JACC: CARDIOVASCULAR IMAGING, VOL. 13, NO. 2, 2020
FEBRUARY 2020:516–31

Evaluation of a Deep Learning–Based Automated CT Coronary Artery Calcium Scoring Algorithm

The measurement of coronary artery calcium scores (CACS) usually requires the manual input of a human operator to identify and mark calcified coronary lesions in each image section (1–3). Because this approach is labor intensive and time consuming, a more automated work flow is desirable to reduce the need for human interaction (4). The aim of this study was to evaluate novel deep learning–based software for automated coronary artery calcium scoring on noncontrast electrocardiographically triggered cardiac computed tomographic (CT) imaging.

We retrospectively included 511 consecutive patients (211 men, mean age 56.4 ± 10.2 years, mean body mass index 29.7 ± 8.3 kg/m^2, mean heart rate 69 ± 14 beats/min) with clinically indicated coronary calcium CT imaging between September 2014 and May 2018. Twenty-two patients showed mitral valve calcifications, and 41 patients had aortic valve calcifications. All data were acquired on a dual-source CT scanner (SOMATOM Force, Siemens Healthineers, Erlangen, Germany), and coronary artery calcium scoring was performed using a prospective electrocardiographically triggered noncontrast sequential acquisition. Coronary artery calcifications were manually selected in the coronary arteries by subspecialty-trained cardiac imagers during clinical routine and served as the reference standard. On the basis of the resulting Agatston score, patients were classified according to their CACS in standard risk categories (5). The deep learning–based automated calcium analysis was performed using a research prototype (Automated CaScoring, Siemens Healthineers). This application is based on a convolution neural network with a ResNet architecture for image features, as well as a fully connected neural network for spatial coordinate features. First, a deep-learning model is used to identify and segment the region belonging to the coronary artery and to ensure that calcium within the aorta, aortic valve, and mitral valve are not included in the CACS calculations. A list of potential calcium candidates with a threshold of >130 Hounsfield units is then implemented into the convolutional deep-learning model. In the final layer, the activations from the ResNet and the fully connected network are combined and fed to a single output neuron with a sigmoid activation function, giving the probability of the voxel's being coronary calcium (**Figure 1**). The algorithm was trained on 2,000 annotated datasets using the standard binary cross-entropy loss function and the Adam optimizer for adjusting model weights.

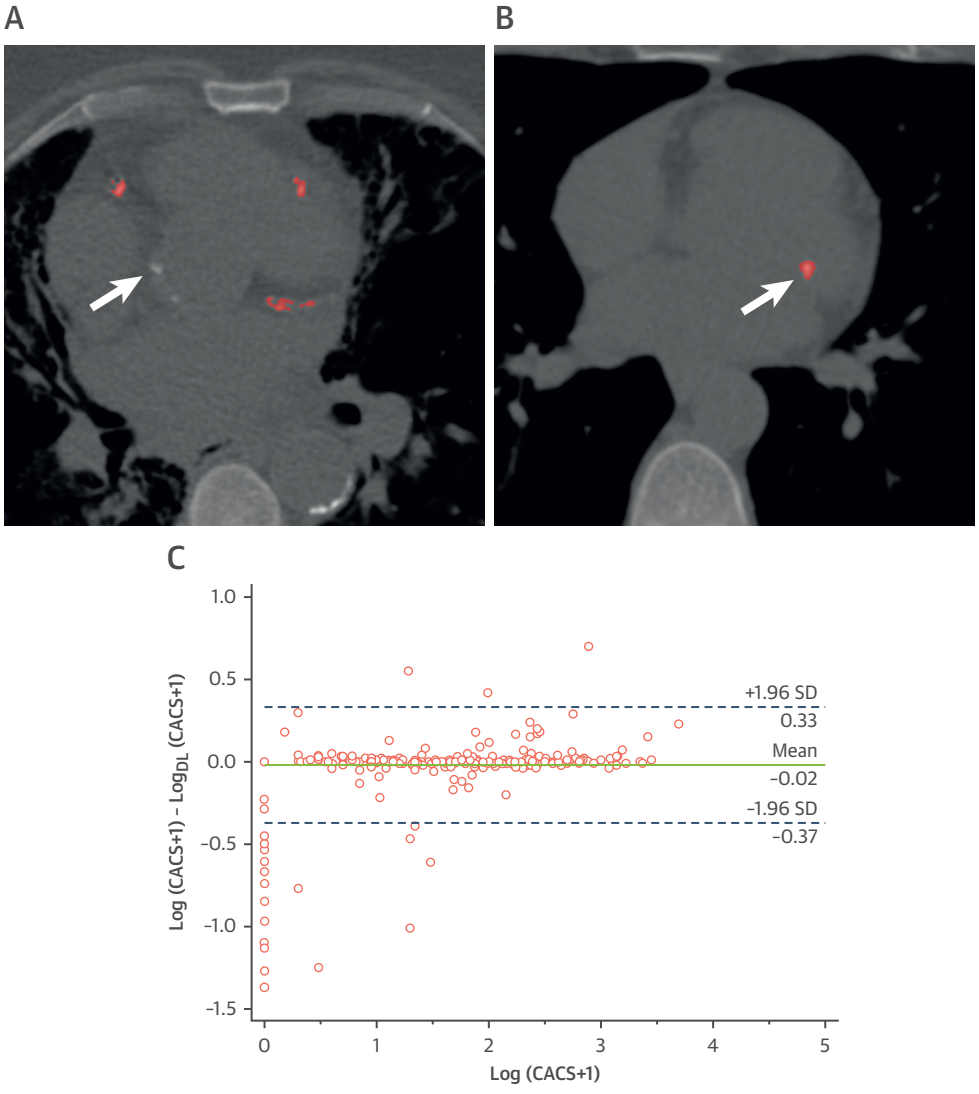
The median time needed for the deep-learning application to generate the final CACS value was 2.7 s in the independent-test group (n = 511). There were no significant differences for log(CACS + 1) values between the automated algorithm and the reference standard (p = 0.725). The coronary artery calcium scoring results of the automated algorithm correlated highly with the reference standard (Spearman's rho = 0.97, intraclass correlation coefficient = 0.985). In a patientwise analysis, the fully automated software classified 476 of 511 patients (93.2%) into the same risk category as the human observers. The Dice similarity coefficient for a CACS >0 was 0.95. The diagnostic accuracy for the different risk categories (0, 1 to 10, 11 to 100, 101 to 400, and >400) were 94.9% (258 of 272), 73.5% (36 of 49), 95.6% (87 of 91), 98.3% (59 of 60), and 92.3% (36 of 39), respectively. In addition, subgroup analysis revealed diagnostic accuracy of 94.3% for male patients and 92.3% for female patients, whereas the results for patients with body mass index greater and <30 kg/m^2 were 91.3% and 94.7%, respectively. Heart rates of ≤80 beats/min and >80 beats/min were associated with diagnostic performance of 92.4% and 96.5%, respectively.

The results of our study indicate that this application could be integrated into routine clinical use to quantify coronary calcium, provided there is a final quality check of deep-learning results by human observers. Nevertheless, substantial work-flow improvements by minimizing the need for human interaction can be anticipated by the clinical integration of artificial intelligence–based algorithms such as the one introduced here. Moreover, the adoption of such an algorithm could potentially reduce costs and make calcium scoring more widely available.

Limitations of our study were that many patients had CACS of 0, as we evaluated a dataset of consecutive patients, and that the current version of the prototype is not capable of branchwise calcium scoring, as a coronary labeling component is not yet implemented in the program.

In conclusion, deep learning–based automated CT coronary calcium scoring shows high accuracy, and the use of this fully automated software application may enhance work-flow efficiencies for this growing CT application.

FIGURE 1 Deep Learning–Based Calcium Scoring



(A) The deep-learning (DL) model for aorta segmentation ensured that calcium within the aorta is not included in the coronary artery calcium score calculations **(arrow)**. **(B)** In this case, the algorithm wrongly classified mitral valve calcification as coronary calcium **(arrow)**. **(C)** Bland-Altman plots of the reference standard and the DL-based calcium scoring method.

Simon S. Martin, MD
Marly van Assen, PhD
Saikiran Rapaka, PhD
H. Todd Hudson, Jr, MS
Andreas M. Fischer, MD
Akos Varga-Szemes, MD, PhD
Pooyan Sahbaee, PhD
Chris Schwemmer, MS
Mehmet A. Gulsun, PhD
Serkan Cimen, PhD
Puneet Sharma, PhD
Thomas J. Vogl, MD
U. Joseph Schoepf, MD*

*Medical University of South Carolina
Department of Radiology and Radiological Science
Division of Cardiovascular Imaging
25 Courtenay Drive
Charleston, South Carolina 29425
E-mail: schoepf@musc.edu

https://doi.org/10.1016/j.jcmg.2019.09.015

JACC: CARDIOVASCULAR IMAGING, VOL. 13, NO. 2, 2020
FEBRUARY 2020:516–31

Please note: Dr. Schoepf has received institutional research support and/or personal fees from Astellas, Bayer, Bracco, Elucid Bioimaging, Guerbet, HeartFlow, and Siemens. Dr. Varga-Szemes has received institutional research support and/or personal fees from Elucid Bioimaging, Guerbet, and Siemens. Dr. Rapaka, Dr. Sahbaee, Mr. Schwemmer, Dr. Gulsun, Dr. Cimen, and Dr. Sharma are employees of Siemens. All other authors have reported that they have no relationships relevant to the contents of this paper to disclose.

REFERENCES

1. Wolterink JM, Leiner T, de Vos BD, et al. An evaluation of automatic coronary artery calcium scoring methods with cardiac CT using the orCaScore framework. Med Phys 2016;43:2361.

2. Brunner G, Chittajallu DR, Kurkure U, Kakadiaris IA. Toward the automatic detection of coronary artery calcification in non-contrast computed tomography data. Int J Cardiovasc Imaging 2010;26:829–38.

3. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M Jr., Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll Cardiol 1990;15:827–32.

4. Wolterink JM, Leiner T, de Vos BD, van Hamersvelt RW, Viergever MA, Išgum I. Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks. Med Image Anal 2016;34:123–36.

5. Greenland P, Bonow RO, Brundage BH, et al. ACCF/AHA 2007 clinical expert consensus document on coronary artery calcium scoring by computed tomography in global cardiovascular risk assessment and in evaluation of patients with chest pain. J Am Coll Cardiol 2007;49:378–402.

Out of Focus: Increase of the Excess Longitudinal Range in Coronary Computed Tomographic Angiography

Coronary computed tomographic angiography (CTA) has emerged as an important noninvasive imaging method in cardiology, but exposure to potentially harmful radiation is of significant concern. The 2017 international PROTECTION (Prospective Multicenter Registry on Radiation Dose Estimates of Cardiac CT Angiography in Daily Practice) VI study demonstrated a 78% reduction of the median dose-length product associated with coronary CTA compared with a similar survey in 2007 (PROTECTION I) (1). Scan length directly contributes to the magnitude of radiation exposure (2), but awareness of reducing scan length in coronary CTA is currently unknown. Therefore, coronary computed tomographic angiographic studies from the PROTECTION I (n = 1,597) and PROTECTION VI (n = 3,724) studies were analyzed, and scan length parameters were determined centrally. The necessary coronary scan range was defined as the longitudinal extent of the coronary arteries plus cranial and caudal safety margins of 15 mm to accommodate differences in inspiratory holds between the localizing scout imaging and coronary CTA. Any scan length beyond these margins was defined as excess longitudinal scan range (**Figure 1A**). Variables are expressed as counts with percentages or as median (interquartile range [IQR]). Comparison of groups was performed using the Wilcoxon-Mann-Whitney *U* test or chi-square test. Univariate and multivariate linear regression analyses were performed using a generalized estimation equation model to account for the clustering effect of this multicenter trial.

The median necessary coronary scan range remained similar between 2007 and 2017 (120 mm [IQR: 115 to 127 mm] vs. 120 mm [IQR: 114 to 128 mm]; $p = 0.40$). In contrast, the median excess longitudinal scan range significantly increased by 75% from 8 mm (IQR: −3 to 23 mm) in 2007 to 14 mm (IQR: 4 to 26 mm) ($p < 0.001$) in 2017 (**Figure 1B**). The median excess longitudinal scan ranges differed significantly among study sites (ranging from −19 to 58 mm in the 2017 dose survey). The proportion of sites that routinely limit scan length to the defined required minimum (median excess longitudinal scan range ≤0 mm) decreased from 30% in 2007 to only 10% in 2017 ($p < 0.05$). Univariate linear regression analysis demonstrated that patient height and weight were not associated with excess longitudinal range. Furthermore, the level of clinical expertise, measured by the performance of coronary CTA in years or the volume of computed tomographic studies per month, was not an influencing factor. In contrast, there was a significant impact of scanner manufacturer on excess longitudinal scanning. Compared with the results with Toshiba scanners, excess longitudinal scan ranges were similar with Siemens (+2%; $p = 0.95$) and Philips (+29%; $p = 0.48$) but significantly increased with systems from GE (+146%; $p < 0.001$). In a multivariate linear regression analysis, the use of ultrawide coverage computed tomographic detectors exceeding 192 slices was associated with a 60% increase in excess longitudinal range ($p < 0.01$). Independent predictors associated with reduced excess longitudinal scanning were high-pitch helical scanning (−36%), preceding coronary artery calcium scanning (−28%), and low-tube potential imaging (−8% per 10 kVp) ($p < 0.05$ for all). Theoretically, limiting scan length to the defined required minimum in each study would have lowered the median dose-length product by 14% from 195 mGy · cm (IQR: 110 to 338 mGy · cm) to 167 mGy · cm (IQR: 94 to 290 mGy · cm) ($p < 0.001$). This corresponds to a reduction of effective dose from 5.1 to 4.3 mSv (conversion factor $k = 0.026$ mSv/mGy · cm).

The present study demonstrates that limiting scan length to a minimum is feasible and effective to lower radiation dose. However, the results also demonstrate that this strategy is largely