Re-evaluating Right Heart Strain at CT Pulmonary Angiography with Artificial Intelligence-Driven Cardiac Chamber Segmentation

Bardia Nadim, MD, Resident Physician, Radiology, Brigham and Women's Hospital
Ian Pan, MD

Introduction

Pulmonary embolism (PE) is a life-threatening condition and the third most common cardiovascular disease. Prompt diagnosis is critical. Computed tomography pulmonary angiography (CTPA) is the mainstay of diagnosis, allowing for PE detection and assessment of right heart strain (RHS) by measuring the right ventricular (RV) to left ventricular (LV) diameter ratio. An RV:LV < 1 is considered normal; however, manual 2D assessment is prone to interobserver variability. Artificial intelligence (AI) offers promising automated cardiac segmentation methods. This study evaluates the concordance of AI volumetric segmentation with manual RV:LV ratings and explores AI's value in assessing RHS on CTPA.

Hypothesis

AI volumetric cardiac segmentation will demonstrate high concordance with manual assessments of the RV:LV on CTPA and offer a standardized evaluation of RHS.

Methods

7,111 CTPAs from the RSNA PE Detection Challenge underwent automated cardiac segmentation by the publicly available TotalSegmentator algorithm. Manual review of a randomly selected subset demonstrated 98% of segmentations were satisfactory. RV:LV was calculated using: 1) the greatest short-axis chamber diameter (2D) and 2) segmentation volumes (3D). 2D RV:LVs were binarized, and concordance with the provided manual ratings of RV:LV ≥1 was calculated using Cohen's kappa for PE-positive examinations. Two-sample t-test compared AI-driven ratios between manual groups. Pearson correlation coefficient compared the 2D and 3D AI ratios. One-way ANOVA compared AI ratios between negative PE, acute peripheral PE, and acute central PE examinations.

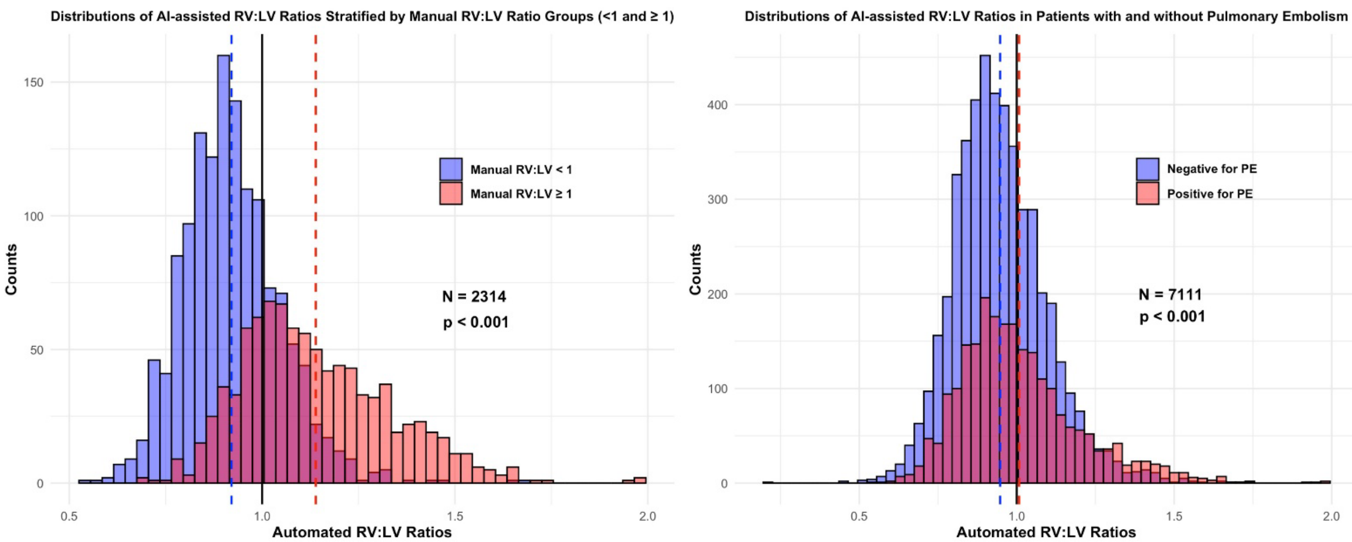
Results

2,314 examinations were positive for PE. The 2D AI ratio means for the manual RV:LV < 1 or ≥1 groups were 0.92 and 1.14, respectively ($p < 0.001$). Cohen's kappa between 2D AI ratio ≥1 and manual ratio ≥1 was 0.492. The 2D AI ratio means for negative and positive PE groups were 0.95 and 1.01, respectively ($p < 0.001$), whereas 3D ratio means were 1.36 and 1.55, respectively ($p < 0.001$). The Pearson correlation coefficient between the 2D and 3D ratios was 0.8. One-way ANOVA of negative PE (n = 5,380, mean = 0.954), acute peripheral PE (n=1,408, mean=0.98), and acute central PE (n=323, mean=1.12) examinations was significant ($p < 0.001$).

Conclusion

AI-driven cardiac segmentation is accurate and reproducible, though only fair-to-moderate concordance was seen between AI and manual RV:LV. There was high correlation between 2D and 3D AI RV:LV, indicating that greatest short-axis can serve as a proxy for volume; however, 2D methods overall underestimated RV:LV. Patients with acute central PE had the highest RV:LV, indicating that this may serve as a measure of RHS and PE severity.

Figure(s)



A B

Figure 1A Distributions of 2D AI RV:LV ratios stratified across binary manual RV:LV ratio < 1 and ≥ 1 groups with a statistically significant difference in the mean 2D AI ratios (0.92 [blue dashed line] vs 1.14 [red dashed line]). However, there is still considerable overlap between the two distributions. (B) Distributions of 2D AI RV:LV ratios in negative and positive pulmonary embolism (PE) groups with a statistically significant difference in the mean ratios (0.94 [blue dashed line] vs 1.01 [red dashed line]). Note the wider overlap between the two distributions, suggesting low correlation between RV:LV ratios and presence of PE.

Keywords

Artificial Intelligence/Machine Learning; Clinical Workflow & Productivity; Quality Improvement & Quality Assurance