CT and MRI of chest | 930 | 69 | 67 | Sensitivity<br>Specificity | 97.1% (89.9%–99.6%)<br>99.3% (98.5%–99.7%) |
| Either DVT or PE | All 4 types above | 2083 | 181 | 176 | Sensitivity<br>Specificity | 97.2% (93.7%–99.1%)<br>99.3% (98.9%–99.7%) |

CI indicates confidence interval; DVT, deep vein thrombosis; PE, pulmonary embolism.

## TABLE 3

Performance of IDEAL-X System in Controlled Vocabulary Mode, Compared With AHRQ PSI-12, Analyzed by Patient Surgical Encounter

| Event | Total Patients | Events by Manual Review | Events by IDEAL-X | Events by AHRQ PSI-12 | Measure | IDEAL-X (95% CI) | AHRQ PSI-12 (95% CI) | *P* |
|---|---|---|---|---|---|---|---|---|
| DVT | 422 | 17 | 16 | 13 | Sensitivity | 94.1% (71.2%–99.0%) | 76.5% (50.1%–93.0%) | 0.38 |
| | | | | | Specificity | 100.0% (99.1%–100.0%) | 96.1% (93.7%–97.7%) | < 0.01 |
| PE | 422 | 25 | 25 | 25 | Sensitivity | 100.0% (86.2%–100.0%) | 100.0% (86.16%–100.0%) | 1.00 |
| | | | | | Specificity | 99.8% (98.6%–100.0%) | 95.7% (93.2%–97.5%) | < 0.01 |
| Either DVT or PE | 422 | 42 | 41 | 39 | Sensitivity | 97.6% (87.4%–99.6%) | 92.9% (80.5%–98.4%) | 0.63 |
| | | | | | Specificity | 99.8% (98.7%–100.0%) | 92.9% (89.8%–95.3%) | < 0.01 |

AHRQ PSI-12 indicates Agency for Healthcare Research & Quality Patient Safety Indicator 12; CI, confidence interval; DVT, deep vein thrombosis; PE, pulmonary embolism.