Reassessing Anticoagulation Risk in Incidental Pulmonary Embolism: A Multicenter Study Using Unsupervised Machine Learning and Bayesian Methods

Sophie Samuel, Ada Selina Jutba, Haley Parker, Gregory Pon, Jennifer Cortes
Memorial Hermann Health System, Department of Pharmacy

BACKGROUND

- Incidentally diagnosed pulmonary embolism (PE), found on imaging for other reasons (e.g., cancer staging, infection workup), presents unique management challenges.
- Unlike PE diagnosed on clinical suspicion via CT pulmonary angiography, incidental PE lacks specific management guidance.
- Although tools like sPESI and institutional bleeding scores are commonly used, their predictive value in incidental PE is uncertain.
- We assessed bleeding outcomes in incidental PE and explored data-driven phenotyping to better stratify bleeding risk.

STUDY OVERVIEW /METHODS

Integrated health system (4 hospitals: 1 academic + 3 community)
Study period: January 1, 2023 – September 30, 2024 • Retrospective cohort

Population: Adults with acute pulmonary embolism Identified by CT pulmonary angiography (CTPA) or ventilation–perfusion scan (V/Q), or incidentally on non–PE-protocol CT	**Key exclusions:** Prior anticoagulation • Outside-hospital diagnosis/transfer No PE • Contraindication to anticoagulation
Incidental PE: (found on non–PE-protocol imaging)	**Suspected PE:** (diagnosed by CT pulmonary angiography or V/Q)

Primary outcome: Major bleeding (composite)
Secondary outcomes: Bleeding components; in-hospital mortality

Analytic framework (Bayesian):
- Binomial models (pooled; site-standardized) → posterior risk and P(Incidental > Suspected)
- Multivariable logistic (association) → posterior odds ratio
- Unsupervised latent class analysis (pre-treatment) → physiology-abnormal vs comorbidity-predominant

RESULTS

Variable	Overall	Incidental	Suspected	p-value
Site distribution	TMC=193, MC=105, W=21, SW=116 (n=435)	TMC=45, MC=5, W=2, SW=19 (n=71)	TMC=148, MC=100, W=19, SW=97 (n=364)	—
Age (years) median (IQR)	68.0 [55.5-77.0]	69.0 [60.0-82.0]	67.0 [55.0-76.0]	0.224
Female, n (%)	226 (52.0%)	40 (56.3%)	186 (51.1%)	0.439
Race/Ethnicity, n (%)				
African American	142 (32.6%)	26 (36.6%)	116 (31.9%)	0.489
White	131 (30.1%)	19 (26.8%)	112 (30.8%)	0.573
Other	162 (37.2%)	26 (36.6%)	136 (37.4%)	1.000
Ethnicity: Hispanic	51 (11.7%)	6 (8.5%)	45 (12.4%)	0.424
Past medical history, n (%)				
Hypertension	163 (37.5%)	26 (36.6%)	137 (37.6%)	0.894
Diabetes Mellitus	160 (36.8%)	35 (49.3%)	125 (34.3%)	0.022
Congestive heart disease	140 (32.2%)	20 (28.2%)	120 (33.0%)	0.489
Chronic kidney disease	108 (24.8%)	24 (33.8%)	84 (23.1%)	0.071
COPD	52 (12.0%)	3 (4.2%)	49 (13.5%)	0.027
Liver disease	33 (7.6%)	5 (7.0%)	28 (7.7%)	1.000
VTE chemical prophylaxis, n (%)				
Heparin	45 (10.3%)	13 (18.3%)	32 (8.8%)	0.030
Enoxaparin	58 (13.3%)	13 (18.3%)	45 (12.4%)	0.184
PE diagnosis, n (%)				
CTPA	364 (83.7%)	0 (0.0%)	364 (95.0%)	0.001
V/Q	18 (4.1%)	0 (0.0%)	18 (4.9%)	0.054
Incidental	71 (16.3%)	71 (100.0%)	0 (0.0%)	0.001

RESULTS

Variable	Overall	Incidental	Suspected	p-value
Site distribution	TMC=193, MC=105, W=21, SW=116 (n=435)	TMC=45, MC=5, W=2, SW=19 (n=71)	TMC=148, MC=100, W=19, SW=97 (n=364)	—
Contrast imaging performed, n (%)				
Chest	17 (3.9%)	13 (18.6%)	4 (1.1%)	0.001
Abdomen/Pelvis	18 (4.1%)	15 (21.1%)	3 (0.8%)	0.001
Chest/Abd/Pelvis	60 (13.8%)	33 (46.5%)	27 (7.4%)	0.001
sPESI, n (%)				
age > 80	90 (20.7%)	20 (28.2%)	70 (19.2%)	0.108
Cancer	92 (21.1%)	18 (25.4%)	74 (20.3%)	0.343
HF or chronic lung disease	148 (34.0%)	17 (23.9%)	131 (36.0%)	0.056
HR > 110 bpm	123 (28.3%)	19 (26.8%)	104 (28.6%)	0.886
SBP < 100 mmHg	60 (13.8%)	14 (19.7%)	46 (12.6%)	0.131
o2 saturation < 90%	106 (24.4%)	5 (7.0%)	101 (27.7%)	0.001
sPESI total, median[IQR]	1 (0-1)	1(0-2)	1(0-2)	0.171
Clinical presentation, n (%)				
Chest pain	95 (21.8%)	5 (7.0%)	90 (24.7%)	0.001
Dyspnea	254 (58.4%)	10 (14.1%)	244 (67.0%)	0.001
Syncope	43 (9.9%)	6 (8.5%)	37 (10.2%)	0.653
Hemoptysis	7 (1.6%)	1 (1.4%)	6 (1.7%)	1.000
DVT	65 (14.9%)	7 (9.9%)	58 (15.9%)	0.273
PE classification, n (%)				
Saddle	31 (7.1%)	2 (2.8%)	29 (8.0%)	0.203
Main PA	77 (17.7%)	10 (14.1%)	67 (18.4%)	0.497
Lobar	123 (28.3%)	23 (32.4%)	100 (27.5%)	0.392
Segmental	247 (56.8%)	39 (54.9%)	208 (57.1%)	0.794
Subsegmental	149 (34.3%)	24 (33.8%)	125 (34.3%)	1.000
Echo: Right heart strain	107 (24.6%)	12 (16.9%)	95 (26.1%)	0.131
Worsening clinical presentation and escalation of therapy, n (%)				
Progression on imaging	3 (0.7%)	0 (0.0%)	3 (0.8%)	1.000
Hemodynamic instability	15 (3.4%)	3 (4.2%)	12 (3.3%)	0.721
Echo	5 (1.1%)	1 (1.4%)	4 (1.1%)	0.592
Escalation of therapy (tPA)	21 (4.8%)	2 (2.8%)	19 (5.2%)	0.551
Catheter-directed thrombolysis	17 (3.9%)	2 (2.8%)	15 (4.1%)	1.000
Surgical Thrombectomy	31 (7.1%)	4 (5.6%)	27 (7.4%)	0.802

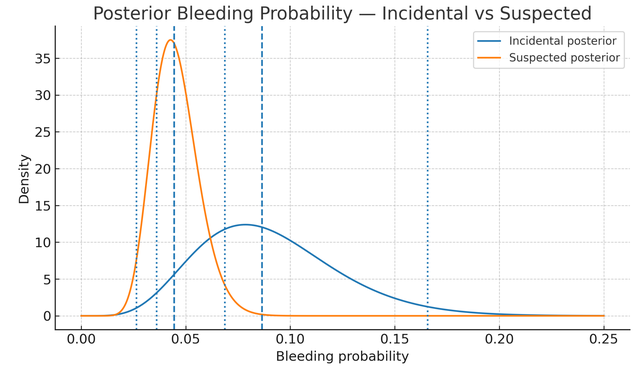
Risk categories (reported as 0 vs ≥1): **low risk (sPESI=0)** and **higher risk (sPESI≥1)** — overall 29.7% / 70.3% ; incidental 33.8% / 66.2%; suspected 28.8% / 71.2%

CONCLUSION

- **Where incidental PE is found:** cancer staging/screening, abdominal/post-op issues, GI bleed/hemoptysis, aortic/renal evaluations, hypoxia/pleural effusion, infection/sepsis & endocarditis work-ups.
- **Pooled & site-standardized binomial:** both show a high posterior probability (>0.9) that incidental PE has more bleeding than suspected PE across sites.
- **Adjusted multivariable (Bayesian logistic):** the direction persists but remains uncertain due to few events → wide credible intervals and only moderate posterior probability; chronic kidney disease is the only covariate with a consistent independent association.
- **Why LCA (phenotyping):** the cohort is heterogeneous; standard scores (e.g., sPESI) don't separate physiologic instability from comorbidity burden. LCA, using pre-treatment features (no outcome leakage), identifies patterns to generate testable hypotheses about bleeding under anticoagulation.
- **Phenotype finding (interpretation):** a physiology-abnormal phenotype showed more bleeding than a comorbidity-predominant phenotype. This suggests, not proves, a framework to study timing decisions: pair phenotype with PE location (segmental/subsegmental vs central), right-heart strain, and the underlying diagnosis prompting imaging (e.g., infection, malignancy) to prospectively test whether immediate inpatient anticoagulation vs brief stabilization/stepwise start affects bleeding without worsening PE outcomes.
- **Bottom line:** signals are hypothesis-generating. Use phenotype + CKD + PE anatomy/echo + clinical context to inform future prospective evaluation of initiation timing, not to replace guideline-based care today.

Major bleeding (composite) & components — Bayesian pooled binomial

Outcome	Incidental n (%)	Suspected n (%)	RR (95% CrI)	P(Inc>Sus)
Major bleed	6 (8.5%)	16 (4.4%)	1.94 (0.73–4.50)	0.92
GI bleed	2 (2.8%)	8 (2.2%)	1.36 (0.24–5.02)	0.66
ICH	1 (1.4%)	2 (0.5%)	2.78 (0.21–23.23)	0.81
Retroperitoneal bleed	0 (0.0%)	4 (1.1%)	0.28 (0.00–4.02)	0.22
PRBC ≥10 in 24 h	1 (1.4%)	7 (1.9%)	0.84 (0.07–4.14)	0.43
PRBC ≥4 in 1 h	2 (2.8%)	4 (1.1%)	2.67 (0.43–12.43)	0.87
In-hospital mortality	8 (11.3%)	26 (7.1%)	1.59 (0.71–3.19)	0.88



Major bleeding — Bayesian binomial (neutral priors)

Analysis	Posterior risk (%)		Risk difference % (95% CrI)	P(Inc>Sus)
	Incidental (95% CrI)	Suspected (95% CrI)		
Pooled binomial	8.6% (3.6–16.6%)	4.4% (2.6–6.9%)	4.2% (−1.4–12.3%)	0.92
Site-adjusted binomial	9.9% (3.8–22.9%)	4.8% (2.9–7.3%)	5.1% (0.1–17.3%)	0.96

Major bleeding — Bayesian multivariable logistic (neutral priors)

Predictor	Posterior OR	95% CrI	P(OR>1)
Incidental (vs suspected)	1.76	0.65–4.76	0.87
sPESI, per point	1.03	0.71–1.50	
CKD (yes vs no)	3.10	1.27–7.52	
Male (vs female)	0.86	0.36–2.06	

Phenotypes (LCA, pre-treatment; K=2)

Phenotype	n (%)	Major bleed %	Top indicators
Physiology-abnormal	61 (18.9)	11.5	HR ≥110; SBP <100; O_2 <90
Comorbidity-predominant	374 (81.1)	5.6	Cancer; CKD; HF/CLD

LCA = latent class analysis; CLD = chronic lung disease; CKD = chronic kidney disease; classes derived from sPESI items (age >80, cancer, HF/CLD, HR ≥110, SBP <100, O_2 <90) plus CKD; outcomes not used to form classes; bleeding shown descriptively (no p-values). Percentages are of total cohort (N=435).

DISCLOSURES

- Authors have nothing to disclose