

Predicting pulmonary embolism among hospitalized patients with machine learning algorithms

Logan Ryan | Jenish Maharjan | Samson Mataraso | Gina Barnes  |
Jana Hoffman | Qingqing Mao | Jacob Calvert | Ritankar Das

Dascena, Inc., Houston, Texas, USA

Correspondence

Gina Barnes, Dascena, Inc., 8515 Fannin
St. Suite 110 Houston, Texas 77054, USA.
Email: gbarnes@dascena.com

Abstract

Background: Pulmonary embolisms (PE) are life-threatening medical events, and early identification of patients experiencing a PE is essential to optimizing patient outcomes. Current tools for risk stratification of PE patients are limited and unable to predict PE events before their occurrence.

Objective: We developed a machine learning algorithm (MLA) designed to identify patients at risk of PE before the clinical detection of onset in an inpatient population.

Materials and Methods: Three machine learning (ML) models were developed on electronic health record data from 63,798 medical and surgical inpatients in a large US medical center. These models included logistic regression, neural network, and gradient boosted tree (XGBoost) models. All models used only routinely collected demographic, clinical, and laboratory information as inputs. All were evaluated for their ability to predict PE at the first time patient vital signs and lab measures required for the MLA to run were available. Performance was assessed with regard to the area under the receiver operating characteristic (AUROC), sensitivity, and specificity.

Results: The model trained using XGBoost demonstrated the strongest performance for predicting PEs. The XGBoost model obtained an AUROC of 0.85, a sensitivity of 81%, and a specificity of 70%. The neural network and logistic regression models obtained AUROCs of 0.74 and 0.67, sensitivity of 81% and 81%, and specificity of 44% and 35%, respectively.

Conclusions: This algorithm may improve patient outcomes through earlier recognition and prediction of PE, enabling earlier diagnosis and treatment of PE.

KEYWORDS

anticoagulants, artificial intelligence, clinical, decision support systems, thrombosis

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 Dascena, Inc. *Pulmonary Circulation* published by John Wiley & Sons Ltd on behalf of Pulmonary Vascular Research Institute.

INTRODUCTION

A pulmonary embolism (PE) is an obstruction of a blood vessel in the branching arteries of the lung.¹ The obstruction is usually caused by the embolization of a distal blood clot (or thrombus) originating in a deep vein. A PE is a life-threatening condition associated with high morbidity and mortality; mortality rates are estimated to be 30% in untreated PE and 8% in treated PE.² Recognition of PE is made difficult by a significant overlap between symptoms of PE and symptoms of other conditions, such as acute coronary syndrome, heart failure, pneumonia, and exacerbation of chronic obstructive pulmonary disease (COPD).³ Clinical signs and symptoms of PE are considered to be limited in terms of both sensitivity and specificity.¹ Accurate prediction of inpatient PE remains an unmet need.

This study describes the development of machine learning algorithms (MLAs) to predict PE in hospitalized patients. Predictions may then be used to either hasten identification of a PE which has occurred but has not yet been detected or to predict the future development of PE during a patient's hospital stay. Identification of patients likely to experience a PE would enable increased monitoring of high-risk patients and earlier diagnosis. In addition, high-risk patients with no contraindications could begin anticoagulants earlier in the disease course or prophylactically, potentially reducing the need for higher-risk procedures such as catheter-directed thrombolysis.⁴ This tool is designed to enable earlier PE diagnosis and more timely intervention, and provide clinicians with the opportunity to improve patient outcomes.

METHODS

Data processing

Data used for model development were extracted from the electronic health record (EHR) system at a large, tertiary medical center in the western United States. Data were extracted from medical and surgical patients admitted to the hospital between May 2011 and November 2017. This data set contained patient data including demographics, lab results, vital sign measurements, medication usage, and patient diagnoses.

Data were collected passively and were deidentified in compliance with the Health Insurance Portability and Accountability Act. Because this study was performed on deidentified data, and therefore, constitutes a nonhuman subject study as per the definition of human subjects

research put forth in 45 Code of Federal Regulations 46, it was exempt from Institutional Review Board approval.

Cohort definition and algorithm run time

We included all patients who had at least one recorded measure of all vital signs present in the chart (systolic and diastolic blood pressure, heart rate, respiratory rate, and temperature), at least one of the laboratory measurements used in the models (Table S1) present in the chart, and who were 40 years or older. The age criterion was included to minimize the probability of false alerts. The risk of PE increases significantly with age and is a less likely diagnosis in young adult patients.⁵

The algorithm was designed to use only the first 3 h of data after any vital or lab measurement was recorded during a patient's hospital stay. Any data collected after the 3-h mark were excluded. All patients meeting the gold standard definition of PE before the algorithm was able to generate predictions (i.e., before all required measurements were present in the chart) were excluded, therefore, no patients with a known prevalent PE were included in the study cohort.

The gold standard definition of PE was identified by the presence of an International Classification of Disease, Ninth Revision, Clinical Modification (ICD-9) or ICD, Tenth Revision (ICD-10) code for PE in the patient chart, together with the presence of an order for a therapeutic regimen of anticoagulants in terms of type and dosage,⁶ order of a thrombolytic medication, or insertion of an inferior vena cava (IVC) filter. Although ICD-10 codes have high sensitivity (~90%) for detecting PE, false positives may be present.^{7,8} Therefore, the use of ICD codes and therapeutic treatment together as the gold standard enhanced our ability to capture true positive class PE patients. ICD codes and medication orders used to identify PE are presented in Table S2. We also included thrombolytics in our definition of the gold standard, as these medications may be urgently ordered for PE patients who are hemodynamically unstable.⁹ We note that we included a broader list of thrombolytics than are currently approved by the Food and Drug Administration or recommended for use in treating acute PE to account for variability in real-world clinical practice and to ensure that patients experiencing a PE were unlikely to be missed by our gold standard definition of PE. Treatment with therapeutic anticoagulants, thrombolytics, or IVC filter placement was also used to determine the onset time of the clinical detection of PE, as ICD codes are not reliably assigned a time-stamp and so could not be used to determine which patients to exclude due to diagnosis of PE before the time of a prediction by the algorithm. All

patients who met the PE gold standard during their hospitalization were considered to be positive for PE. All other patients were considered to be negative.

MLAs

MLAs were developed to predict the development of PE at any point during the patient stay. We compared the performance of models developed using logistic regression, neural network (using a multilayer perceptron method), and gradient boosted decision trees (implemented using XGBoost in Python).^{10,11} For consistency with regard to the inputs, the vitals and lab measurements from the 3-h input period were binned and averaged for every hour. The difference between the measurements of the first and second hours and second and third hours were calculated and provided as new inputs to the model to provide information about the change in measurements over time. Since logistic regression and neural network models are unable to incorporate missing data or not a number (*NaN*) values, such values were replaced with the average value for that feature across the entire data set. The algorithm generated PE risk predictions the first time all patient vitals and at least one laboratory measure was present in the patient chart. The features included in all models are presented in Table S1.

Before model training, the development data set was randomly split in a ratio of 80:20, with these partitions forming the training and test sets, respectively. The XGBoost model was trained with 100 estimators and a maximum depth of four nodes. The learning rate, gamma, colsample by tree, and L2 regularization (lambda) values were set to 0.08, 0.2, 0.6, and 3, respectively. The value for the hyperparameter scale_pos_weight was set to 12.8 which accounted for the high-class imbalance in the data set. All of the hyperparameters were selected for the XGBoost by performing a cross-validated grid search. A neural network model was trained with 1 hidden layer of 100 neurons for a maximum of 300 iterations with ReLU activation function to introduce nonlinearity. A learning rate of 0.001 and tolerance for optimization of 0.001 was used for training the neural network. Model performance on the hold-out test set was assessed with regard to area under the receiver operating characteristic curve (AUROC), sensitivity, specificity, positive and negative likelihood ratios, and diagnostic odds ratio. A Shapley additive explanations (SHAP) analysis¹² was performed to evaluate the feature importance used by the best performing model in generating predictions.

RESULTS

In total, 60,297 patients were included in the experiments to develop and test the three algorithms, 309 of whom experienced a PE while hospitalized. On average, patients experiencing a PE were likely to be older, have a history of cancer, have experienced past venous thromboembolism (VTE), or be diagnosed with pneumonia. Patient demographic information for the full data set is presented in Table 1. Demographic information for the hold-out test set only is presented in Table S3.

Of the three machine learning (ML) models examined, XGBoost demonstrated the highest performance in terms of AUROC (Figure 1). XGBoost achieved an AUROC of 0.85, while the neural network and logistic regression models achieved AUROC values of 0.74 and 0.67, respectively. At a constant sensitivity of 81% across models, the XGBoost model also obtained superior specificity, positive and negative likelihood ratios, and diagnostic odds ratio values (Table 2).

The feature importance plot for XGBoost was generated using the TreeSHAP algorithm (Figure 2). The SHAP summary plot ranks the most important input features based on the contribution to the decision-making process of the algorithm. Record of recent fracture, history of surgery, and history of deep vein thrombosis were the top three most important features for generating accurate predictions of PE at any point during the patient's stay.

To study the effects of anticoagulant usage before the algorithm runtime, we also trained an XGBoost model with an additional boolean input indicating whether a patient was administered anticoagulants during the hospital stay before the algorithm run time. However, the addition of this extra input did not significantly improve the model's performance.

DISCUSSION

In this study, we demonstrated the ability of ML-based models to identify patients at high risk of experiencing a PE before the event occurred or was clinically detected. Because PE can rapidly become a life-threatening event, early detection or advanced prediction can optimize care by enabling rapid diagnosis and treatment or prophylaxis. The present study determined that our gradient boosted algorithm is capable of accurately predicting development of PE before the clinical detection of onset.

Although several ML models were evaluated for their ability to predict PE in hospitalized patients, the gradient boosted decision tree algorithm developed using

	Characteristic	Non PE encounters (%) <i>n</i> = 60,297	PE encounters (%) <i>n</i> = 309
Age	40–49	10542 (17.48)	37 (11.97)
	50–59	15090 (25.03)	73 (23.62)
	60–69	17540 (29.09)	108 (34.95)
	70–79	9994 (16.57)	63 (20.39)
	80+	7131 (11.83)	28 (9.06)
Sex	Male	30942 (51.32)	152 (49.19)
	Female	29355 (48.68)	157 (50.81)
Ethnicity	Hispanic	6933 (11.50)	22 (7.12)
	Not Hispanic	53364 (88.50)	287 (92.88)
Race	White	34678 (57.51)	193 (62.46)
	Black	5645 (9.36)	38 (12.30)
	Asian	8832 (14.65)	42 (13.59)
	Pacific Islander	843 (1.40)	4 (1.29)
	Native American	276 (0.46)	0 (0.00)
	Other/unknown	10023 (16.62)	32 (10.36)
Comorbidities	HIV/AIDS	1472 (2.44)	9 (2.91)
	Renal disease	5205 (8.63)	15 (4.85)
	Liver disease	6397 (10.61)	33 (10.68)
	Prior organ transplant	6405 (10.62)	19 (6.15)
	Diabetes	14059 (23.32)	68 (22.01)
	COPD	5037 (8.35)	42 (13.59)
	History of cancer	22354 (37.07)	168 (54.37)
	Alcohol use disorder	2601 (4.31)	13 (4.21)
	Pneumonia	5474 (9.08)	34 (11.00)
	Heart failure	7216 (11.97)	60 (19.42)
	Myocardial infarction	2161 (3.58)	8 (2.59)
	Previous VTE	3369 (5.59)	133 (43.04)
	Recent/active pregnancy (within 2 months)	7276 (12.07)	44 (14.24)
	Recent surgery (within 1 month)	1794 (2.98)	5 (1.64)

Abbreviations: COPD, chronic obstructive pulmonary disease; HIV/AIDS, human immunodeficiency virus/acquired immunodeficiency syndrome; PE, pulmonary embolism; VTE, venous thromboembolism.

TABLE 1 Demographic and clinical characteristics of the study sample used for training and testing the machine learning models

XGBoost performed most accurately, which may be attributed to XGBoost's superior handling of missing data. It is possible that other neural networks could be more customized to this prediction task with additional research to outperform the multilayer perceptron model used in the present study. The accuracy of predictions made by XGBoost makes the use of this model advantageous.

A SHAP analysis was used to evaluate the contributions of individual features to model predictions.¹² Several features identified as important for model predictions in the SHAP summary plot of the XGBoost model (Figure 2) have previously been identified as provoking factors for provoked PE, or PE precipitated by identifiable, major risk factors. Major trauma is considered a risk factor for provoked PE,⁵ and recent fracture

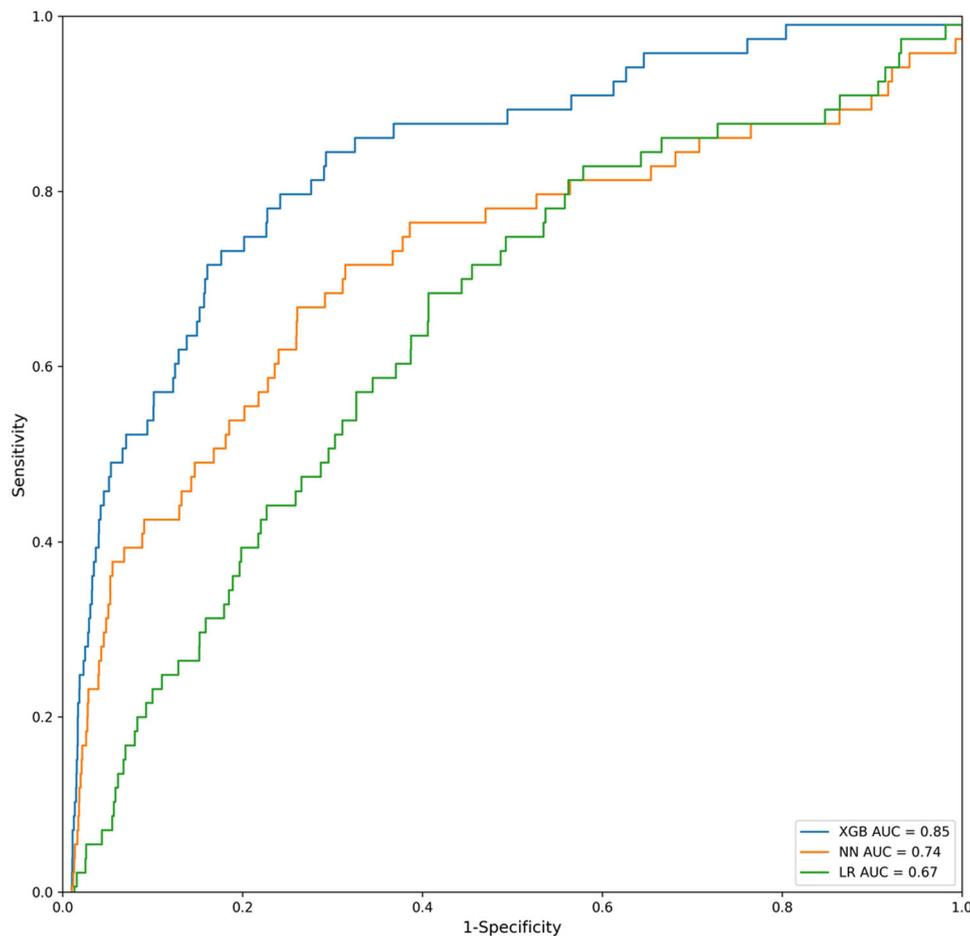


FIGURE 1 Receiver operating characteristic (ROC) curves for the XGBoost, neural network, and logistic regression machine learning models. AUROC, area under the receiver operating characteristic curve; LR, logistic regression; NN, neural network; XGB, XGBoost

TABLE 2 Comparison of performance metrics of the three machine learning models

Model	AUC	Sensitivity	Specificity	LR +	LR -	DOR
XGBoost	85%	81%	77%	3.48	0.25	13.80
Neural network	74%	81%	48%	1.56	0.40	3.89
Logistic regression	67%	81%	45%	1.47	0.43	3.43

Abbreviations: AUROC, area under the receiver operating characteristic curve; DOR, diagnostic odds ratio; LR, likelihood ratio.

may act as a proxy for recent trauma. The SHAP analysis also identified additional, nonprovoking risk factors which have been linked to increased risk of PE. Previous DVT, a known PE risk factor, was one of the five most important features for PE prediction.¹³ Further, patient fluid status impacts hemoconcentration, which has been linked to increased risk of thromboembolic events.¹⁴ Urine output, change in urine output and receipt of a fluid bolus were among the most important features, which, in combination, may reflect whether an individual is dehydrated and hemoconcentrated, or in a more balanced, net fluid positive situation in which

additional IV fluids are not required. Obesity has also been associated with increased risk for VTE and PE¹⁵; higher weight was associated with higher risk of a prediction of PE, suggesting that weight as a feature may have indirectly represented obesity.

The MLA developed in this study offers many advantages over alternative risk stratification methods. Unlike many existing rules-based risk stratification methods, this algorithm requires no additional clinician inputs or workflow disruption and automatically screens a broad inpatient population based only on data taken from the EHR. Several rules-based tools have been

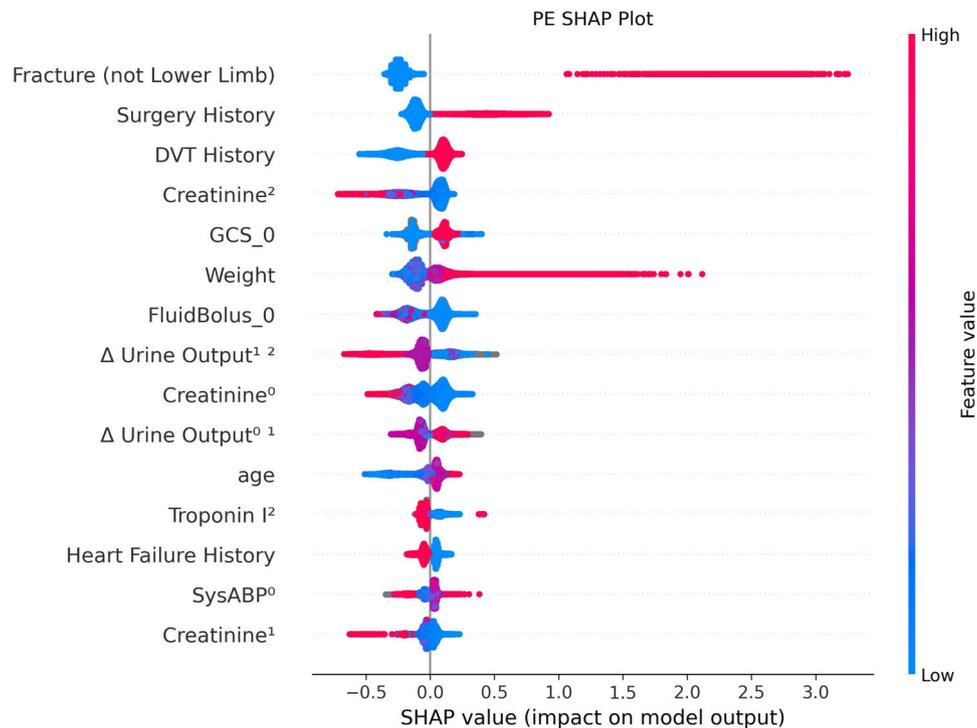


FIGURE 2 SHAP summary plot for the XGBoost model. The x axis of the plot shows the SHAP value for each of the features. The color of a point is indicative of the feature value, where red is a high value and blue is a low value. The y axis lists feature names in descending order of importance to the model's decision-making process. Superscripts in the feature names denote the hour of the patient's stay at which the measure was recorded. Delta symbols (Δ) are used when a feature captures the hourly change in a measure, with superscripts denoting the hours under consideration, for example, Δ Urine Output^{1,2} is the change in urine output measured during the first hour and the second hour in the 3-h input period. DVT, deep vein thrombosis; GCS, Glasgow coma scale; SysABP, systolic arterial blood pressure

developed to aid in the stratification of patients with suspected PE; the two most commonly used risk scores being the Wells criteria for PE¹⁶ and the revised Geneva score.¹⁷ However, these tools were designed for use in patients who are suspected to have an existing PE, and were not designed to predict the future occurrence of PE. Furthermore, these scoring systems have only been validated in assessing the risk of PE in outpatients and not for hospitalized patients.¹⁸

A performance accuracy meta-analysis of the Wells and revised Geneva score for diagnosing PEs reported that sensitivity ranged from 63% to 79% and 55% to 74%, respectively.¹⁹ The specificity for the Wells score ranged from 49% to 90% and had an AUROC of 0.78 while the Geneva score had an AUROC of 0.69. Furthermore, as the use of the Wells score has been validated in discharged patients, one study investigated its performance for hospitalized patients suspected of having deep vein thrombosis and deemed that it was inaccurate, and thus unsafe for inpatient use.²⁰ Additional risk assessment tools developed for VTE, as opposed to strictly PE, are also commonly used. These include the International Medical Prevention Registry on Venous Thromboembolism (IMPROVE) score, used for determining VTE in

hospitalized patients,²¹ the Padua Prediction Score (PPS), which assesses VTE risk in inpatient populations, and the Caprini Score, designed for use with surgical patients.²² However, each of these scorings has its limitations. For example, efforts to enhance the accuracy of VTE identification with the IMPROVE tool have been achieved by incorporating D-dimer lab values, which require laboratory tests that may not be readily available.^{23,24} The Padua score has not been extensively validated, which limits the generalizability of the tool²²; though the Caprini score has been validated, this has been achieved in specific subpopulations, including surgical patients, patients hospitalized with serious illness, and recently, those with COVID-19.^{25–27} Therefore, this tool may also lack generalizability. Our MLA has been developed and evaluated on a broad hospital inpatient population and demonstrates high accuracy in screening this population for PE risk.

ML methods have been explored for PE prediction in the past; however, prior work has focused on the ability of ML to assist with interpreting chest images^{28–30} or their accompanying radiology reports.^{31–34} Although these methods have the potential to increase the accuracy and timeliness of definitive PE diagnosis, MLAs for interpretation of radiological

imaging cannot be applied to patients who have not yet undergone a formal diagnostic workup for PE. The method described in this study adds to the existing literature, by demonstrating a means through which ML methods can improve PE prediction and early detection upstream of the confirmatory diagnostic process. Such prediction may improve patient outcomes by enabling not only earlier therapeutic treatment of PE but also prophylactic use of anticoagulants. Interestingly, a recent study by Nafee et al. developed MLAs to predict the risk of VTE in hospitalized patients with certain acute medical illnesses, and these outperformed the IMPROVE score.³⁵ This study obtained data from the APEX clinical trial.³⁶ However, certain lab measurements used for the APEX trial may not be routinely taken, and thus could limit the MLA's use.

Our study has several limitations. First, we evaluated the MLAs in a retrospective setting and we, therefore, cannot determine how any algorithm will perform in novel, prospective settings. Also related to the retrospective nature of this study, it is possible that not all patients' PE statuses were correctly classified by the gold standard. ICD codes have known limitations for identifying acute conditions, including PE, in chart data.⁷ However, ICD codes for PE have previously been shown to identify PE within hospitalization data with sensitivities and specificities ranging from 88% to over 90%.⁸ We attempted to mitigate this limitation by requiring that all identified PE cases additionally documented accepted treatments for PE (therapeutic anticoagulant regimens, thrombolytic medications, or placement of an IVC filter) during their hospital stay, to increase the diagnostic specificity. This retrospective study was conducted on data from a single hospital center. We cannot predict how this algorithm would perform in patient populations that have different demographics or a different incidence of PE, or those who reside in another geographic location. It is necessary to evaluate the algorithm on a range of additional sites to determine its generalizability. The incorporation of imaging data into the model may improve its performance and is an area for further investigation. Evaluation of the MLA in a prospective clinical setting is required to evaluate any effect on clinician actions and impact on patient outcomes.

In this study, an MLA capable of accurately predicting which patients will develop a PE during their hospitalization was developed. This gradient boosted algorithm utilized only routinely collected health data from inpatient EHR data to predict risk of inpatient development of PE. The algorithm described in this study may be able to improve patient outcomes through earlier identification of at-risk patients, allowing for earlier confirmatory diagnostic testing and treatment of PE.

ACKNOWLEDGMENTS

The authors would like to sincerely thank and acknowledge Anna Siefkas for her writing and revision of the manuscript. This study received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CONFLICT OF INTERESTS

All authors associated with Dascena are employees of Dascena.

ETHICS STATEMENT

The data in this study were collected passively and deidentified in compliance with the Health Insurance Portability and Accountability Act and therefore did not require Institutional Review Board approval.

AUTHOR CONTRIBUTIONS

Data analysis and revision of the manuscript: Logan Ryan. *Data analysis and revision of the manuscript:* Jenish Maharjan. *Study conception and design:* Samson Mataraso. *Revision of manuscript:* Gina Barnes. *Revision of manuscript:* Jana Hoffman. *Study conception and revision of the manuscript:* Qingqing Mao. *Study conception:* Jacob Calvert and Ritankar Das.

ORCID

Gina Barnes  <http://orcid.org/0000-0002-8093-0245>

REFERENCES

1. Pulmonary embolism. *Nat Rev Dis Primer.* 2018;4:18031.
2. Beckman MG, Hooper WC, Critchley SE, Ortel TL. Venous thromboembolism: a public health concern. *Am J Prev Med.* 2010;38:S495–501.
3. Bělohávek J, Dytrych V, Linhart A. Pulmonary embolism, part I: epidemiology, risk factors and risk stratification, pathophysiology, clinical presentation, diagnosis and non-thrombotic pulmonary embolism. *Exp Clin Cardiol.* 2013;18:129–38.
4. Rivera-Lebron B, McDaniel M, Ahrar K, Alrifai A, Dudzinski DM, Fanola C, Blais D, Janicke D, Melamed R, Mohrien K, Rozycki E, Ross CB, Klein AJ, Rali P, Teman NR, Yarboro L, Ichinose E, Sharma AM, Bartos JA, Elder M, Keeling B, Palevsky H, Naydenov S, Sen P, Amoroso N, Rodriguez-Lopez JM, Davis GA, Rosovsky R, Rosenfield K, Kabrhel C, Horowitz J, Giri JS, Tapson V, Channick R. PERT C Diagnosis, treatment and follow up of acute pulmonary embolism: Consensus Practice from the PERT Consortium. *Clin Appl Thromb Off J Int Acad Clin Appl Thromb.* 2019;25:1076029619853037.
5. Duffett L, Castellucci LA, Forgie MA. Pulmonary embolism: update on management and controversies. *BMJ.* 2020;370:m2177.
6. Kearon C, Akl EA, Comerota AJ, Prandoni P, Bounameaux H, Goldhaber SZ, Nelson ME, Wells PS, Gould MK, Dentali F,

- Crowther M, Kahn SR. Antithrombotic therapy for VTE disease: Antithrombotic Therapy and Prevention of Thrombosis, 9th ed: American College of Chest Physicians Evidence-Based Clinical Practice Guidelines. *Chest*. 2012;141:e419S–96S.
7. Burles K, Innes G, Senior K, Lang E, McRae A. Limitations of pulmonary embolism ICD-10 codes in emergency department administrative data: let the buyer beware. *BMC Med Res Methodol*. 2017;17:89.
 8. Casez P, Labarère J, Sevestre M-A, Haddouche M, Courtois X, Mercier S, Lewandowski E, Fauconnier J, François P, Bosson JL. ICD-10 hospital discharge diagnosis codes were sensitive for identifying pulmonary embolism but not deep vein thrombosis. *J Clin Epidemiol*. 2010;63:790–7.
 9. Martin C, Sobolewski K, Bridgeman P, Boutsikaris D. Systemic Thrombolysis for pulmonary embolism: a review. *Pharm Ther*. 2016;41:770–5.
 10. Python Package Introduction—xgboost 1.4.0-SNAPSHOT documentation [cited 2021 Jan 19]. Available from: https://xgboost.readthedocs.io/en/latest/python/python_intro.html
 11. Chen T, Guestrin C. XGBoost: A Scalable Tree Boosting System. In: *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. San Francisco California USA: ACM. p. 785–94.
 12. Lundberg S, Lee S-I. A Unified approach to interpreting model predictions. *ArXiv170507874 Cs Stat*. 2017 [cited 2021 May 3]. Available from: <http://arxiv.org/abs/1705.07874>
 13. Moheimani F, Jackson DE. Venous thromboembolism: classification, risk factors, diagnosis, and management. *ISRN Hematol*. 2011;2011:124610.
 14. Braekkan SK, Mathiesen EB, Njolstad I, Wilsgaard T, Hansen JB. Hematocrit and risk of venous thromboembolism in a general population. the Tromsø study. *Haematologica*. 2010;95:270–5.
 15. Stein PD, Beemath A, Olson RE. Obesity as a risk factor in venous thromboembolism. *Am J Med*. 2005;118:978–80.
 16. Wells PS, Anderson DR, Rodger M, Stiell I, Dreyer JF, Barnes D, Forgie M, Kovacs G, Ward J, Kovacs MJ. Excluding pulmonary embolism at the bedside without diagnostic imaging: management of patients with suspected pulmonary embolism presenting to the emergency department by using a simple clinical model and d-dimer. *Ann Intern Med*. 2001;135:98–107.
 17. Le Gal G, Righini M, Roy P-M, Sanchez O, Aujesky D, Bounameaux H, Perrier A. Prediction of pulmonary embolism in the emergency department: the revised Geneva score. *Ann Intern Med*. 2006;144:165–71.
 18. Douma RA. Performance of 4 clinical decision rules in the diagnostic management of acute pulmonary embolism: a prospective cohort study. *Ann Intern Med*. 2011;154:709–18.
 19. Shen J-H, Chen H-L, Chen J-R, Xing JL, Gu P, Zhu BF. Comparison of the Wells score with the revised Geneva score for assessing suspected pulmonary embolism: a systematic review and meta-analysis. *J Thromb Thrombolysis*. 2016;41:482–92.
 20. Ageno W. The Wells rule is not accurate in hospitalized patients. *Nat Rev Cardiol*. 2015;12:449–50.
 21. Spyropoulos AC, Anderson FA Jr, FitzGerald G, Decousus H, Pini M, Chong BH, Zotz RB, Bergmann JF, Tapson V, Froehlich JB, Monreal M, Merli GJ, Pavanello R, Turpie A, Nakamura M, Piovella F, Kakkar AK, Spencer FA, IMPROVE I Predictive and associative models to identify hospitalized medical patients at risk for VTE. *Chest*. 2011;140:706–14.
 22. Zhou H, Hu Y, Li X, Wang L, Wang M, Xiao J, Yi Q. Assessment of the risk of venous thromboembolism in medical inpatients using the Padua Prediction Score and Caprini Risk Assessment Model. *J Atheroscler Thromb*. 2018;25:1091–104.
 23. Gibson CM, Spyropoulos AC, Cohen AT, Hull RD, Goldhaber SZ, Yusen RD, Hernandez AF, Korjian S, Daaboul Y, Gold A, Harrington RA, Chi G. The IMPROVEDD VTE Risk Score: incorporation of D-Dimer into the IMPROVE Score to improve venous thromboembolism risk stratification. *TH Open*. 2017;01:e56–65.
 24. Spyropoulos AC, Lipardi C, Xu J, Peluso C, Spiro TE, De Sanctis Y, Barnathan ES, Raskob GE. Modified IMPROVE VTE Risk Score and elevated D-Dimer identify a high venous thromboembolism risk in acutely ill medical population for extended thromboprophylaxis. *TH Open Companion J Thromb Haemost*. 2020;4:e59–65.
 25. Tsaplin S, Schastlivtsev I, Zhuravlev S, Barinov V, Lobastov K, Caprini JA. The original and modified Caprini score equally predicts venous thromboembolism in COVID-19 patients. *J Vasc Surg Venous Lymphat Disord*. 2021;9(6):1371–81.
 26. Fu Y, Liu Y, Chen S, Jin Y, Jiang H. The combination of Caprini risk assessment scale and thrombotic biomarkers to evaluate the risk of venous thromboembolism in critically ill patients. *Medicine (Baltimore)*. 2018;97:e13232.
 27. Cronin M, Dengler N, Krauss ES, Segal A, Wei N, Daly M, Mota F, Caprini JA. Completion of the Updated Caprini Risk Assessment Model (2013 Version). *Clin Appl Thromb*. 2013;25:1076029619838052. <https://doi.org/10.1177/1076029619838052>
 28. Zhang C, Sun M, Wei Y, Zhang H, Xie S, Liu T. Automatic segmentation of arterial tree from 3D computed tomographic pulmonary angiography (CTPA) scans. *Comput Assist Surg*. 2019;24:79–86.
 29. Liu W, Liu M, Guo X, Zhang P, Zhang L, Zhang R, Kang H, Zhai Z, Tao X, Wan J, Xie S. Evaluation of acute pulmonary embolism and clot burden on CTPA with deep learning. *Eur Radiol*. 2020;30:3567–75.
 30. Banerjee I, Sofela M, Yang J, Chen JH, Shah NH, Ball R, Mushlin AI, Desai M, Bledsoe J, Amrhein T, Rubin DL, Zamanian R, Lungren MP. Development and Performance of the Pulmonary Embolism Result Forecast Model (PERFORM) for computed tomography clinical decision support. *JAMA Netw Open*. 2019;2, 198719:e198719.
 31. Banerjee I, Ling Y, Chen MC, Hasan SA, Langlotz CP, Moradzadeh N, Chapman B, Amrhein T, Mong D, Rubin DL, Farri O, Lungren MP. Comparative Effectiveness of Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) Architectures for Radiology Text Report Classification. *Artif Intell Med*. 2019;97:79–88.
 32. Banerjee I, Chen MC, Lungren MP, Rubin DL. Radiology report annotation using intelligent word embeddings: applied to multi-institutional chest CT cohort. *J Biomed Inform*. 2018;77:11–20.
 33. Chen MC, Ball RL, Yang L, Moradzadeh N, Chapman BE, Larson DB, Langlotz CP, Amrhein TJ, Lungren MP. Deep learning to classify radiology free-text reports. *Radiology*. 2018;286:845–52.

34. Weikert T, Nestic I, Cyriac J, Bremerich J, Sauter AW, Sommer G, Stieltjes B. Towards automated generation of curated datasets in radiology: application of natural language processing to unstructured reports exemplified on CT for pulmonary embolism. *Eur J Radiol.* 2020;125:108862.
35. Nafee T, Gibson CM, Travis R, Yee MK, Kerneis M, Chi G, AlKhalfan F, Hernandez AF, Hull RD, Cohen AT, Harrington RA, Goldhaber SZ. Machine learning to predict venous thrombosis in acutely ill medical patients. *Res Pract Thromb Haemost.* 2020;4:230–7.
36. Cohen AT, Harrington R, Goldhaber SZ, Hull R, Gibson CM, Hernandez AF, Kitt MM, Lorenz TJ. The design and rationale for the Acute Medically Ill Venous Thromboembolism Prevention with Extended Duration Betrixaban (APEX) study. *Am Heart J.* 2014;167:335–41.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Ryan L, Maharjan J, Mataraso S, Barnes G, Hoffman J, Mao Q, Calvert J, Das R, Predicting pulmonary embolism among hospitalized patients with machine learning algorithms. *Pulmonary Circulation.* 2022;12:e12013. <https://doi.org/10.1002/pul2.12013>