Short-term Clinical Deterioration After Acute Pulmonary Embolism

Anthony J. Weekes, MD, MSc (Principal Investigator), Jason T. Nomura, MD, Dasia Esener, MD, Jeremy S. Boyd, MD, Patrick M. Ockerse, MD, Stephen Leech, MD, H. James Norton, PhD

The Charlotte-Mecklenburg Hospital Authority d/b/a Atrium Health

Project Period: 9/30/2018–8/31/2021

Federal Project Officer: Darryl T. Gray, MD, ScD, FACC, FAHA

Funding: This project was supported by grant number 5R01HS025979 from the Agency for Healthcare Research and Quality. The content is solely the responsibility of the authors and does not necessarily represent the official views of the Agency for Healthcare Research and Quality. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of manuscripts.

Grant Number: 5R01HS025979

STRUCTURED ABSTRACT (250 word maximum)

PURPOSE: Determine prognostic performance of more than 120 clinically available candidate predictor variables for short-term serious adverse events in pulmonary embolism (PE) patients. We included biochemical and imaging assessments for right ventricular (RV) abnormalities. Outcomes included quality of life at 30 days.

SCOPE: We conducted a prospective, multicenter, observational study of patients diagnosed with acute PE at emergency departments (EDs) at teaching hospitals within six separate US healthcare systems. All study sites participate in the Pulmonary Embolism Short-term Clinical Outcomes Registry.

METHODS: We used prospective registry data from six EDs. We aimed to evaluate prognostic performances of biochemical, imaging, and ECG assessments for RV abnormalities for clinical deterioration endpoints within 5 days of PE and to develop and validate a prediction model for clinical deterioration within 5 days of PE. Primary outcome was a composite of death, delayed circulatory or respiratory dysfunction, respiratory failure intervention, and escalated PE interventions within 5 days of PE diagnosis. Secondary outcomes were nonfatal bleeding, recurrence of venous thromboembolism, and hypoxia requiring oxygen supplementation.

RESULTS: We enrolled 1008 patients and included 935 patients in final analyses, 209 (22.4%) of whom had primary composite outcomes. Patient-reported quality of life was affected as early as 30 days after acute PE. We showed better performance of an RVD inclusive model over an RVD exclusive model after vetting over 100 candidate variables. RV assessments added significant prognostic value. Our research culminated in the development and validation of a prognostic model, which we converted into a simple points model for clinical application.

KEY WORDS: pulmonary embolism, echocardiography, right ventricle, clinical prediction rule, prognosis, validation study, outcomes, clinical deterioration, risk assessment

PURPOSE

Our long-range goal was to improve the quality and safety of care provided to patients presenting to emergency departments (EDs) nationwide with pulmonary embolism (PE), which would result in improved health outcomes for this patient population. The objective of this study was to compare right ventricular dysfunction dependent and independent prognostic models for short-term serious adverse events in PE patients. We determined the prognostic performance of more than 120 clinically available candidate predictor variables for short-term serious adverse events in pulmonary embolism (PE) patients. We included biochemical and imaging assessments for right ventricular abnormalities as predictor variables.

SCOPE

Background

An important indicator of acute PE of moderate to high severity is an acute increase in right ventricular pressure or size or decreased systolic function. PE-provoked right ventricle (RV) abnormality is commonly assessed in two ways: 1) laboratory surrogates of myocardial stretch and injury and 2) imaging assessments for RV dilatation, pressure increases, and decreased systolic function. The most common diagnostic tests are natriuretic peptide, troponin, and imaging by computed tomography (CT) and echocardiography. Assessments for abnormal RV (abnlRV) are absent in validated clinical prognostic models, such as the original and simplified Pulmonary Embolism Severity Index (PESI and sPESI) and Hestia.[1–3] These prognostic prediction models utilize a limited set of candidate variables without pertinent imaging and laboratory measurements.[4] Risk of early clinical deterioration from worsening RV function is not captured in current prediction models.[5–7]

Although newer anticoagulants offer an improved efficacy, convenience, and safety profile in PE treatment than previous anticoagulants, there is still hesitancy to identify PE patients eligible for early discharge with outpatient management soon after diagnosis. Hospitalization for PE is as high as 90%–95% in the US and Europe, despite 41%–51% of PE patients being classified as low-risk by existing clinical prediction models.[8–12] Clinical algorithms, checklists, and prognostic models are being developed and updated to optimize the safety of outpatient management, improve prognostic accuracy for outcome(s), and provide guidance to reduce practice variation. Incorporation of imaging and laboratory assessments for PE-provoked abnlRV have now been incorporated into hybrid clinical algorithms,[1,7,13–16] and some meta-analyses now support use of one or multiple RV assessment methods.[4,17,18] A consistent definition of PE-provoked abnlRV, however, is lacking.[19–22]

Context

Thus, acute care providers are challenged to identify PE patients who are considered low-risk (and safe for early discharge) and those at greater risk of clinical deterioration without a clear guideline on RV assessment in acute PE. Providers must make disposition decisions driven by concerns for acute deterioration (respiratory failure, cardiac arrest, new dysrhythmia, sustained

hypotension, and rescue reperfusion intervention) within the first days of PE diagnosis rather than events at 30 days or later.

Settings

We conducted a prospective, multicenter, observational study of patients diagnosed with acute PE at six EDs across the US. The participating EDs are located at teaching hospitals within six separate healthcare systems. Each healthcare system had an electronic medical record (EMR) that allowed integrated access to affiliated centers in each region. Each site had an academic emergency medicine residency program and an ultrasound fellowship program. All study sites are part of a PE research consortium that populates a longstanding PE short-term clinical outcomes registry (PESCOR), which is registered on clinicaltrials.gov. The overall goal of the registry is to optimize risk stratification of ED patients with acute PE to identify patient need for hospital-based monitoring and interventions within 5 days of PE diagnosis.

Participants

The study population was ED patients with confirmed acute PE within 12 hours of ED presentation. The primary outcome was a composite of death, delayed circulatory or respiratory dysfunction, hypoxia, and reperfusion intervention within 5 days of PE diagnosis. Secondary outcomes were nonfatal bleeding, recurrence of venous thromboembolism, and hypoxia requiring oxygen supplementation.

METHODS

Study Design, Data Sources, and Data Collection

This was a prospective, observational, multicenter study using two registry databases. The first database was the Pulmonary Embolism Short-term Clinical Outcomes Registry (PESCOR; clinicaltrials.gov NCT02883491), a prospective registry of patients who presented to six urban, academic EDs in the following locations during the pilot: San Diego, California; Newark, Delaware; Orlando, Florida; Charlotte, North Carolina; Nashville, Tennessee; and Salt Lake City, Utah. The cohort was chosen to allow for broad generalizability. By enrolling patients from a diverse set of EDs with geographic spread, we expected to capture the full spectrum of demographics and acute PE severity at presentation. The second registry was created after federal funding was secured for development of the prediction model (Short-term Clinical Deterioration After Acute Pulmonary Embolism; clinicaltrials.gov NCT03915925). The unfunded initial registry (PESCOR) was used for the validation. Both registries were populated by the same six EDs and had similar variables, data recording instruments, and outcome variables.

The development database was prospectively accrued between September 18, 2018, and December 14, 2020. The validation database was built between August 2016 to March 2019. The central site (located in Charlotte, North Carolina) prospectively enrolled consecutive patients; the other five sites prospectively enrolled on a convenience basis. During the early stages of the unfunded registry, the central site enrolled patients with written informed consent until its

institutional review board (Atrium Health IRB) approved waiver of informed consent. The other five sites enrolled with written informed consent with approval from each of their institutional review boards. Once federal funding was secured, all sites used the central IRB (Advarra IRB), which approved the study protocol and waiver of written informed consent for enrollments at all sites (Advarra approval code PRO-00029256).

The electronic case report included over 400 variable entry fields for prognostic model testing and other aims of the registry. For the prognostic tool, we collected 138 data elements on each patient, including vital signs at presentation, risk factors for PE, comorbidities, contemporaneous measurements of cardiac biomarkers [troponin and brain natriuretic peptide (BNP)], and CT and goal-directed echocardiography (GDE) evaluations performed early in ED management of the index PE event.

The prognostic model was developed from 935 PE patients. Univariable analysis of 138 candidate variables was followed by penalized and standard logistic regression on 26 retained variables and then tested with a validation database (N = 801).

Measures

The primary composite outcome was death (all cause or PE-related) or clinical deterioration within 5 days of index PE confirmation. Deaths were classified as PE-related when the site investigator reviewed the case and determined death was not likely to be due to another cause, such as septic shock or acute myocardial infarction. Elements of clinical deterioration included respiratory failure, cardiac arrest, new dysrhythmia, sustained hypotension requiring intravenous volume expansion or adrenergic medication, and rescue reperfusion intervention.

Our secondary outcome included all components of the primary composite outcome plus major bleeding, recurrence of venous thromboembolism (VTE), or subsequent hospitalization within 30 days of the index PE.

Aim-specific Analyses

Aim 1: We risk stratified each patient with the sPESI, modified HESTIA, and modified European Society of Cardiology (ESC) approaches. For each risk category, we compared proportions of patients with the primary and secondary outcomes within five days of PE diagnosis, as well as the sensitivity, specificity, and positive and negative likelihood ratios. The proportion of reclassifications was determined. After data cleaning, we closed data entry and created reports for sPESI, Hestia, and by 2019 ESC 'low risk' and 'NOT low risk' categorization criteria. We created reports of patients with and without the primary outcome. We reported on the diagnostic accuracy of each of the different PE risk stratification approaches and reported the proportion with disagreements.

Aim 2: We determined functional outcomes of PE patients 30 days after PE using the validated pulmonary embolism quality of life (PEMQoL) questionnaire and compared the incidence of variables associated with subjects' PEMQoL scores. We reported PEMQoL scores for patients who completed PEMQoL questionnaires.

Aim 3: The primary analyses used logistic regression with clinical deterioration (yes/no) as the outcome variable. Independent variables for the first regression model included the components of the sPESI and HESTIA criteria. Sensitivity, specificity, positive predictive value, and negative predictive value (with their corresponding 95% confidence intervals) were reported for each of the four RVD assessment methods and compared with the algorithm developed from the logistic regression. We reported and compared the diagnostic accuracy for the outcome variable of the presence or absence of elevated serum troponin, serum brain natriuretic peptide, CT RV:LV ratio, and goal directed echocardiography derived assessment of the presence or absence of right ventricular dysfunction. We determined which of these approaches to assessing right ventricular dysfunction best improved the prediction model.

Aim 4: Following the analyses described above for Aim 3, a second regression model added results of the four RVD assessment methods as a modification to the ESC approach. This tested whether any of the RVD assessments could improve the prediction model. We determined simple and weighted agreement of the risk categorization systems. We applied the best performing RVD variable to the PE risk stratification model and then reported on the simple and weighted agreement of sPESI, HESTIA, 2019 ESC, and the derived PE risk prediction model approaches.

Aim 5: We derived a prediction model using the current AHRQ-funded cohort and performed a validation study of the derived prediction model on the cohort from the research consortium's pilot project (unfunded) plus any excess enrollments accrued during the AHRQ funding period. [Validation cohort = prospectively obtained cases entered in the PESCOR registry by the same sites between late 2016 and mid-2018.] We compared predicted and actual outcomes; determined the area under the curve (AUC); and reported demographics, predictors, and outcomes of the derivation vs. validation cohorts for comparison of case mix.

RESULTS and DISCUSSION

We enrolled a total of 1,008 patients and included 935 patients in the final analyses (exceeding the desired sample size of 880).

INCLUSION OF AHRQ PRIORITY POPULATIONS: The mean age of enrollees was 60 $\pm$ 17 years, with minimum and maximum ages of 18 and 104 years, respectively. There were 446 enrollees (44%) aged 65 or older. Children were excluded from this study. Forty-nine percent of enrollees (n = 493) were women, and 34% (n = 345) were racial/ethnic minorities. Fifty-five percent of enrollees (n = 559) had a Charlson comorbidity index of 1 or higher, which may indicate they have special healthcare needs. Our actual enrollment demographics were different from what we anticipated at time of grant application (see **Tables 1** and **2** below). We anticipated more than half of enrollees would be women; however, the proportion of women actually enrolled was 49%. Racial/ethnic representation (34%) was also less than the 52% we anticipated. We did not make any projections in our application about special healthcare needs of enrollees, nor their age range.

Table 1: Planned (**Pre-Award**) Inclusion Enrollment Report Submitted

Racial Categories	**Ethnic Categories**				**Total**
	Not Hispanic or Latino		**Hispanic or Latino**		
	Female	Male	Female	Male	
American Indian/Alaska Native	0	0	0	0	0
Asian	4	2	0	0	6
Native Hawaiian or Other Pacific Islander	0	0	0	0	0
Black or African American	220	210	0	0	430
White	243	175	0	0	418
More than 1 Race	0	0	14	12	26
Total	467	367	14	12	880

Table 2: Cumulative (**Post-Award**) Inclusion Enrollment Report

Racial Categories	**Ethnic Categories**							**Total**
	Not Hispanic/Latino		**Hispanic/Latino**		**Unknown/Not Reported**			
	Female (F)	Male (M)	F	M	F	M	Unknown	
American Indian/Alaska Native	3	3	0	0	1	0	0	7
Asian	8	1	2	0	0	0	0	11
Native Hawaiian or Other Pacific Islander	3	0	0	0	0	0	0	3
Black or African American	131	127	1	1	4	6	0	270
White	275	336	33	22	15	9	0	690
More than 1 Race	0	0	0	0	0	0	0	0
Unknown/Not Reported	2	3	12	4	3	2	1	27
Total	422	470	48	27	23	17	1	**1008**

We will summarize the results of our aim-specific analyses next. We describe the dissemination of our study findings, including the status of public presentations and manuscript development and submission, at the end of this report.

Aim 1:
Information relevant to Specific Aim 1 is reported in a published manuscript (https://doi.org/10.1371/journal.pone.0260036).[23] We reported on the diagnostic accuracy of

each PE risk stratification approach and reported the proportion with disagreements. A univariable analysis reported the proportion of patients stratified by the primary composite outcomes in supplemental Table 1 (https://doi.org/10.1371/journal.pone.0260036.s001). We compared the diagnostic accuracy of sPESI, ESC, and PE-SCORE in supplemental Table 3 shown below and available at https://doi.org/10.1371/journal.pone.0260036.s003.

Supplemental Table 3: Prognostic Performance of sPESI and ESC at Low-Risk Threshold

Low-Risk sPESI on Development Database	**Primary Outcome +**	**Primary Outcome -**
No	178 (85.2%)	443 (61.0%)
Yes (sPESI = 0 points)	31 (14.8%)	283 (39.0%)

sensitivity 85.2% (79.6%–89.7%), specificity 39.0% (35.4%–42.6%), PPV 28.7% (27.0%–30.4%), NPV 90.1% (86.7%–92.8%), and accuracy 49.3% (46.1%–52.6%)

Low-risk ESC on Development Database	**Primary Outcome +**	**Primary outcome -**
No	208 (99.5%)	650 (89.5%)
Yes	1 (0.5%)	76 (10.5%)

sensitivity 99.5% (97.4%–99.9%); specificity 10.5% (8.3%–12.9%); PPV 24.2% (23.8 - 24.7%); NPV 98.7% (91.4%–99.8%), and accuracy 30.4% (27.4%–33.4%)

Last, to determine the difference in prognostic value of components of PE-SCORE and the component of sPESI, we determined upward and downward reclassifications by PE-SCORE followed by determination of prognostic errors or corrections for death or clinical deterioration by PE-SCORE. We used the combined development and validation databases in the PE-SCORE prognostic model report and compared binary classification into low-risk versus not low-risk by sPESI with that of a new 9-variable prognostic points model (PE-SCORE), where a score of 0 points was 'low-risk' and scores of 1–10 points were considered 'not low-risk.'

As mentioned before, of 1,569 PE patients with both PE-SCORE and sPESI scores, 24.5% experienced the primary outcome. sPESI classified 1,011/1,569 (64.4%) patients as low-risk whereas PE-SCORE classified 309/1,569 (19.6%) as low-risk ($p < 0.001$). The proportion of low-risk sPESI reclassified as "not low-risk" by PE-SCORE was 897/1,011 (88.7%), of which 33.8% (303/897) had clinical deterioration. The proportion of those NOT low-risk by sPESI reclassified as low-risk by PE-SCORE was 34.9%, of which 192/195 (98.4%) did not have the primary outcome.

For determining *not low-risk*, the sensitivity of PE-SCORE is 96.1% (94%–98%) [i.e., 370 of 385 patients with clinical deterioration correctly predicted as 'not at low-risk"], specificity 24.8% (22%–27%) [i.e., 294 of 890 *without* outcome were correctly predicted as *low-risk*), positive

predictive value 29.4% (27%–32%) [i.e., false-positive rate of 70.6%), and negative predictive value 95.1% (93%–98%). Alternatively, sPESI had sensitivity of 18.2% (14%–22%), specificity of 58.8% (56%–62%), positive predictive value of 12.5% (10%–15%), and negative predictive value of 68.8% (66%–72%).

Summary: Although prognosis of clinical deterioration events by PE-SCORE had 70% false-positive rate, PE-SCORE had very high sensitivity and negative predictive values, whereas sPESI had very low sensitivity and low specificity.

Aim 2:
Quality of life (QoL) is an important healthcare outcome. The disease-specific pulmonary embolism quality of life survey was used. The PEMQoL questionnaire contains six dimensions created based on the contents of the items: frequency of complaints, activities of daily living limitations, work-related problems, social limitations, intensity of complaints, and emotional complaints. In the questions, no complaint received a score of 1 point whereas severe or highly frequent, intense complaints received scores of 5 or 6 points. Therefore, higher numerical QoL domain or overall PEMQoL scores mean worse patient-reported experience for that domain of quality of life. In addition, we reported on the influence of quality of life on the outcome of subsequent hospitalization within 30 days (recidivism).

We fit independent multivariable linear regression models for each PEMQoL domain outcome as well as a multivariable linear regression model for PEMQoL score (the average score for each patient across standardized domain scores on a scale from 0–100). In each model, we included the four primary independent variables of interest: primary composite outcome, GDE showing RV abnormalities, escalated PE intervention used within 5 days, and PE score (as a continuous variable).

For *Frequency of Complaints*, none of the predictors were significantly related to the domain outcome. For *Activities of Daily Living*, patients with the primary composite outcome had a higher score ($p = 0.034$) than those without, while patients with escalated PE interventions had a lower PEMQoL score than those without. Patients with GDE showing RV abnormalities and who received escalated PE interventions had a significantly lower score relative to patients with GDE showing RV abnormalities ($p = 0.035$). Additionally, total PEMQoL score increased as PE score increased ($p = 0.032$).

For *Work-related Problems*, no predictor was significant except for subsequent hospitalizations. For *Social Limitations*, patients with the primary composite outcome had a higher score (marginally significant; $p = 0.056$) than those without. For combined *Intensity of Complaints,* patients with GDE showing RVD without escalated PE intervention had a marginally significant lower score than those without GDE showing RV abnormalities ($p = 0.07$). For *Emotional Complaints*, no predictors other than subsequent hospitalization was significant. For average PEMQoL score across domains, patients with the primary composite outcome had a higher score (worse QoL experience) with marginal significance ($p = 0.080$). Patients with GDE showing RV abnormalities and who received escalated PE interventions scored lower than patients without GDE showing RV abnormalities (marginally significant; $p = 0.059$). Patients with GDE showing

RV abnormalities who did not receive escalated PE interventions demonstrated significantly lower PEMQoL scores (better QoL) than patients without GDE showing RVD (p = 0.049).

For all domain totals and overall PEMQoL score, subsequent hospitalization was significantly associated with higher PEMQoL scores (worse QoL).

In conclusion, acute PE survivors who experienced one or more of the clinical deterioration events within 5 days of diagnosis had worse quality of life at 30 days in domains of *Activities of Daily Living* and *Social Limitations* whereas patients who received escalated PE interventions had lower QoL scores than those not receiving escalated PE interventions. Worse QoL of life was associated with an increase in subsequent hospitalizations within 30 days.

Aim 3:
The following information pertains to the development database. First, we used univariable analyses to report prognostic accuracy of the biochemical, imaging, and ECG assessments for RV abnormalities for the primary outcome of death or clinical deterioration endpoints. Second, we determined which of these right ventricular assessment methods best improved the prognostic model. We used multivariable logistic regression analysis.

Prognostic Performance of Laboratory Assessment of RV Abnormalities on the Development Database (unpublished data)		
Natriuretic peptide	Accuracy = 63.7% (60.5%–66.9%)	
Elevated BNP	sensitivity 56.2% (49.1–63.2%)	PPV 32.4% (29.0–36.0%)
Normal BNP	specificity 65.9% (62.2–69.4%)	NPV 83.8% (81.5 -86.0%)
True positive=113 False positive=236 False negative=88 True negative=456		
Troponin	Accuracy = 69.6% (66.5–72.5%)	
Elevated troponin	sensitivity 47.1% (40.18% to 54.14%)	PPV 36.4% (32.1 to 41.0%)
Normal troponin	specificity 76.1% (72.8–79.2%)	NPV 83.2% (81.2–85.0%)
True positive=98 False positive=171 False negative=110 True negative=544		

Prognostic Performance of Imaging Assessment of RV Abnormalities on the Development Database (unpublished data)		
CT RV:LV ratio	Accuracy 68.1% (65.0–71.1%)	
Ratio 1.0 or more	sensitivity 53.9% (46.8–60.9%)	PPV 35.6% (31.7– 39.7%)
Ratio < 1.0	specificity 72.1% (68.7–75.4%)	NPV 84.6% (82.4–86.5%)
True positive=110 False positive=199 False negative=94 True negative=515		
Goal directed echocardiography	Accuracy 70.6% (67.5–72.5%)	
Severe RV dilatation present	sensitivity 58.6% (51.5–65.5%)	PPV 38.4% (34.5–42.4%)
Severe RV dilatation absent	specificity 73.9% (70.5–77.0%)	NPV 86.5% (84.4–88.4%)
True positive=119 False positive= 191 False negative=84 True negative=540		

Within the published PE-SCORE manuscript[23] that reported on the development of the final prognostic model, we reported on:

1) Univariable and multivariable logistic regression analyses with clinical deterioration (yes/no) as the outcome variable: https://doi.org/10.1371/journal.pone.0260036.s001
2) Multivariable analysis showed CT imaging and echocardiography (GDE) assessments were better than biochemical or ECG predictors for the primary composite outcome: https://doi.org/10.1371/journal.pone.0260036.t002
    a) https://doi.org/10.1371/journal.pone.0260036.t003

In a separate and currently unpublished manuscript, we reported on electrocardiography findings in PE stratified by primary composite outcome, with significance defined by p value < 0.05. The most common ECG patterns were sinus tachycardia (38.9%); S1-Q3-T3 pattern (16.3%), incomplete or complete right bundle branch block (RBBB) 15.1%, and T wave inversion (TWI) V2-4 (14%). Significant ECG patterns for clinical deterioration by univariable analysis were sinus tachycardia, TWI in V2-4 , ST elevation (STE) in aVR, STE V1, supraventricular tachycardia (SVT) including atrial fibrillation, TWI II,III avF and ST depression V4-6.

All patterns were significant for abnlRV by GDE except left bundle branch block and left ventricular hypertrophy. Multivariable analysis for clinical deterioration showed that sinus tachycardia, TWI V2-4, STE aVR, and SVT were significant. Multivariable analysis for abnlRV by GDE showed sinus tachycardia, incomplete RBBB, S1-Q3-T3, and TWI V2-4 had significance.

In conclusion, the most common abnormal ECG patterns in PE were sinus tachycardia, S1-Q3-T3 RBBB, and TWI. Sinus tachycardia, TWI V2-4, STE aVR, and SVT (including atrial fibrillation) were significant for clinical deterioration whereas sinus tachycardia, incomplete RBBB, S1-Q3-T3, and TWI V2-4 had significance for abnlRV by GDE.

Aim 4:
In the peer-reviewed PE-SCORE manuscript,[23] we described development of the prognostic model. We compared the diagnostic accuracy of sPESI, ESC, and PE-SCORE in a supplemental file (https://doi.org/10.1371/journal.pone.0260036.s003).

In a separate manuscript (submitted but not yet published), we reported on the added prognostic value of RVD assessments for determining an even lower risk for primary outcome over the low-risk sPESI classification for the primary outcome. In that manuscript, we showed better performance of an RVD-inclusive model over an RVD-exclusive model, even after vetting over 100 candidate variables.

From the combined databases of 1,736 patients, we identified 611 (35.2%) low-risk sPESI patients. Of these 611 patients, 75 (12.3%) experienced early deterioration. We developed random forest models [full (with RV variables) and reduced (without RV variables)] and reported variable importance plots from full random forest and log odds from logistic regression models.

Proportions with abnormal RV by computed tomography (CT), echocardiography, troponin, and natriuretic peptide were 26.2%, 20.6%, 17.7%, and 23.1%, respectively. For deterioration, the receiver operating characteristics for full and reduced prognostic models were 0.80 (0.77–0.82) and 0.71 (0.68–0.73), respectively. RV assessments were the top four in the variable importance plot for the random forest model. Echocardiography and CT significantly increased predicted probability of deterioration in the regression model.

In conclusion, RV assessments improved prognostic model accuracy versus models without RV assessments and were the top predictors of deterioration, with echocardiography being ranked first.

Aim 5 (the most important specific aim):
We reported on development of the prognostic model in the following manuscript: https://doi.org/10.1371/journal.pone.0260036**.** Logistic regression yielded a nine-variable model, then simplified to a nine-point tool (PE-SCORE): one point each for abnormal RV by echocardiography, abnormal RV by computed tomography, systolic blood pressure < 100 mmHg, dysrhythmia, suspected/confirmed systemic infection, syncope, medicosocial admission reason, abnormal heart rate, and two points for creatinine greater than 2.0 mg/dL. In the development database, 22.4% had the primary outcome.

Prognostic accuracy of logistic regression model vs. PE-SCORE model: 0.83 (0.80, 0.86) vs. 0.78 (0.75, 0.82) using area under the curve (AUC) and 0.61 (0.57, 0.64) vs. 0.50 (0.39, 0.60) using precision-recall curve (AUCpr). In the validation database, 26.6% had the primary outcome. PE-SCORE had AUC 0.77 (0.73, 0.81) and AUCpr 0.63 (0.43, 0.81). As points increased, outcome proportions increased: a score of zero had 2% outcome whereas scores of six and above had ≥ 69.6% outcomes. In the validation dataset, PE-SCORE zero had 8% outcome [no deaths] whereas all patients with PE-SCORE of six and above had the primary outcome.

We concluded the manuscript by stating that the PE-SCORE model identifies PE patients at low- and high-risk for deterioration and may help guide decisions about early outpatient management versus need for hospital-based monitoring.[23]

In addition, we used the combined development and validation databases to assemble a machine learning prognostic model for the primary outcome using the large number of candidate predictors. We then compared the performance of PE-SCORE with the machine learning prognostic model, as shown below. Methodologic details follow.

PESCORE

Pulmonary Embolism Short-term Clinical Outcomes Risk Estimation

PE SCORE	Points assigned	
Suspected/confirmed systemic infection	No 0	Yes +1
Echo showing PE provoked RV abnormality	No 0	Yes +1
Abnormal Heart rate <50 or >100, bpm	No 0	Yes +1
CT RV:LV ratio ≥ 1.0	No 0	Yes +1
Preceding episode syncope	No 0	Yes +1
Creatinine > 2.0 mg/dL	No 0	Yes +2
Medical or social reason for hospitalization	No 0	Yes +1
Systolic BP < 100 mmHg	No 0	Yes +1
Dysrhythmia	No 0	Yes +1

PE SCORE was created for providers to determine the probability of death or clinical deterioration within 5 days of disposition decisions on emergency department patients with acute pulmonary embolism.

The primary composite outcome = PE-related death or clinical deterioration (respiratory failure, cardiac arrest, new dysrhythmia, sustained hypotension requiring intravenous volume expansion or adrenergic medication, or escalated PE interventions).

PE SCORE total range 0-10 points
Guide to Interpretation: In1625 ED patients diagnosed with acute pulmonary embolism that had complete PE SCORE elements, the overall incidence of primary outcomes= 24%

PE SCORE total points with proportion of patients with primary outcome:
0 points; 4.5% (15/331)
1 point; 11.5% (47/409)
2 points; 20.8% (74/355)
3 points; 39.8% (101/254)
4 points; 44.8% (78/174)
5 points; 70.6% (48/68)
6+points; 79.4% (27/34)

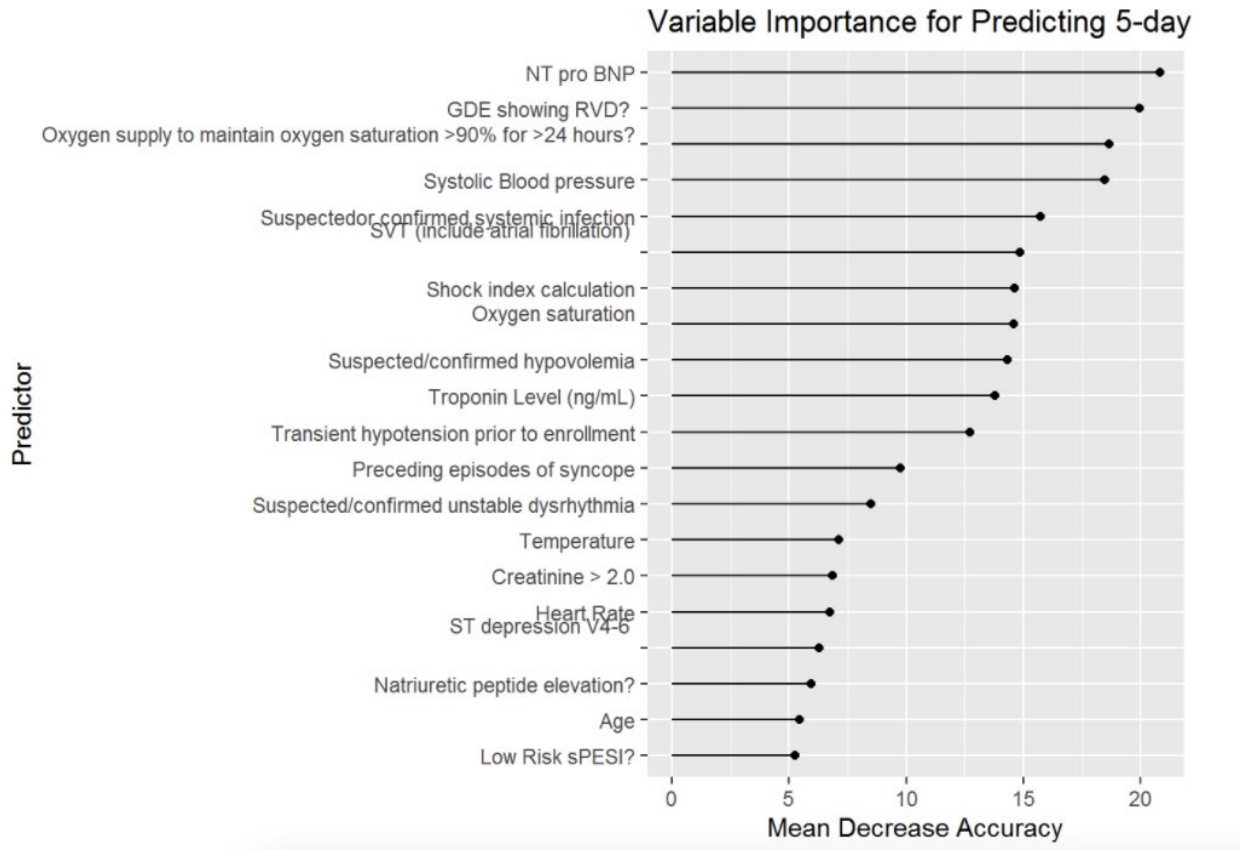
Goal-directed echocardiography (GDE) for PE-provoked RV abnormality as performed by emergency physicians using the following interpretation guidelines:
Severe RV dilatation = RV:LV basal diameter ≥ 1.0 or basal RV diameter > 42 mm with RV apex blunting on two or more different windows.
Severe RV systolic dysfunction:= estimate or measurements of tricuspid annular planar systolic excursion (TAPSE) being 10 mm or less and RV free wall hypokinesis.
Flattening or deviation of inter-ventricular septum towards the left ventricle.
GDE scoring range: 0-3
RV dilatation is a requirement for visual identification of PE-provoked RV abnormalities. The absence of RV dilatation is scored as 0, whereas 1 point each is assigned for RV dilatation, septal flattening or leftward deviation, and RV systolic dysfunction.
GDE scores of 1 to 3 may determined to be either acute or chronic. Chronicity of RV dilatation is based on RV free wall thickness ≥ 7 mm or if abnormal RV is documented on any previous echocardiography report.

The combined dataset contained 1,736 observations with 80 predictor variables, PE-SCORE, and the primary composite outcome. We used imputation by random forest (RF) to impute missing data. We sought to compare a RF prognostic model to the use of PE-SCORE. We split the data 70:30 for training and validation and compared overall model performance in terms of AUC (with 95% confidence intervals for each). For each prognostic model, we chose a decision threshold of 0.50 and compared results in terms of sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV). The results presented here reflect those from the validation dataset.

Results of the comparison of PE-SCORE and RF prognostic models: sensitivity 0.36 (0.28, 0.45) vs 0.69 (0.60, 0.77); specificity 0.89 (0.86, 0.92) vs 0.76 (0.72, 0.80); PPV 0.52 (0.41, 0.62) vs 0.48 (0.41, 0.55); NPV 0.81 (0.78, 0.85) vs 0.88 (0.85, 0.92); and AUC 0.76 (0.71, 0.81) vs 0.80 (0.76, 0.84). Comparing AUCs using Delong's test, the RF model was statistically significantly better than PE-SCORE ($p = 0.02$). Based on the training dataset, the decision threshold corresponding to a > 50% probability of the primary composite outcome corresponded to a PE-SCORE ≥ 4.

In conclusion, with a large dataset and field of candidate variables, the RF model had significantly better prognostic performance than PE-SCORE.

Method	Sensitivity	Specificity	Positive Predictive Value	Negative Predictive Value	AUC
PE Score	0.36 (0.28, 0.45)	0.89 (0.86, 0.92)	0.52 (0.41, 0.62)	0.81 (0.78, 0.85)	0.76 (0.71,0.81)
Random Forest	0.69 (0.6, 0.77)	0.76 (0.72, 0.8)	0.48 (0.41, 0.55)	0.88 (0.85, 0.92)	0.8 (0.76,0.84)



Limitations:

Although the validation was performed on a different database with data collected during a different time period, external validation should be conducted at sites outside the current registry. Our study focused on clinical deterioration and early mortality due to PE *severity*. We did not assess outcomes due to PE *treatment* (e.g., bleeding, bleeding risk, compliance with treatment), which would influence disposition decisions and need for safety outcomes. The study setting was focused on ED patients and ambulatory care settings where the cadence and feasibility of testing may not be generalizable to patients developing acute PE while already in the inpatient setting. Already hospitalized patients who develop acute PE may have different risk factors or susceptibilities to PE-associated deterioration from those diagnosed in an outpatient setting.

Our a priori study design included using troponin measurements as continuous data; however, institutional change in troponin assay at the central site interrupted plans to perform linear regression on the troponin variable. Similarly, two of the six sites used NT proBNP, but others used point-of-care BNP assay measurements. Therefore, we used institutional assay cutoffs to create categorical variables (troponin and natriuretic peptide elevation). Univariable analyses showed significant differences in mean troponin, point-of-care BNP, and NT proBNP measurements between outcome groups in both databases. Valuable information, however, may have been lost by converting a continuous variable into a categorical variable.[24]

Univariable analysis identified the clinical site itself as a variable of importance. The logistic regression model, therefore, has a random effects intercept for clinical sites. The random intercept cannot be used in a risk calculation on patients at sites outside of the six sites of this study, as the random effect of the new site is unknown. Thus, only the fixed intercept of the random effects model is used in the risk calculation. Model performance at a clinical site outside the six sites in this study may differ. Other discrete variables that may be of interest (e.g., insurance status, other social determinants of health) were not included in this study. Despite significant differences in patient characteristics between sites, the prognostic model performed well on patients.

Another possible limitation is that machine-based learning derivation techniques may offer better management of multiple variables (including those with interactions); however, our preliminary steps with classification tree analysis were not helpful. The logistic regression model we developed had an AUC of 0.83 (95% CI 0.80–0.86), whereas the PE-SCORE yielded an AUC of 0.78 on the development database. Although PE-SCORE had lower prognostic performance than the logistic regression model, PE-SCORE performed similarly by AUC on both databases and offers real-world usefulness at the site of care.

Although this study was performed at academic centers, competency in GDE has been expected of those emerging from emergency medicine residency training for the past decade. Our results may indicate an opportunity to study the impact of employing GDE into PE risk stratification. Upon external validation, any real-world application of PE-SCORE would include recommendation that technically difficult or uninterpretable GDE images limit full use of PE-SCORE. None of the other eight variables used to calculate PE-SCORE were missing during development.

When faced with absent GDE scores, providers should use available clinical information, recognizing that the worst-case scenario (that GDE is abnormal) has not been ruled out. Providers may either add a point or consider the partial PE-SCORE a minimum score. The other real-world option is to consider comprehensive echocardiography (by cardiology service).

Potential benefits of PE-SCORE include early detection of deterioration and avoidance of misclassification of patients who experience the outcome but would have been classified as low-risk by another prognostic tool. Potential harms may include unnecessary testing or interventions in those who did not experience any clinical deterioration outcomes despite higher-risk classification, subjecting them to potential adverse events of the interventions and increased lengths of stay and medical costs. After external validation, we anticipate use of the PE-SCORE tool in acute care settings with similar prevalence of early clinical deterioration to identify PE patients likely to benefit from early discharge and those who may need higher-level monitoring and escalated PE interventions. However, incorporation of any new prognostic tool into clinical practice requires implementation and impact studies to better understand the clinical consequences.[25]

Conclusions:

The prognostic models were assembled using a large field of candidate variables and identified 24.3% prevalence of important clinical deterioration endpoints within 5 days. RV assessments added significant prognostic value regardless of logistic regression or machine learning model. Patient-reported quality of life was affected as early as 30 days after acute PE based on these outcomes.

Significance and Implications:

The outcomes and the timeframe of 5 days that we focused on are important, are common, and impact quality of life. These outcomes should be factored into PE management considerations. We identified high-value predictors by both logistic regression and machine learning prognostic models. We demonstrated very good to strong prognostic performance with a simplified points model PE-SCORE (user-friendly), logistic regression model, and more complex out-of-bag random forest model.

REFERENCES

1. Jiménez D, Kopecna D, Tapson V, et al. Derivation and validation of multimarker prognostication for normotensive patients with acute symptomatic pulmonary embolism. Am J Respir Crit Care Med [Internet] 2014;189(6):718–26. Available from: http://dx.doi.org/10.1164/rccm.201311-2040OC

2. Aujesky D, Obrosky DS, Stone RA, et al. Derivation and validation of a prognostic model for pulmonary embolism. Am J Respir Crit Care Med [Internet] 2005;172(8):1041–6. Available from: http://dx.doi.org/10.1164/rccm.200506-862OC

3. Zondag W, Mos IC, Creemers-Schild D, et al. Outpatient treatment in patients with acute pulmonary embolism: the Hestia Study. J Thromb Haemost [Internet] 2011;9(8):1500–7. Available from: http://dx.doi.org/10.1111/j.1538-7836.2011.04388.x

4. Elias A, Mallett S, Daoud-Elias M, Poggi J-N, Clarke M. Prognostic models in acute pulmonary embolism: a systematic review and meta-analysis. BMJ Open [Internet] 2016;6(4):e010324. Available from: http://dx.doi.org/10.1136/bmjopen-2015-010324

5. Kabrhel C, Sacco W, Liu S, Hariharan P. Outcomes considered most important by emergency physicians when determining disposition of patients with pulmonary embolism. Int J Emerg Med [Internet] 2010;3(4):239–64. Available from: http://dx.doi.org/10.1007/s12245-010-0206-8

6. Kabrhel C, Okechukwu I, Hariharan P, et al. Factors associated with clinical deterioration shortly after PE. Thorax [Internet] 2014;69(9):835–42. Available from: http://thorax.bmj.com/content/69/9/835.long

7. Vinson DR, Drenten CE, Huang J, et al. Impact of relative contraindications to home management in emergency department patients with low-risk pulmonary embolism. Ann Am Thorac Soc [Internet] 2015;12(5):666–73. Available from: http://dx.doi.org/10.1513/AnnalsATS.201411-548OC

8. Singer AJ, Thode HC Jr, Peacock WF th. Admission rates for emergency department patients with venous thromboembolism and estimation of the proportion of low risk pulmonary embolism patients: a US perspective. Clin Exp Emerg Med [Internet] 2016;3(3):126–31. Available from: https://www.ncbi.nlm.nih.gov/pubmed/27752630

9. Dentali F, Di Micco G, Giorgi Pierfranceschi M, et al. Rate and duration of hospitalization for deep vein thrombosis and pulmonary embolism in real-world clinical practice. Ann Med [Internet] 2015;47(7):546–54. Available from: http://dx.doi.org/10.3109/07853890.2015.1085127

10. Jiménez D, de Miguel-Díez J, Guijarro R. Trends in the management and outcomes of acute pulmonary embolism: analysis from the RIETE registry. Journal of the American [Internet] 2016;Available from: https://www.jacc.org/doi/abs/10.1016/j.jacc.2015.10.060

11. Konstantinides SV. Trends in Pulmonary Embolism Outcomes: Are We Really Making Progress? [Internet]. J. Am. Coll. Cardiol. 2016;67(2):171–3. Available from: http://dx.doi.org/10.1016/j.jacc.2015.10.062

12. Mastroiacovo D, Dentali F, di Micco P, et al. Rate and duration of hospitalisation for acute pulmonary embolism in the real-world clinical practice of different countries: analysis from the RIETE registry [Internet]. European Respiratory Journal. 2019;53(2):1801677. Available from: http://dx.doi.org/10.1183/13993003.01677-2018

13. Kabrhel C, Rosovsky R, Baugh C, et al. Multicenter Implementation of a Novel Management Protocol Increases the Outpatient Treatment of Pulmonary Embolism and Deep Vein Thrombosis [Internet]. Academic Emergency Medicine. 2018;Available from: http://dx.doi.org/10.1111/acem.13640

14. Bledsoe JR, Woller SC, Stevens SM, et al. Management of Low-Risk Pulmonary Embolism Patients Without Hospitalization: The Low-Risk Pulmonary Embolism Prospective

Management Study. Chest [Internet] 2018;154(2):249–56. Available from: http://dx.doi.org/10.1016/j.chest.2018.01.035

15. Konstantinides S, Meyer G, Becattini C. 2019 ESC Guidelines for the diagnosis and management of acute pulmonary embolism developed in collaboration with the European Respiratory Society (ERS …. Eur Heart J [Internet] 2020;Available from: https://researchportal.helsinki.fi/en/publications/2019-esc-guidelines-for-the-dignosis-and-management-of-acute-pulm

16. Bova C, Sanchez O, Prandoni P, et al. Identification of intermediate-risk patients with acute symptomatic pulmonary embolism. Eur Respir J [Internet] 2014;44(3):694–703. Available from: http://dx.doi.org/10.1183/09031936.00006114

17. Andrade I, Mehdipoor G, Le Mao R, et al. Prognostic significance of computed tomography-assessed right ventricular enlargement in low-risk patients with pulmonary embolism: Systematic review and meta-analysis. Thromb Res [Internet] 2021;197:48–55. Available from: https://www.sciencedirect.com/science/article/pii/S0049384820305909

18. Barco S, Mahmoudpour SH, Planquette B, Sanchez O, Konstantinides SV, Meyer G. Prognostic value of right ventricular dysfunction or elevated cardiac biomarkers in patients with low-risk pulmonary embolism: a systematic review and meta-analysis. Eur Heart J [Internet] 2019;40(11):902–10. Available from: https://www.ncbi.nlm.nih.gov/pubmed/30590531

19. Lahm T, Douglas IS, Archer SL, et al. Assessment of Right Ventricular Function in the Research Setting: Knowledge Gaps and Pathways Forward. An Official American Thoracic Society Research Statement. Am J Respir Crit Care Med [Internet] 2018;198(4):e15–43. Available from: http://dx.doi.org/10.1164/rccm.201806-1160ST

20. Huang SJ, Nalos M, Smith L, Rajamani A, McLean AS. The use of echocardiographic indices in defining and assessing right ventricular systolic function in critical care research. Intensive Care Med [Internet] 2018;Available from: https://www.ncbi.nlm.nih.gov/pubmed/29789861

21. Cho JH, Kutti Sridharan G, Kim SH, et al. Right ventricular dysfunction as an echocardiographic prognostic factor in hemodynamically stable patients with acute pulmonary embolism: a meta-analysis. BMC Cardiovasc Disord [Internet] 2014;14:64. Available from: http://dx.doi.org/10.1186/1471-2261-14-64

22. Sanchez O, Trinquart L, Colombet I, et al. Prognostic value of right ventricular dysfunction in patients with haemodynamically stable pulmonary embolism: a systematic review. Eur Heart J [Internet] 2008;29(12):1569–77. Available from: http://eurheartj.oxfordjournals.org/content/ehj/29/12/1569.full.pdf

23. Weekes AJ, Raper JD, Lupez K, et al. Development and validation of a prognostic tool: pulmonary embolism short-term clinical outcomes risk estimation (PE-SCORE). PLoS One [Internet] 2021;Available from: https://doi.org/10.1371/journal.pone.0260036

24. Royston P, Altman DG, Sauerbrei W. Dichotomizing continuous predictors in multiple regression: a bad idea. Stat Med [Internet] 2006;25(1):127–41. Available from: http://dx.doi.org/10.1002/sim.2331

25. Vickers AJ, Van Calster B, Steyerberg EW. Net benefit approaches to the evaluation of prediction models, molecular markers, and diagnostic tests. BMJ [Internet] 2016;352:i6. Available from: http://dx.doi.org/10.1136/bmj.i6

LIST OF PUBLICATIONS and PRODUCTS: Bibliography of outputs from the study.

1. Development and validation of a prognostic tool: **p**ulmonary **e**mbolism **s**hort-term **c**linical **o**utcomes **r**isk **e**stimation (PE-SCORE) manuscript submitted to PLOS ONE journal.[23]

2. Added prognostic value of right ventricular assessments for low-risk pulmonary embolism - manuscript submitted to Academic Emergency Medicine journal. Virtual Oral Abstract presented at American College of Emergency Physicians (ACEP) Oct 2021.

3. Electrocardiographic findings in acute pulmonary embolism (*AHRQ/PESCOR databases*) -- manuscript in development still. Goals: abstract submission to Society of Academic Emergency Medicine and manuscript submission end of Dec 2021.

4. Association of pulmonary embolism presentation characteristics with quality of life at one month (*AHRQ/PESCOR databases*). Goals: abstract submission to Society of Academic Emergency Medicine by early Dec 2021, and manuscript submission end of Dec 2021.