

Prospective Clinical Evaluation of the Miss Rate of Unassisted and AI-Assisted Radiologists for Detection of Pulmonary Emboli on Dedicated Chest CT Angiography

Steven Rothenberg, MD, Assistant Professor, University of Alabama at Birmingham

Mostafa Abozeed; Satinder Singh; Omar Hamki; Cody Savage; Srinidhi Tridandapani; Andrew Smith

Introduction

Chest CT angiography (CTA) is the gold standard for diagnosing pulmonary emboli (PE), and both radiologists and artificial intelligence (AI) algorithms have high diagnostic accuracy for detection of PE on a dedicated chest CTA.

Hypothesis

The miss rate of unassisted radiologists for detection of PE on a dedicated chest CTA will be higher than the miss rate of radiologists assisted by a dedicated AI algorithm for detection and triage.

Methods

For this IRB-approved prospective single-center study, 1529 consecutive patients who underwent dedicated chest CTA for suspected PE were recruited in two phases. In phase 1, consecutive chest CTAs (N=503) were evaluated independently by a radiologist and by a FDA-cleared AI triage algorithm. This data was used to establish a baseline for unassisted radiologist and AI interpretations. Following this, all radiologists interpreting chest CTAs (N=41) were trained to review the results of the AI algorithm and triage system concurrently during image evaluation and reporting. In phase 2, consecutive chest CTAs (N=1026) were evaluated by radiologists using AI assistance for all cases. Reports and addendums were manually reviewed for the presence or absence of PE. A chest radiologist determined the ground truth by reviewing all chest CTAs with discrepancies between the radiologist and AI algorithm results. The miss rate (false negative rate), accuracy, sensitivity, and specificity were calculated.

Results

In phase 1 and 2, 12.7% (64/503) and 14.3% (147/1026) of chest CTAs were positive for PE, respectively ($p=0.395$). In phase 1 and 2, 4.97% (25/503) and 3.31% (34/1026) of exams had discrepant results between the radiologist's final report and the AI algorithm result, respectively ($p=0.126$). The miss rate by unassisted radiologists in phase 1 was 0.60% (3/503), statistically higher than 0.10% (1/1026) for radiologists with AI assistance in phase 2 ($p=0.037$). Accuracy, sensitivity, and specificity were 0.99, 0.95 and 0.99 for the unassisted radiologists in phase 1, 0.96, 0.72, and 0.96 for the AI algorithm in phase 1, and 0.99, 0.99, and 1.00 for the radiologists with AI assistance in phase 2.

Conclusion

In a single-center prospective real-world study, the miss rate of unassisted radiologists for detection of PE on a dedicated chest CTA was statistically higher than the miss rate of radiologists assisted by an AI algorithm for detection and triage.

Figures

Phase	Arm	Sens	Spec	Accuracy	Miss Rate	PE	Discrepant
1	Unassisted Rad	0.95	0.99	0.99	0.6% (3/503)	12.7% (64/503)	5.0% (25/503)
	AI Alone	0.72	0.96	0.96	3.6% (18/503)		
2	AI Assisted Rad	0.99	1.00	0.99	0.1% (1/1026)	14.3% (147/2026)	3.3% (34/1026)
	AI Alone	0.86	0.98	0.97	1.9% (20/1026)		

Table 1. Comparative effectiveness results for the detection of PE on dedicated chest CTA for pulmonary arterial vasculature evaluation. The miss rate is also defined as the false negative rate. Discrepant results are discordances between AI results and manually reviewed reports.

Keywords

Artificial Intelligence; Quality Improvement & Quality Assurance

SIIM22 Presentation