

UC San Diego

UC San Diego Previously Published Works

Title

Automated Deep Learning Analysis for Quality Improvement of CT Pulmonary Angiography.

Permalink

<https://escholarship.org/uc/item/5456g4sm>

Journal

Radiology Artificial Intelligence, 4(2)

ISSN

2638-6100

Authors

Hahn, Lewis D
Hall, Kent
Alebdi, Thamer
[et al.](#)

Publication Date

2022-03-01

DOI

10.1148/ryai.210162

Peer reviewed

Automated Deep Learning Analysis for Quality Improvement of CT Pulmonary Angiography

Lewis D. Hahn, MD • Kent Hall, MD • Thamer Alebdi, MD • Seth J. Kligerman, MD • Albert Hsiao, MD, PhD

From the Department of Radiology, University of California San Diego School of Medicine, 9300 Campus Point Dr, MC 0841, La Jolla, CA 92037-0841 (L.D.H., T.A., S.J.K., A.H.); and Naval Hospital Camp Pendleton, Oceanside, Calif (K.H.). Received June 15, 2021; revision requested July 27; revision received January 24, 2022; accepted February 3. Address correspondence to L.D.H. (e-mail: Lehahn@health.ucsd.edu).

Supported by UCSD Health Sciences Research Award RG100830.

Conflicts of interest are listed at the end of this article.

Radiology: Artificial Intelligence 2022; 4(2):e210162 • <https://doi.org/10.1148/ryai.210162> • Content codes: **AI** **CT** **VA**

CT pulmonary angiography (CTPA) is the first-line imaging test for evaluation of acute pulmonary emboli. However, diagnostic quality is heterogeneous across institutions and is frequently limited by suboptimal pulmonary artery (PA) contrast enhancement. In this retrospective study, a deep learning algorithm for measuring enhancement of the central PAs was developed and assessed for feasibility of its use in quality improvement of CTPA. In a convenience sample of 450 patients, automated measurement of CTPA enhancement showed high agreement with manual radiologist measurement ($r = 0.996$). Using a threshold of less than 250 HU for suboptimal enhancement, the sensitivity and specificity of the automated classification were 100% and 99.5%, respectively. The algorithm was further evaluated in a random sampling of 3195 CTPA examinations from January 2019 through May 2021. Beginning in January 2021, the scanning protocol was transitioned from bolus tracking to a timing bolus strategy. Automated analysis of these examinations showed that most suboptimal examinations following the change in protocol were performed using one scanner, highlighting the potential value of deep learning algorithms for quality improvement in the radiology department.

© RSNA, 2022

CT pulmonary angiography (CTPA) is the first-line imaging test for diagnosis of acute pulmonary embolism (PE). However, accurate diagnosis of PE with CTPA can be marred by patient and technical factors, including suboptimal contrast enhancement of the pulmonary arteries (PAs) (1). The degree of PA enhancement on CTPA images can be evaluated qualitatively through visual assessment or quantitatively by drawing a region of interest measurement in the main PA to measure the attenuation in Hounsfield units (2,3). Both methods may limit quality initiatives, however, because they are often performed nonsystematically and do not allow for routine evaluation of PA enhancement across an organization.

This study reports the accuracy of a deep learning method for assessment of central PA enhancement (CPAE) relative to manual measurement. The algorithm was used to analyze more than 3000 CTPA examinations performed since 2019 for quality improvement purposes. In particular, this method was used to assess CTPA quality during a change in our scanning protocol as we transitioned intravenous contrast timing from bolus tracking to a timing bolus strategy.

Materials and Methods

Dataset

This was a Health Insurance Portability and Accountability Act–compliant, institutional review board–approved study (protocol number 191797) with waiver of informed consent owing to use of retrospectively collected anonymized data. All data were obtained from the University of California San Diego picture archiving

and communication system. Inclusion criterion for the study was CTPA performed between 2016 and 2021 without exclusion based on age or indication. CT examinations were performed with CT scanners at two hospitals and one outpatient site, including three 64-section scanners (Discovery 750 HD; GE Healthcare), two 64-section scanners (Revolution EVO; GE Healthcare), one 256-section scanner (Revolution; GE Healthcare), and one 320-section scanner (Aquilion One; Canon Medical Systems USA). CTPA images were de-identified using Arterys software (version 26; Arterys) with the Digital Imaging and Communications in Medicine PS3.15 2016e specification (National Electrical Manufacturers Association).

Algorithm development dataset.— We obtained a convenience sample of 125 CTPA examinations for algorithm development and assigned 90 (34319 images) to the training set, 10 (6765 images) to the validation set, and 25 (9212 images) to the test set. The convenience sample was derived from a larger group of cardiovascular imaging CT studies that were readily accessible for deep learning algorithm development. The CT scans were manually inspected to ensure representation of a wide range of contrast opacification. The central (main, left, and right), interlobar, and left descending PAs were manually segmented. Ground truth segmentations were manually performed in ITK-SNAP software (version 3.6.0; www.itksnap.org) by two radiologists with 3 years' postresidency experience (K.H., L.D.H.) and a radiology resident (T.A.). All segmentations were reviewed by a fellowship-trained cardiothoracic radiologist (L.D.H.) for consistency.

Abbreviations

CPAE = central PA enhancement, CNN = convolutional neural network, CTPA = CT pulmonary angiography, DSC = Dice similarity coefficient, ICU = intensive care unit, PA = pulmonary artery, PE = pulmonary embolism

Summary

Deep learning algorithms may facilitate automation and ensure diagnostic quality of CT pulmonary angiograms.

Key Points

- Deep learning convolutional neural networks were used to accurately segment the central pulmonary arteries and measure central pulmonary artery enhancement on CT pulmonary angiograms.
- An automated method for identification of suboptimal contrast enhancement had a sensitivity of 100% and a specificity of 99.5% when compared with manual assessment as a reference standard.
- Automated central pulmonary enhancement measurement was used as a performance metric to identify targets for a CT pulmonary angiogram quality improvement initiative.

Keywords

CT Angiography, Pulmonary Arteries

Neural network development.— A two-dimensional convolutional neural network (CNN) modifying a U-Net style architecture (4) was developed by the lead author (L.D.H.) to perform segmentation of the central PA, interlobar or left descending PA, and background using the manual segmentations as ground truth labels. CNN input consisted of axial images at their original resolution (512×512). Voxel density was clipped and normalized to the range of -1200 HU to 3000 HU. The CNN was initialized with random initial weights. The optimizer used was RMSprop (<https://keras.io/api/optimizers/rmsprop/>), batch size was 32, and the loss function was sparse categorical cross-entropy. The algorithm was trained for 30 epochs with early stopping. No augmentation was used during training. The algorithm was implemented in Python (version 3.8.3; Python Software Foundation; <https://www.python.org/>) using Keras (version 2.4; <https://keras.io>) with TensorFlow (version 2.4; Google; <https://www.tensorflow.org>) backend. Dice similarity coefficient (DSC) was used to assess PA segmentation CNN performance on a per-study basis in the test set of 25 CTPA studies.

Assessment of central PA enhancement measurement accuracy.— In an independent set of 450 CTPA studies that were not used for algorithm development, a fellowship-trained cardiothoracic radiologist (L.D.H.) performed CPAE manual measurement by drawing an area of interest at the bifurcation of the main PA to serve as a reference standard. Automated CPAE was performed using the CNN by computing the average attenuation within the inferred central PA segmentation volume. The automated CPAE measurement was compared with manual measurement by computing the Pearson correlation coefficient (r). To assess the effectiveness of CNN-based classification of enhancement, a central PA attenuation threshold below 250 HU using a manual measurement was considered suboptimal (3).

Application of CNN for automated quality assessment.— By using the CNN, we retrospectively evaluated CPAE for 3195 randomly selected CTPA examinations performed between January 2019 and May 2021. In January 2021, a test bolus strategy was implemented as a quality improvement initiative, and CPAE of every examination was measured over the entire period to estimate the rate of suboptimal PA enhancement. Subgroup analyses were performed to assess variation in frequency of suboptimal enhancement based on scanner site, whether the CT was performed on the weekend versus performed on a weekday, and clinical setting. In August 2021, scan delay was increased by 2 seconds for a specific scanner with a large proportion of suboptimal examinations, and CPAE was further analyzed for 1 month.

Statistical analysis.— Statistical comparison of enhancement and rate of suboptimal examinations between bolus tracking and test bolus was performed using a two-sided unpaired Student t test and two-proportion Z test using Excel 2019 (Microsoft) and R Foundation (version 3.5.3; The R Project for Statistical Computing), respectively. A P value of less than .05 indicated a significant difference.

Data Availability

The code for the model used in this study has been made publicly available (<https://github.com/ldhahn/CPAsegmentation>).

Results

Algorithm Performance

Segmentation performance was good, with a mean \pm standard deviation DSC of 0.89 ± 0.04 for the validation set and 0.85 ± 0.14 for the test set. Visually, central PA segmentation generally correlated well with PA borders in the test set (Fig 1).

CPAE Measurement Accuracy

Manual and automated measurements of CPAE in 450 examinations correlated strongly ($r = 0.996$) (Fig 2A). The automated method demonstrated sensitivity of 100% (39 of 39) for suboptimal examinations and a true-negative rate of 99.5% (409 of 411) (Fig 2B). Total accuracy was 99.6%, with manual and automated measurements agreeing in 448 of 450 examinations. In the two inaccurately classified examinations, the automated measurements of 248 HU and 243 HU were slightly lower than the manual measurements of 258 HU and 268 HU, respectively.

Automated Quality Assessment Using CNN

Figure 3 shows CPAE for 3185 CTPA examinations performed between January 2019 and May 2021. Average CPAE using test bolus was lower than that for bolus-tracking studies (412 HU ± 147 vs 449 HU ± 150 , respectively; $P < .001$). The overall rate of suboptimal examinations was 7.1% (226 of 3195), with 6.8% (195 of 2880) and 9.8% (31 of 315) for suboptimal bolus-tracking and timing bolus examinations, respectively ($P = .06$). CPAE analysis revealed that suboptimal

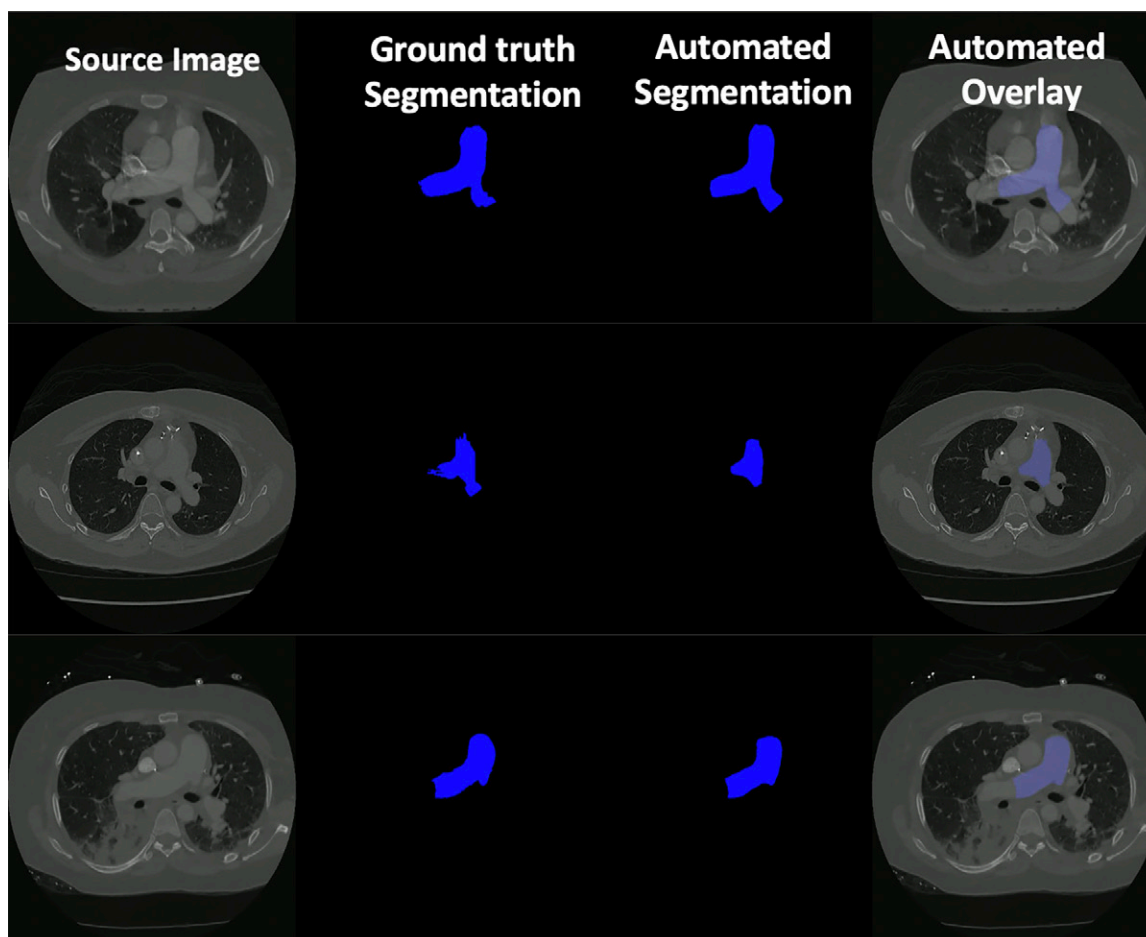
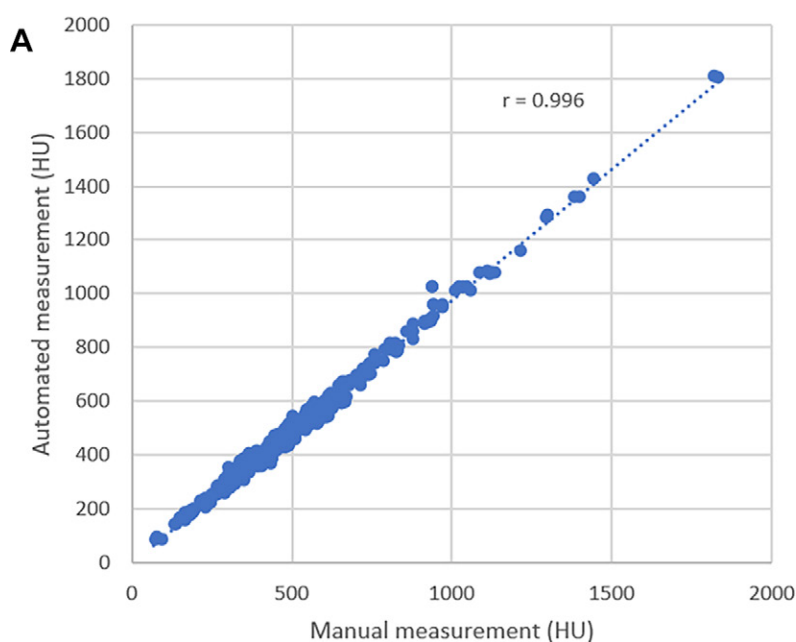


Figure 1: Example results of automated central pulmonary artery (PA) segmentation using a convolutional neural network. Automated segmentation of the PAs agreed well with ground truth segmentation in these examples with streak artifact (top row), poor PA contrast opacification (middle row), and lung parenchymal abnormalities (bottom row).



B

		Automated	
		<250 HU	≥250 HU
Manual	< 250 HU	39	0
	≥ 250 HU	2	409

Figure 2: (A) Automated measurement of central pulmonary artery enhancement (CPAE) strongly correlated with manual measurement ($r = 0.996$). (B) Confusion matrix for CPAE classification of optimal versus suboptimal using a threshold of 250 HU demonstrated excellent agreement between automated and manual measurements.

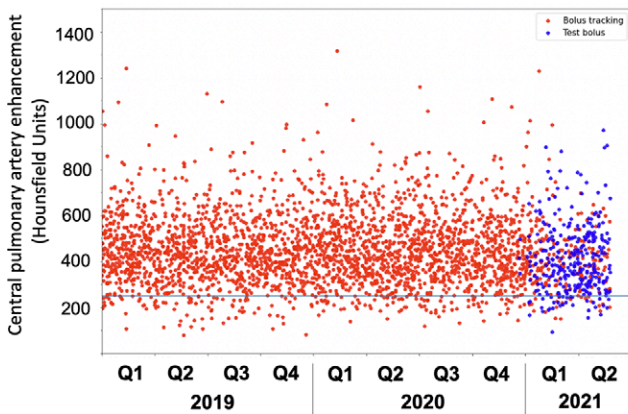


Figure 3: Automated measurement of central pulmonary artery enhancement in 3185 CT pulmonary angiography examinations performed between January 2019 and May 2021. These data were used to evaluate pulmonary artery enhancement during a transition from contrast timing using bolus tracking (red) to a test bolus (blue) beginning in January 2021.

test bolus examinations were predominantly localized to one scanner, which accounted for 32% (102 of 315) of all test bolus examinations but 65% (20 of 31) of suboptimal studies (Fig 4A). Furthermore, although only 28% (88 of 315) of all test bolus studies were performed on weekends, 42% (13 of 31) of suboptimal test bolus studies were performed on weekends (Fig 4B).

Subgroup Analyses

A hospital-based scanner was identified as being associated with the highest rate of nondiagnostic examinations, and the patient distribution was as follows: 79.4% in the emergency department (81 of 102), 7.8% in the intensive care unit (ICU) (eight of 102), 12.7% inpatients (13 of 102), and no outpatients. For the remaining scanners, the patient distribution was 47.9% in the emergency department (102 of 213), 11.7% in the ICU (25 of 213), 30.5% inpatients (65 of 213), and 9.9% outpatients (21 of 213). On weekdays, the distribution was 56.8% in the emergency department (129 of 227), 9.7% in the ICU (22 of 227), 24.2% non-ICU inpatients (55 of 227), and 9.3% outpatients (21 of 227). On weekends, the distribution was 61% in the emergency department (54 of 88), 13% in the ICU (11 of 88), 26% non-ICU inpatients (23 of 88), and no outpatients.

Analysis of CTPA studies from the scanner that generated a disproportionate number of suboptimal studies showed early contrast bolus timing in most examinations, and subsequently, scan delay was increased by 2 seconds. Prior to that intervention, 20 of 102 studies (19.6%) performed on this scanner were suboptimal. One month after intervention, 12 of 76 studies (15.8%) performed using that machine were suboptimal. The average CPAE increased from $378 \text{ HU} \pm 152$ to $390 \text{ HU} \pm 158$. However, there was no statistically significant difference in suboptimal examination rate or CPAE.

Discussion

This study used a deep learning algorithm to facilitate automated image-based analysis of CTPA quality by directly seg-

menting and measuring CPAE. Although we implemented the change in contrast timing strategy from bolus tracking to timing bolus to improve contrast opacification, we were surprised to find an initial decrease in average CPAE. Application of the deep learning algorithm allowed us to identify a specific hospital-based scanner with a disproportionate number of suboptimal examinations after this change. We increased the bolus timing scan delay, recognizing that early bolus timing was a contributing factor to suboptimal opacification. This resulted in a small, but not statistically significant, improvement in frequency of suboptimal enhancement.

Other factors, including patient complexity, also may have contributed to suboptimal opacification. The CT scanner in question was primarily used for patients evaluated in the emergency department and was used in the evaluation of fewer outpatients compared with other scanners, although the greater proportion of ICU patients examined using other scanners could have had an opposing effect. The difference in suboptimal examinations on weekends was likely partly related to absence of outpatient CTPA being performed.

Prior works have used automated and semiautomated methods for PA segmentation at CT. Many earlier works used rules-based computer vision methods, including fast marching level sets combined with active contours (5) and serial erosion and dilatation (6). In recent years, there has generally been a shift from rules-based methods to machine learning for segmentation of vascular structures (7). Machine learning methods may be more robust to variations in anatomy and disease compared with rules-based methods if the training set accounts for such variation. One group has previously used a deep learning method to segment the central PA, using a three-dimensional CNN (8). Although direct comparison is difficult because of differences in data sources, Román et al (8) observed performance similar to that of our method, achieving a mean DSC of 0.89.

An absolute CPAE threshold of 250 HU was used to determine adequate PA opacification, as commonly used in the literature (2,3,9). It is important to acknowledge the limitations of this metric. Early bolus or transient interruption of the contrast bolus may lead to adequate central PA opacification but poor opacification of segmental and subsegmental arteries. Moreover, in cases in which the central PAs demonstrate an average attenuation greater than 250 HU, it may still be possible to diagnostically evaluate subsegmental arteries, particularly if the central PAs are opacified at the tail end of the bolus. In fact, we subjectively found that most cases identified as suboptimal still had diagnostic contrast opacification. Nondiagnostic PE study rates vary from 3.3% to 6.6% (3,10,11) based on imaging report review. We suspect that our rate of suboptimal examination is higher likely owing to use of an absolute threshold for attenuation rather than radiologist-based assessment. Although this measure is imperfect, we find it useful for comparing general trends, as done in our study for contrast timing methods, scanners, and days of the week. Additionally, this was a single-institution study, and these results may not hold on datasets from other institutions. CPAE can also be affected by factors not evaluated in our study, including larger patient body habitus and kilovoltage peak (12,13), which we presumed was proportionately represented across our sample population.

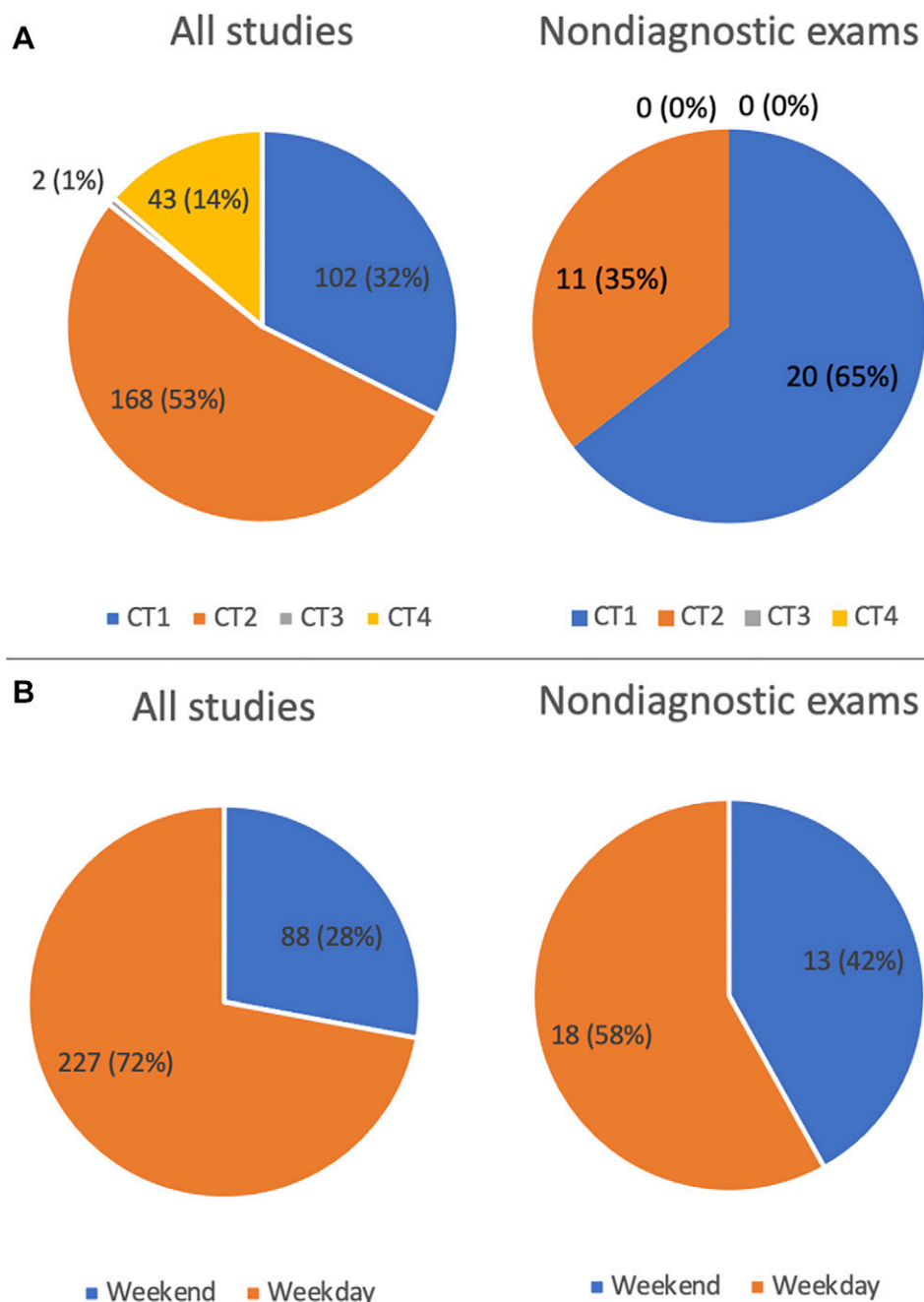


Figure 4: Subgroup analysis of CT pulmonary angiography performed using a test bolus for scan timing. **(A)** Analysis by scanner (ie, CT1, CT2, CT3, or CT4) demonstrated that CT performed using the CT1 scanner accounted for a disproportionate number of suboptimal examinations (exams). **(B)** Suboptimal examinations were performed more frequently on weekends relative to the overall proportion of CT examinations performed on weekends.

We believe this approach may be readily reproduced and valuable for automating quality monitoring and improvement across several examinations at multiple sites, which may be particularly helpful at large-volume centers with a variety of scanners, scanning protocols, and technologists. For practical application, this system could be applied to a set number of examinations performed each month, in parallel with standard workflow for clinical interpretation.

Beyond observing trends in contrast opacification for quality improvement, the automated analysis may eventually integrate with daily workflow for technologists, alerting them to possible

suboptimal examinations and prompting radiologist consultation. Because most examinations labeled as suboptimal were in fact diagnostic, it may be necessary to lower the 250 HU threshold to reduce alert fatigue in this setting. Additionally, the algorithm could be refined to address other reasons for nondiagnostic examinations, including motion artifact. Nonetheless, this PA segmentation algorithm may facilitate other applications, such as for measurement of PA size, as previously performed for other vascular structures (14).

In summary, we conducted a quality improvement analysis facilitated by a deep learning algorithm for CPAE measurement.

This method could be used in future studies to prospectively determine whether interventions result in improved CPAE.

Acknowledgments: We would like to thank Abraham Noorbakhsh, MD, MPH, and Brian Hurt, MD, MS, for their assistance with curating CT data used in this study.

Author contributions: Guarantors of integrity of entire study, **L.D.H., S.J.K., A.H.;** study concepts/study design or data acquisition or data analysis/interpretation, all authors; manuscript drafting or manuscript revision for important intellectual content, all authors; approval of final version of submitted manuscript, all authors; agrees to ensure any questions related to the work are appropriately resolved, all authors; literature research, **L.D.H., T.A., S.J.K., A.H.;** clinical studies, **L.D.H., K.H., T.A., S.J.K.;** experimental studies, **L.D.H., K.H., S.J.K.;** statistical analysis, **L.D.H.;** and manuscript editing, all authors

Disclosures of conflicts of interest: **L.D.H.** UCSF Health Sciences Research Award RG100830. **K.H.** No relevant relationships. **T.A.** No relevant relationships. **S.J.K.** Deputy editor for *Radiology: Cardiothoracic Imaging*. **A.H.** Research grants from GE Healthcare and Bayer to institution (unrelated to presented work); consulting fees paid to institution from Bayer (unrelated to presented work); co-founder shareholder in Arterys (unrelated to presented work).

References

1. Wittram C. How I do it: CT pulmonary angiography. *AJR Am J Roentgenol* 2007;188(5):1255–1261.
2. Bae KT, Tao C, Gürel S, et al. Effect of patient weight and scanning duration on contrast enhancement during pulmonary multidetector CT angiography. *Radiology* 2007;242(2):582–589.
3. Jones SE, Wittram C. The indeterminate CT pulmonary angiogram: imaging characteristics and patient clinical outcome. *Radiology* 2005;237(1):329–337.
4. Chollet F. Image segmentation with a U-Net-like architecture. https://keras.io/examples/vision/oxford_pets_image_segmentation/. Published 2020. Accessed February 1, 2021.
5. Linguraru MG, Pura JA, Van Uitert RL, et al. Segmentation and quantification of pulmonary artery for noninvasive CT assessment of sickle cell secondary pulmonary hypertension. *Med Phys* 2010;37(4):1522–1532.
6. Moses D, Sammut C, Zrimec T. Automatic segmentation and analysis of the main pulmonary artery on standard post-contrast CT studies using iterative erosion and dilation. *Int J Comput Assist Radiol Surg* 2016;11(3):381–395.
7. Zhao F, Chen Y, Hou Y, He X. Segmentation of blood vessels using rule-based and machine-learning-based methods: a review. *Multimedia Syst* 2019;25(2):109–118.
8. Román KL, de La Bruere I, Onieva J, et al. 3D Pulmonary Artery Segmentation from CTA Scans Using Deep Learning with Realistic Data Augmentation. *Image Anal Mov Organ Breast Thorac Images* (2018) 2018;11040:225–237.
9. Kim C, Lee CW, Hong GS, Kim G, Lee KY, Kim SS. Assessment of pulmonary arterial enhancement on CT pulmonary angiography using a leg vein for contrast media administration. *Medicine (Baltimore)* 2017;96(49):e9099.
10. Bates DD, Tkacz JN, LeBedis CA, Holalkere N. Suboptimal CT pulmonary angiography in the emergency department: a retrospective analysis of outcomes in a large academic medical center. *Emerg Radiol* 2016;23(6):603–607.
11. Lazarus MS, Kim Y, Mathai B, et al. Diagnostic Performance of Pulmonary Embolism Imaging in Patients with History of Asthma. *J Nucl Med* 2021;62(3):399–404.
12. Megyeri B, Christe A, Schindera ST, et al. Diagnostic confidence and image quality of CT pulmonary angiography at 100 kVp in overweight and obese patients. *Clin Radiol* 2015;70(1):54–61.
13. Chen EL, Ross JA, Grant C, et al. Improved Image Quality of Low-Dose CT Pulmonary Angiograms. *J Am Coll Radiol* 2017;14(5):648–653.
14. Hahn LD, Mistelbauer G, Higashigaito K, et al. CT-based True- and False-Lumen Segmentation in Type B Aortic Dissection Using Machine Learning. *Radiol Cardiothorac Imaging* 2020;2(3):e190179.