Using a deep learning algorithm to score coronary artery calcium in myocardial perfusion imaging: A real opportunity or just a new hype?

Jan Stassen, MD,[a,b] Pieter van der Bijl, MD, PhD,[a] and Jeroen J. Bax, MD, PhD[a,c,d]

[a] Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands
[b] Department of Cardiology, Jessa Hospital, Hasselt, Belgium
[c] Department of Cardiology, University of Turku and Turku University Hospital, Turku, Finland
[d] Department of Cardiology, Heart Lung Centre, Leiden University Medical Centre, Leiden, The Netherlands

Received May 5, 2022; accepted May 5, 2022
doi:10.1007/s12350-022-03009-1

See related article, pp. 239–250

Ischemic heart disease is still the leading cause of death worldwide.[1] Functional testing, including myocardial perfusion imaging with positron emission tomography (PET), remains the preferred method to evaluate coronary artery disease and is often recommended to risk-stratify and identify patients with ischemia before considering invasive coronary angiography.[2,3] However, with the increasing referral of patients at low to intermediate risk, the diagnostic and prognostic value of functional testing is declining.[4] Coronary artery calcium (CAC) scoring, a surrogate for coronary atherosclerotic plaque burden, may provide additional prognostic value over functional testing in these low to intermediate risk patients who are asymptomatic or present with new-onset chest pain.[5,6] In clinical practice, a PET scan for myocardial perfusion imaging is preceded by a non-ECG triggered low dose computed tomography (CT) scan for attenuation correction of the data. Previous studies have shown that CAC scoring is feasible on almost all diagnostic, non-contrast CT scans.[7] Assessment of CAC from these low dose CT scans could therefore provide additional information over myocardial perfusion imaging and may improve risk stratification in lower risk patients. Traditionally, the CAC score is calculated manually, using the Agatston scoring method.[8] However, this method is time consuming, preventing its widespread use in clinical practice. Visual scoring (scoring CAC by 'eye balling') and the use of artificial intelligence-based algorithms with deep learning abilities (scoring CAC 'automatically') are less time consuming and could therefore help to facilitate the more widespread introduction of CAC scoring after PET in clinical practice.[9,10] However, the accuracy of patients' CAC scoring according to manual, visual and automatic performance have not been compared previously in patients receiving a low-dose CT scan prior to myocardial perfusion imaging with PET.

In the current issue of the journal, Dobrolinska et al. evaluated 213 patients undergoing a low dose CT scan, prior to a 13N-ammonia PET scan for myocardial perfusion imaging.[11] All patients also received an ECG-triggered calcium scoring CT scan as the gold standard. Both low dose and calcium scoring CT scan were scored manually, visually and automatically in all patients. For visual scoring, a 6-points risk scale (according to the Agatston score) was used, whereas the automatic scoring was performed with deep-learning software. Manual scoring on the ECG-triggered, calcium scoring CT was used as the reference. The authors found a strong correlation between the automatic (weighted kappa 0.95, 95% CI 0.92-0.97) and visual (weighted kappa 0.88, 95% CI 0.85-0.92) scores on the cardiac scoring CT, and the reference technique. In addition, all three scoring methods (manual, visual and automatic) used in the low dose CT, correctly identified patients with CAC, as reflected by the high positive predictive values (100% for all). However, none of the scoring methods reliably excluded the presence of calcium, as reflected by the low

Reprint requests: Jan Stassen, MD, Department of Cardiology, Leiden University Medical Center, Leiden, The Netherlands; *J.Stassen@lumc.nl*
J Nucl Cardiol 2023;30:251–3.
1071-3581/$34.00

negative predictive values (57%, 31% and 63% for manual, automatic and visual methods, respectively). The agreement of manual (weighted kappa 0.59, 95% CI 0.53-0.65) and automatic (weighted kappa 0.50, 95% CI 0.44-0.56) scoring on the low dose CT with the reference technique was low, whereas agreement of the visual scoring on the low dose CT with the reference technique was much higher (weighted kappa 0.82; 95% CI 0.77-0.86).

The use of artificial intelligence with deep learning algorithms in the field of cardiovascular imaging is growing fast and it has the potential to be implemented in standard clinical practice. However, with this increased interest comes the risk of unrealistic expectations and subsequent disappointment, especially when this technology fails to deliver the expected results. The current study of Dobrolinska et al indeed showed that visual assessment outperformed manual, as well as automatic analysis of CAC scoring in low dose CT scans. However, instead of focusing on visual calcium scoring in low dose CT scans, clinicians should learn how to improve their knowledge of artificial intelligence algorithms and understand what these algorithms need for their successful implementation in different medical settings. Indeed, machine learning algorithms are not a magical solution, suitable for all clinical scenarios. Deep learning algorithms are crucially dependent on the raw input that is provided. The automatic method used in the current study was trained on ECG-gated CT scans, whereas non-gated CT scans were used for CAC scoring. The lack of ECG-triggering increases the amount of motion artifacts and decreases the accuracy to detect calcium in the coronary system, especially when the deep learning algorithm is not trained on this type of data.[12] This knowledge might at least partially explain the discrepancy that was seen between this study and others, which showed a better agreement between automated CAC scoring on low-dose CT scans and the gold standard (i.e., manual scoring on a calcium scoring CT).[13] Deep learning algorithms should also undergo thorough internal and external validation, similar to all other diagnostic algorithms, before being used in clinical practice. Furthermore, the implementation of deep learning algorithms is an ongoing process and training, as well as the continuous input of more general data, could certainly help to further improve their performance in different clinical settings. Despite the low agreement observed in the current study, the risk reclassification did not vary by more than one risk group in $>90\%$ of the patients and the high positive predictive value of the automatic method demonstrated a correct identification of patients with CAC. Taking the above-mentioned limitations into account, these findings highlight the potential value of deep learning algorithms when used in the correct setting.

Although PET myocardial perfusion imaging certainly has value for the risk stratification of patients with suspected coronary artery disease, this technique may underestimate the cardiovascular risk in patients with non-flow limiting atherosclerosis. The additional information from CAC scoring on a low dose CT scan could certainly help to improve the risk stratification and change the behavior of patients with cardiovascular risk factors who have a 'normal' functional test.[14] Due to the time consuming character of reading these scans manually, we can only hope that the use of deep learning algorithms will help to implement CAC scoring from low dose CT scans into clinical practice. Recognizing the appropriate setting and the current limitations of artificial intelligence-based algorithms, as well as having realistic expectations will certainly have an impact on whether these novel developments will have a long-lasting effect in this field of cardiovascular imaging.

Disclosures

The Department of Cardiology of Leiden University Medical Centre received research grants from Abbott Vascular, Bayer, Biotronik, Bioventrix, Boston Scientific, Edwards Lifesciences, GE Healthcare and Medtronic. Jeroen Bax received speaker fees from Abbott Vascular. Jan Stassen and Pieter van der Bijl have nothing to disclose.

References

1. Yusuf S, Joseph P, Rangarajan S, et al. Modifiable risk factors, cardiovascular disease, and mortality in 155 722 individuals from 21 high-income, middle-income, and low-income countries (PURE): A prospective cohort study. Lancet 2020;395:795-808.
2. Visseren FLJ, Mach F, Smulders YM, et al. 2021 ESC Guidelines on cardiovascular disease prevention in clinical practice. Eur Heart J 2021;42:3227-337.
3. Neumann FJ, Sousa-Uva M, Ahlsson A, et al. 2018 ESC/EACTS Guidelines on myocardial revascularization. Eur Heart J 2019;40:87-165.
4. Rozanski A, Gransar H, Hayes SW, et al. Temporal trends in the frequency of inducible myocardial ischemia during cardiac stress testing: 1991 to 2009. J Am Coll Cardiol 2013;61:1054-65.
5. Stone NJ, Robinson JG, Lichtenstein AH, et al. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: A report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. Circulation 2014;129:S1-45.
6. Mouden M, Timmer JR, Reiffers S, et al. Coronary artery calcium scoring to exclude flow-limiting coronary artery disease in symptomatic stable patients at low or intermediate risk. Radiology 2013;269:77-83.
7. Hecht HS, Cronin P, Blaha MJ, et al. 2016 SCCT/STR guidelines for coronary artery calcium scoring of noncontrast noncardiac chest CT scans: A report of the Society of Cardiovascular Computed Tomography and Society of Thoracic Radiology. J Cardiovasc Comput Tomogr 2017;11:74-84.

8. 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.
9. Einstein AJ, Johnson LL, Bokhari S, et al. Agreement of visual estimation of coronary artery calcium from low-dose CT attenuation correction scans in hybrid PET/CT and SPECT/CT with standard Agatston score. J Am Coll Cardiol 2010;56:1914-21.
10. Vonder M, Zheng S, Dorrius MD, et al. Deep Learning for Automatic Calcium Scoring in Population-Based Cardiovascular Screening. JACC Cardiovasc Imaging 2022;15:366-7.
11. Dobrolinska M,et al. Performance of visual, manual, and automatic coronary calcium scoring of cardiac 13N-ammonia PET/low dose CT. Journal of Nuclear Cardiology 2022. Online ahead of print.
12. Xie X, Greuter MJ, Groen JM, et al. Can nontriggered thoracic CT be used for coronary artery calcium scoring? A phantom study. Med Phys 2013;40:081915.
13. Išgum I, de Vos BD, Wolterink JM, et al. Automatic determination of cardiovascular risk by CT attenuation correction maps in Rb-82 PET/CT. J Nucl Cardiol 2018;25:2133-42.
14. Chang SM, Nabi F, Xu J, et al. The coronary artery calcium score and stress myocardial perfusion imaging provide independent and complementary prediction of cardiac risk. J Am Coll Cardiol 2009;54:1872-82.