The Need for Human Oversight in AI: False Positive Findings of Incidental Pulmonary Embolism and Intracranial Hemorrhage.

Kirby Quinn MD, Ayden Jacob MD, Cinthia Del Toro BS, Thiago Braga MD, Felipe Munera MD, Alexander McKinney MD, Jean Jose DO

University of Miami Miller School of Medicine, Department of Radiology

PURPOSE

- Evaluate incidence and characteristics of false positive findings generated by an AI-guided triage system designed to detect incidentally found PE (iPE) and intracranial hemorrhage (ICH)

MATERIAL & METHODS

- Data from the AI software implemented at our institution was analyzed
- February 6th-September 30th 2024
- AI-software triaged routine chest and abdomen CT performed with contrast
- Positive AI findings → cross-referenced with radiologist interpretation

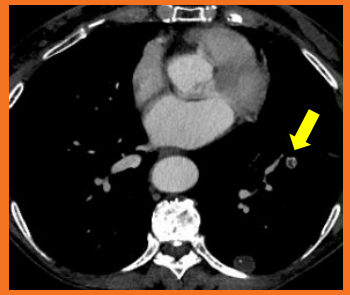
EXCLUSION CRITERIA:

- Studies performed in the inpatient or emergency room setting
- CT angiography (CTA) chest studies performed specifically to evaluate for PE

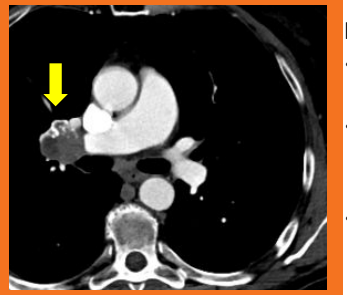
Kquinn@med.miami.edu

	Total # cases identified by the AI algorithm	False Positives
iPE	n=88	20 (15%)
ICH	n=90	27 (30%)

iPE False Positives	CH False Positives
Hilar lymphadenopathy (25%)	Hyperdense and/or calcified lesion (33%)
Pulmonary mass or nodule (20%)	Algorithm flagging normal DaTscan studies (22%)
Pulmonary vein or venous structure (10%)	Vascular structure or lesion (15%)



Pulmonary nodule/mass false identified by the AI algorithm as iPE.



Hilar lymphadenopathy falsely identified by the AI algorithm as iPE.



Left frontal AVM falsely identified by AI algorithm as ICH.



Extra-axial metastases falsely identified by AI algorithm as ICH.

RESULTS:

- N= 9,590 CT evaluated by AI algorithm
- **90 ICH identified**, 63 (70%) confirmed by radiologist, 27 (30%) false positives
- **88 iPE identified**, 68 (75%) confirmed by radiologist, 20 (15%) false positives
- ICH False Positives: 33% hyperdense or calcified lesion, 22% algorithm flagging normal DaTscan studies, 15% vascular structure or lesion.
- iPE False Positives: 25% hilar lymphadenopathy, 20% pulmonary mass or nodule, 10% pulmonary vein

CONCLUSION:

- AI-driven triage systems can enhance identification and prioritization of iPE and ICH
- However, our false positive findings highlight the need for *Human Intervention (HI)* in combination with AI to ensure accurate triage and diagnosis