Correspondence

Performance evaluation of deep learning-based post-processing and diagnostic reporting system for coronary CT angiography: a clinical comparative study

Nan Luo1, Yi He1, Jitao Fan2, Ning Guo2, Guang Yang2, Yuanyuan Kong3, Jianyong Wei2, Tao Bi4, Jie Zhou4, Jiaxin Cao1, Xianjun Han1, Fang Li3, Shiyu Zhang1, Rujing Sun1, Zhaozhao Wang1, Tian Ma1, Lixue Xu1, Hui Chen6, Hongwei Li6, Zhenchang Wang1, Zhenghan Yang1

1Department of Radiology, Beijing Friendship Hospital of Capital Medical University, Beijing 100050, China; 2Shukun (Beijing) Technology Company Ltd., Beijing 100102, China; 3Statistics Section, Beijing Friendship Hospital of Capital Medical University, Beijing 100050, China; 4Department of Radiology, Beijing Anzhen Hospital of Capital Medical University, Beijing 100011, China; 5Department of Radiology, Beijing Chest Hospital of Capital Medical University, Beijing 100010, China; 6Department of Cardiology, Beijing Friendship Hospital of Capital Medical University, Beijing 100050, China.

To the Editor: Coronary computed tomography angiography (CCTA) has been increasingly widely performed for diagnosing coronary artery disease. Anatomical diagnosis, that is, stenosis grading, is still the main diagnostic index provided by most CCTA tests. Post-processing and interpretation of stenosis are 2 essential steps that need to be performed by cardiovascular imaging professionals from scan completion to diagnosis conclusion, which is repetitive and time-consuming, taking an average of 30 minutes each case in China and becoming the bottleneck of repetitive and time-consuming, taking an average of 30 minutes each case in China and becoming the bottleneck problem. We developed an AI system for automating post-processing and diagnostic reporting of CCTA data using deep learning algorithms to establish a new 1-click workflow for everyday use, namely, CCTA-AI (Figure 1). To further assess its capabilities, this study intends to answer 2 following questions: To what extent can it improve the efficiency of post-processing? To what extent can CCTA-AI detect and calculate coronary artery stenosis due to each atherosclerotic plaque? A total of 10,410 CCTA cases from 18 hospitals were collected and randomly split into 3 sets (training, tuning, and validation test sets) at a 7:2:1 ratio. In the training stage, each of the 7287 CCTA cases was annotated by a multi-seniority annotation group consisting of 3 layers of trained graders with ≥2 years, 5 years, and 10 years of experience in cardiovascular imaging diagnosis. The first layer of graders conducted initial quality control and excluded images with poor quality. The second layer independently annotated the aorta, coronary artery, and plaques. Finally, the third layer verified the annotations for each image. For the steps of the algorithm and the processing interface of CCTA-AI [Supplementary Figure 1, http://links.lww.com/CM9/A869].

In the clinical validation of this study, patients were eligible if they underwent CCTA and invasive coronary angiography (ICA) in sequence within 6 months from January 2017 to January 2019 in Beijing Friendship Hospital. A total of 335 patients were enrolled for the comparison of diagnostic accuracy. Another 350 independent cases were continuously collected for the comparison of time efficiency. All CCTA images were acquired with a GE Revolution 256-row multidetector CT scanner (GE Healthcare, Waukesha, Wisconsin, US), or a Philips Brilliance 128-row multidetector CT scanner (Philips Medical Systems, Eindhoven, The Netherlands), and a GE LightSpeed VCT 64-row multidetector CT scanner (GE Healthcare, Waukesha, Wisconsin, US). We used a prospective electrocardiography (ECG)-triggered CCTA technique with a 0.625 mm slice thickness and 0.625 mm intervals for all scanners. Non-ionic contrast media (iopromide 370 or iohexol 350) was continuously injected into the antecubital vein at 5 mL/s with a volume of 60 mL when the body weight was <100 kg or 6 mL/s with 80 mL when the body weight was ≥100 kg, followed by a 50 mL bolus of saline at 5 mL/s using a dual-head injector (Stellant; MEDRAD, Warrendale, Pennsylvania). ICA was performed using an Inova 2100 digital X-ray system (GE Healthcare, Waukesha) according to standard methods. At least 2 different views were obtained.
for each major vessel. Quantitative coronary angiography was performed using angiogram vendor-integrated quantitative coronary angiography software (GE Healthcare, Waukesha).

In the conventional CCTA path, 8 experienced radiographers and 4 radiologists routinely performed their respective tasks and recorded the time required for each case. In accuracy comparison, CCTA and ICA reading was accomplished by 2 groups of cardiovascular imaging professionals to evaluate lumen stenosis on CCTA and ICA, respectively. Agreeing results automatically becoming the final results and any specific disagreement was resolved by arbitration in CCTA or by quantitative measurement in ICA. The so-called final result in this study is the result that has been arbitrated, whereas the results obtaining from different readers were defined as the preliminary results. CCTA image quality was assessed on a per-segment basis on a 5-point Likert scale. Segments scored 1 were excluded from the following evaluation. The coronary arteries on CCTA were visually segmented and graded on a per-segment basis using the Society of Cardiovascular Computed Tomography classification with 18 segments to most and stenosis grading from 0 to 5.[2] Segments with a diameter <1.5 mm by CCTA were excluded. Segments with severe calcification were excluded by subjective visual assessment of the 2 preliminary readers or by arbitration. CCTA-AI reading was fulfilled without human interaction after CCTA data being transferred to the CCTA-AI platform directly after CT scanning [Supplementary Figure 2, http://links.lww.com/CM9/A869]. Paired t test was used to compare the post-processing time of the 2 paths. Diagnostic performance was compared between AI and professionals on a patient, vessel, and segment basis in terms of ≥50% and ≥70% stenosis separately. The sensitivity, specificity, negative predictive value (NPV), and positive predictive value of CCTA-AI were compared against those of the ICA. The area under the receiver operating characteristic curve (AUC) was calculated for both CCTA-AI and CCTA-professionals (both preliminary and final results were included) on a patient, vessel, and segment basis at the 50% and 70% threshold. Significance was set at the level of 0.05.

CCTA-AI showed a high completion rate in post-processing. Except for a case in which an anomalous origin of the coronary artery that had been unknown to the researchers before enrolment was not correctly recognized and analyzed by AI, the remaining cases were all post-processed by CCTA-AI. The average duration of manual post-processing was 16,011 seconds, while CCTA-AI took an average of 4,019 seconds (P < 0.0001) [Supplementary Figure 3, http://links.lww.com/CM9/A869]. In terms of diagnostic performance comparison, a total of 335 patients, with 1222 vessels and 3359 segments, were finally enrolled for comparative analysis. In the patient-based analysis, CCTA-AI correctly detected more stenosis lesions (89.3% and 72.4%) than the professional group (82.6% and 62.0%) at either 50% or 70% stenosis (P = 0.019, P = 0.038). However, the specificity was 55.9% at ≥50% and 64.5% at ≥70% stenosis by CCTA-AI, both lower than the 71.2% and 85.5% demonstrated by professionals (P = 0.006 and
The AUC of CCTA-AI was 0.76 (95% confidence interval, 0.70 to 0.81) for the diagnosis of a patient with at least 1 site of ≥50% coronary stenosis, which is inferior to the final results \( (P = 0.002) \) but equivalent to either of the preliminary results \( (P = 0.983) \). In detecting ≥70% stenosis, the AUC of CCTA-AI was roughly the same as its performance at ≥50%. In the vessel-based analysis, when compared to final results, the sensitivity of CCTA-AI was 74.6% for ≥50% and 54.7% for ≥70%, both of which were similar to that of the professionals \( (P = 0.061, 0.615) \). The specificity of CCTA-AI was 80.8% and 87.8%, both lower than that of the professionals, at 90.0% and 95.4% \( (P < 0.0001) \). The NPV of CCTA-AI was 88.8% and 89.1%, basically identical to the 87.8% and 89.5% of the professionals \( (P = 0.383 \) and \( P = 0.682) \). The AUC of CCTA-AI on the vessel basis was still marginally worse than the professional results. On a segment basis, the NPV of CCTA-AI was lower than that of professionals at ≥50% stenosis \( (P = 0.003) \) and similar to that of professionals at ≥70% stenosis \( (P = 0.074) \). The specificity at both ≥50% and ≥70% stenosis was somewhat lower than the professional results \( (P < 0.0001) \). The AUC of CCTA-AI showed underperformance compared to either the arbitrated or the unilateral results [Supplementary Figure 4, http://links.lww.com/CM9/A869 and Supplementary Tables 1 and 2, http://links.lww.com/CM9/A869].

There is no doubt that AI is more efficient than professionals in terms of time. In a country like China with a large population and patients, this is certainly meaningful. The improvement in time efficacy reduces the workload of radiologists, shortens the queue of patients, and facilitates the decision-making of physicians by providing results sooner. It is worth noting that data from 3 different CT vendors were adopted in the study and could be recognized and processed by the CCTA-AI system, which shows its robustness to some extent. The current data demonstrate the superior performance of CCTA-AI in detecting obstructive coronary artery stenosis at both the 50% and 70% thresholds. On a patient basis, the low specificity of CCTA-AI probably contributes to the high population incidence.\(^{[1-3]}\) A total of 66.9% of the included patients had ≥50% stenosis in at least 1 coronary artery, showing a high coronary disease burden. For at least the first phase of AI technologies, it would be valued more on sensitivity and NPV over specificity and positive predictive value by warranting non-significant lesions and leaving them to radiologists. The NPV of CCTA-AI at ≥50% stenosis was comparable to that of the professional group on a patient and vessel basis but was slightly lower on a segment basis. A possible reason might be the issue of differences in coronary artery segmentation between the algorithm and visual assessment. Although CCTA-AI’s segmentation is entirely based on the standard method, in actual execution, however, the naked eye does not fully follow predetermined rules when performing segmentation as the algorithm does. It would induce dislocation of lesions and therefore the misjudgment of segment-based stenosis.

CCTA-AI’s performance at patient basis reached the level of an attending physician. Nevertheless, the AUC of CCTA-AI showed its slightly inferior diagnostic performance compared to the final results on any basis. The unsatisfactory results may be related to the inter-evaluator difference. In our study, there was moderate consistency among professionals, even senior evaluators working in Grade-A Tertiary Hospitals with rich experience, indicating the real dilemma of visual interpretation in the current clinical procedure, that is, subjectivity in interpretation is inevitable. This implied that embedding CCTA-AI into the CCTA workflow may enable quantitative evaluation as a routine application. It has been expected for years to get more accurate measurements to minimizing inconsistencies between observers, whereas manual editing has always been required in previous attempts at software support.\(^{[4-6]}\)

Overall, the diagnostic performance of CCTA-AI in coronary stenosis still needs to be improved, so it is not yet ready to make an independent diagnosis. Due to its superior diagnostic sensitivity and equal performance compared to an attending radiologist, the CCTA-AI model can be considered a preliminary screening tool to optimize the clinical workflow.

**Funding**

This study has received funding from the National Key Research and Development Program of China (No. 2019YFE0107800) and the Beijing Municipal Science and Technology Commission (No. Z201100005620009).

**Conflicts of interest**

None.

**References**

1. Knuuti J, Wijns W, Saraste A, Capodanno D, Barbato E, Funck-Brentano C, et al. 2019 ESC guidelines for the diagnosis and management of chronic coronary syndromes. Eur Heart J 2020; 41:407–477. doi: 10.1093/eurheartj/ehz425.

2. Abbara S, Blanken P, Maroules CD, Cheezum M, Choi AD, Han BK, et al. SCCT guidelines for the performance and acquisition of coronary computed tomographic angiography: a report of the society of Cardiovascular Computed Tomography Guidelines Committee: endorsed by the North American Society for Cardiovascular Imaging (NASCI). J Cardiovasc Comput Tomogr 2016;10:435–449. doi: 10.1016/j.jcct.2016.10.002.

3. Wu Y, Benjamin EJ, MacMahon S. Prevention and control of cardiovascular disease in the rapidly changing economy of China. Circulation 2016;133:2345–2460. doi: 10.1161/CIRCULATIONAHA.115.008728.

4. Boogers MJ, Schuijf JD, Kietslaar PH, van Werkhoven JM, de Graaf FR, Boersma E, et al. Automated quantification of stenosis severity on 64-slice CT: a comparison with quantitative coronary angiography. JACC Cardiovasc Imaging 2010;3:699–709. doi: 10.1016/j.jcmg.2010.01.010.

5. Halpern EJ, Halpern DJ. Diagnosis of coronary stenosis with CT angiography. Acad Radiol 2011;18:324–333. doi: 10.1016/j.acra.2010.10.014.

6. Kang D, Slomka PJ, Nakazato R, Arsanjani R, Cheng VY, Min JK, et al. Automated knowledge-based detection of nonobstructive and obstructive arterial lesions from coronary CT angiography. Med Phys 2013;40:041912. doi: 1118/1.4794480.