Interobserver agreement in the diagnosis of acute pulmonary embolism from computed tomography pulmonary angiography and on the effectiveness of computer-aided diagnosis

To the Editor,

We read with great interest the article by Costantino et al [1] and their response [2] to the correspondence by Chartrand-Lefebvre [3] regarding interobserver agreement in the interpretation of computed tomography (CT) pulmonary angiography (CTPA) for the diagnosis of acute pulmonary embolism (PE). The interobserver agreement among radiologists is high for the diagnosis of massive (ie, large central) PE but is diminished for the diagnosis of segmental and subsegmental PEs. A similar hierarchy of agreement has been demonstrated for conventional pulmonary angiography [4,5]. The observations of Costantino et al are convincing and reflect the reality of interpreting CTPA examinations.

CTPA is an outstanding diagnostic tool for patients suspected of PE, but the accuracy of the interpretation of CTPA depends on the radiologist's training and expertise; it has been shown that a lack of dedicated experience in CTPA interpretation results in poorer interpretative performance relative to expert CT interpreters [6,7]. Outside academic practice, general radiologists are frequently called upon to interpret CTPA studies on an emergency basis. It is therefore likely that false-positive and false-negative CTPA interpretations are not infrequent in clinical practice, as illustrated in Fig. 1. The potential for overlooked patients with PE or unnecessary anticoagulation in patients caused by false-positive interpretation of CTPA is increased with the rapid proliferation of this technology [8,9]. These data are difficult to quantify because CTPA has been accepted as the gold standard for the diagnosis of PE, and catheter pulmonary angiograms are rarely performed for the diagnostic confirmation or exclusion of PE. Accordingly, incorrectly interpreted CTPA examinations are discovered only when CTPA studies are presented to an expert thoracic radiologist for review. As noted by Costantino et al [1,2] and Chartrand-Lefebvre [3], CTPA interpretation by a "second expert" would be useful in the case of nonmassive PE. However, the number of patients undergoing CTPA examinations has increased by an order of magnitude over the past decade, and the rate of positive CTPA studies is only 5% to 10% [10,11] in clinical practice; therefore, it may not be feasible to have expert CT radiologists review all negative or nonmassive PE CTPA diagnoses in clinical practice.

Computer-aided diagnosis (CAD) is a nascent field [12,13], but it promises to improve the efficiency and accuracy of the

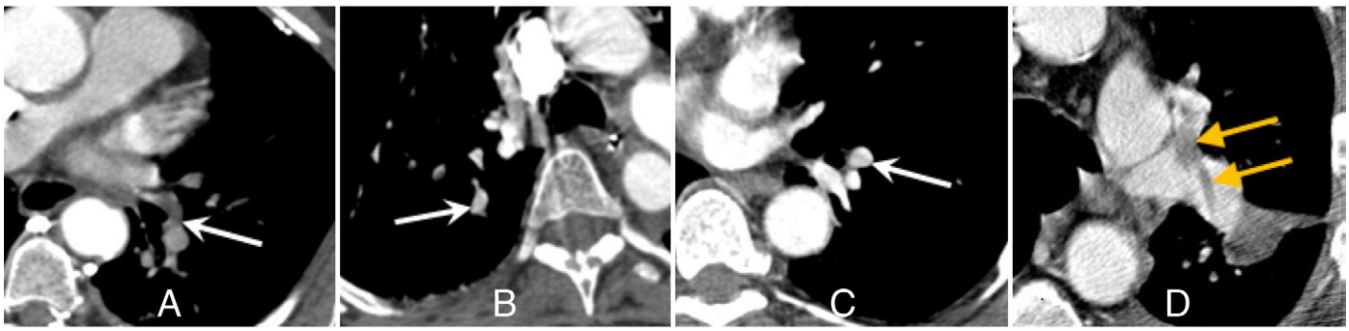

Fig. 1 Examples of false-positive (A-C) and false-negative (D) interpretations for PE on CTPA by general radiologists. The examination was ordered from the emergency department (patient A) and the inpatient units (patients B, C, and D). Examinations A, B, and C were initially interpreted as positive (arrows), and patient D was interpreted as negative for PE by general radiologists. Anticoagulation was begun for patients A, B, and C, and withheld for patient D. An experienced pulmonologist was consulted who then found patients A, B, and C were at low risk but patient D was at high risk for PE. Review of these examinations by an experienced thoracic radiologist found no evidence of PE for patients A, B, and C but confirmed PE for patient D in the left pulmonary artery (arrows). The finding mistaken for PE by general radiologists was a lymph node (patient A) and motion artifact for patients B and C.

CTPA diagnosis of PE. We have been developing a CAD system [14-16] for the diagnosis of PE, which has undergone preliminary evaluation in several clinical studies:

(1) Clinical study 1 [17] evaluated our system with 43 patients suspected of PE. The mean overall sensitivity for our system alone was 83% at a specificity of 80%. All CTPA studies were also read by 3 radiologists (R1, R2, R3), and their sensitivities were 87%, 82%, and 77%, respectively. With the aid of our system, reader sensitivities increased to 98%, 93%, and 92%, respectively ($P < .0001$), indicating that the detection performance of radiologists can be improved through the use of our CAD system.

(2) Clinical study 2 [18] tested our system with 40 patients, each read by 6 general radiologists and simultaneously and automatically processed by our CAD system. The readers detected 157 (74%) of 212 emboli, with a sensitivity of 97% (63/65) for central and 70% (103/147) for peripheral emboli with 9 false-positive findings. Our PE CAD algorithm detected 168 (79%) of 212 emboli, reaching a sensitivity of 74% (48/65) for central and 82% (120/147) for peripheral emboli. A total of 154 CAD PE candidates were considered false positives, yielding an average of 3.85 false positives per case. This was an earlier version of our system that was tuned to detect peripheral emboli. These results indicate that our system offers benefits for CTPA interpretation when used as a second reader, especially for peripheral emboli.

(3) Clinical study 3 [19] examined the influence of our PE CAD system on interpretations performed by radiology residents in 25 patients with a total of 107 confirmed emboli as the reference standard. Four radiology residents first independently analyzed all CTPA studies and then reanalyzed all studies with the aid of our system. PE-level (on a per-embolus basis) and patient-level[1] (on a per-patient basis) analyses were performed. At the PE level, the improvement in radiology resident sensitivity was from 46.7% to 52.3% (reader 1), from 57.9% to 59.8% (reader 2), from 55.1% to 60.7% (reader 3), and from 54.2% to 62.6% (reader 4). The mean PE-level sensitivity increased from 53.5% to 58.9 ($P < .028$). The mean false-positive rate of CAD was 2.4 per case with respect to the reference standard. At the patient level, the improvement was from 70.8% to 83.3% (reader 1), from 79.2% to 87.5% (reader 2), and from 91.7% to 100% (reader 4), whereas reader 3 maintained 100% performance without and with CAD. Our CAD system enabled correct PE diagnosis for an additional 4 patients overall.

(4) Clinical study 4 [23][2] used the PE detection capability of our CAD system to prospectively investigate its potential value in PE severity assessment through the characterization of embolic burden. This study used 58 patients with PE whose CTPA studies were analyzed by 4 observers to assess the embolic burden using the Mastora index [25] and by our CAD system. Interobserver agreement for PE severity was enhanced by consensus with CAD data, and the perceptual scoring errors were significantly decreased after CAD consensus ($P \leq .005$). Misclassifications of PE risk groups occurred in 27.6%, 24.1%, 5.2%, and 5.2% of patients for readers 1 to 4, respectively, and were corrected by CAD consensus in 56.3%, 36%, 33.3%, and 33.3% of misclassified patients, respectively ($P < .05$). These results indicate that CAD provides an incremental improvement in the accuracy of embolic load assessment compared with the radiologist's interpretation and that it may therefore be a valuable adjunct for risk stratification for patients with PE.

Our CAD system is not designed to replace radiologists but rather to enhance and extend the radiologist's capabilities through a synergy between the computer and the radiologist [26]. The 4 clinical studies collectively demonstrate that this CAD system offers capabilities that complement those of radiologists, leading to overall increased sensitivity and reduced variability associated with the CTPA diagnosis of PE, as well as enhanced consistency in the Mastora index for risk stratification in patients with PE. We expect that the improved diagnostic capability will result in fewer complications related to PE overdiagnosis as well as a reduction in complications resulting from PE underdiagnosis, and that the enhanced capability in risk stratification will ultimately lead to the improved management of patients afflicted with life-threatening PE, one of the most difficult diagnostic conditions and a major health problem in the United States according to the Surgeon General [27].

Jianming Liang PhD
Department of Biomedical Informatics
Arizona State University
Arizona Biomedical Collaborative
Phoenix, AZ 85004, USA

Michael B. Gotway MD
Scottsdale Medical Imaging
an affiliate of Southwest Diagnostic Imaging
Scottsdale, AZ, USA

[1] Existing CAD algorithms [12-16,20-22] are designed for PE-level detection (localizing individual emboli). This study simply used the capability offered by our system in PE-level detection [14-16] to assist the residents in achieving their patient-level diagnoses (excluding non–PE patients and dispatching PE patients to treatment).

[2] Note that the segmental vessel-based sensitivity, which is used in Engelke et al [23,24], is not a suitable measure for evaluating CAD performance. An embolus may extend into multiple vessel segments, but there is no need to detect the same embolus multiple times. By design, our CAD algorithm aims to provide a single CAD finding per embolus [14-16]. The embolus-based sensitivity reported by the same investigators in a recent study [23] is more meaningful both clinically and algorithmically.

Diagnostic Radiology/Biomedical Imaging and Pulmonary/Critical Care Medicine
Department of Radiology
University of California
San Francisco, CA, USA
Department of Biomedical Informatics
Arizona State University
Scottsdale, AZ 85258, USA

Demetri Terzopoulos PhD
Computer Science Department University of California
Los Angeles, Los Angeles, CA 90095, USA

H. Dirk Sostman MD
The Methodist Hospital System
Houston, TX 77030, USA
Weill Cornell Medical College
New York, NY 10065, USA

doi:10.1016/j.ajem.2010.12.030

References

[1] Costantino G, Norsa AH, Amadori R, et al. Interobserver agreement in the interpretation of computed tomography in acute pulmonary embolism. Am J Emerg Med 2009;27:1109-11.
[2] Costantino G, Rusconi A, Montano N. Interobserver agreement in computer tomography readings for pulmonary embolism. Am J Emerg Med 2011;29(1):119.
[3] Chartrand-Lefebvre C. Computed tomography angiography in the diagnosis of pulmonary embolism: interobserver agreement. Am J Emerg Med 2011;29(1):118-9.
[4] Quinn MF, Lundell CJ, Klotz TA, Finck EJ, Pentecost M, McGehee WG, Garnic JD. Reliability of selective pulmonary angiography in the diagnosis of pulmonary embolism. AJR 1987;149:469-71.
[5] Stein PD, Henry JW, Gottschalk A. Reassessment of pulmonary angiography for the diagnosis of pulmonary embolism: relation of interobserver agreement to the order of the involved pulmonary arterial branch. Radiology 1999;210:689-91.
[6] Pugliese F, Hunink MG, Gruszczynska K, et al. Learning curve for coronary CT angiography: what constitutes sufficient training? Radiology 2009;251:359-68.
[7] van Rossum AB, van Erkel AR, van Persijn van Meerten EL, Ton ER, Rebergen SA, Pattynama PMT. Accuracy of helical CT for acute pulmonary embolism: ROC analysis of observer performance related to clinical experience. Eur Radiol 1998;8:1160-4.
[8] Trowbridge RL, Araoz PA, Gotway MB, Bailey RA, Auerbach AD. The effect of helical computed tomography on diagnostic and treatment strategies in patients with suspected pulmonary embolism. Am J Med 2004;116(2):84-90.
[9] Donohoo JH, Mayo-Smith WW, Pezzullo JA, Egglin TK. Utilization patterns and diagnostic yield of 3421 consecutive multidetector row computed tomography pulmonary angiograms in a busy emergency department. J Comput Assist Tomogr 2008;32:421-5.
[10] Le Gal G, Bounameaux H. Diagnosing pulmonary embolism: running after the decreasing prevalence of cases among suspected patients. J Thromb Haemost 2004;2(8):1244-6.
[11] Stein PD, Matta F. Acute pulmonary embolism. Curr Probl Cardiol 2010;35(7):314-76.
[12] Bastarrika GM, Blackmon KNB, Ravenel JGM, Costello PM, Schoepf UJM. The diagnostic value of computer-aided detection at multi-detector row spiral computed tomography for pulmonary embolism. Clinical Pulmonary Medicine 2009;16(2):101-5.
[13] Chan HP, Hadjiiski L, Zhou C, Sahiner B. Computer-aided diagnosis of lung cancer and pulmonary embolism in computed tomography—a review. Acad Radiol 2008;15(5):535-55.
[14] Liang J, Bi J. Computer aided detection of pulmonary embolism with tobogganing and multiple instance classification in CT pulmonary angiography. Inf Process Med Imaging 2007;20:630-41.
[15] Bi J, Liang J. Multiple instance learning of pulmonary embolism detection with geodesic distance along vascular structure. Computer Vision and Pattern Recognition, 2007 CVPR '07 IEEE Conference on; 2007. p. 1-8.
[16] Liang J, Bi J. Local characteristic features for computer aided detection of pulmonary embolism in CT angiography. Proceedings of the First MICCAI Workshop on Pulmonary Image Analysis; 2008. p. 263-72.
[17] Das M, Mühlenbruch G, Helm A, Bakai A, Salganicoff M, Stanzel S, Liang J, Wolf M, Günther RW, Wildberger JE. Computer-aided detection of pulmonary embolism: influence on radiologists' detection performance with respect to vessel segments. Eur Radiol 2008;18(7):1350-5.
[18] Buhmann S, Herzog P, Liang J, Wolf M, Salganicoff M, Kirchhoff C, Reiser M, Becker CH. Clinical evaluation of a computer aided diagnosis (CAD) prototype for the detection of pulmonary embolism. Acad Radiol 2007;14(6):651-8.
[19] Lake DR, Seekins JM, Cluver AR, Fernandez MJ, Ravenel JG, Schoepf UJ, Salganicoff M, Liang J, Nguyen S, Dundar M, Wolf M, Costello P. Computer-aided detection of peripheral pulmonary embolus on multi-detector row CT: initial experience and impact on resident diagnosis. The 106th annual meeting of the American Roentgen Ray Society, April 30 - May 5, 2006. Vancouver, BC, Canada; 2006.
[20] Zhou C, Chan H, Hadjiiski LM, Chughtai A, Patel S, Cascade PN, et al. Automated detection of pulmonary embolism (PE) in computed tomographic pulmonary angiographic (CTPA) images: multiscale hierarchical expectation-maximization segmentation of vessels and PEs. Proc SPIE 2007;6514(1):65142F.
[21] Zhou C, Chan HP, Sahiner B, Hadjiiski LM, Chughtai A, Patel S, et al. Computer-aided detection of pulmonary embolism in computed tomographic pulmonary angiography (CTPA): performance evaluation with independent data sets. Med Phys 2009;36(8):3385-96.
[22] Bouma H, Sonnemans JJ, Vilanova A, Gerritsen FA. Automatic detection of pulmonary embolism in CTA images. Med Imaging, IEEE Trans 2009;28(8):1223-30.
[23] Engelke C, Schmidt S, Auer F, Rummeny EJ, Marten K. Does computer-assisted detection of pulmonary emboli enhance severity assessment and risk stratification in acute pulmonary embolism? Clin Radiol 2010;65(2):137-44.
[24] Engelke C, Schmidt S, Bakai A, Auer F, Marten K. Computer-assisted detection of pulmonary embolism: performance evaluation in consensus with experienced and inexperienced chest radiologists. Eur Radiol 2008;18(2):298-307.
[25] Mastora I, Remy-Jardin M, Masson P, Galland E, Delannoy V, Bauchart JJ, et al. Severity of acute pulmonary embolism: evaluation of a new spiral CT angiographic score in correlation with echocardiographic data. Eur Radiol 2003;13(1):29-35.
[26] Doi K. Computer-aided diagnosis in medical imaging: historical review, current status and future potential. Comput Med Imaging Graph 2007;31(4-5):198-211.
[27] US DHHS.. Surgeon general's call to action to prevent deep vein thrombosis and pulmonary embolism. US Department of Health and Human Services; 2008.