

## Deep learning coronary artery calcium scores from SPECT/CT attenuation maps improves prediction of major adverse cardiac events

Running title: Deep learning of CAC on SPECT/CT

Robert JH Miller<sup>1,2\*</sup>, Konrad Pieszko<sup>1,3\*</sup>, Aakash Shanbhag<sup>1</sup>, Attila Feher<sup>4</sup>, Mark Lemley<sup>1</sup>,  
Aditya Killekar<sup>1</sup>, Paul B. Kavanagh<sup>1</sup>, Serge D. Van Kriekinge<sup>1</sup>, Joanna X. Liang<sup>1</sup>, Cathleen  
Huang<sup>1</sup>, Edward J Miller<sup>4</sup>, Timothy Bateman<sup>5</sup>, Daniel S. Berman<sup>1</sup>, Damini Dey<sup>1</sup>, Piotr J.  
Slomka<sup>1</sup>

- 1) Departments of Medicine (Division of Artificial Intelligence in Medicine), Imaging and Biomedical Sciences Cedars-Sinai Medical Center, Los Angeles, CA, USA
- 2) Department of Cardiac Sciences, University of Calgary, Calgary AB, Canada
- 3) Department of Interventional Cardiology and Cardiac Surgery, University of Zielona Gora, Poland
- 4) Section of Cardiovascular Medicine, Department of Internal Medicine, Yale University School of Medicine, New Haven, CT, USA
- 5) Cardiovascular Imaging Technologies LLC, Kansas City, MO, USA

\* Authors contributed equally

Word count: 4986

### Corresponding Author:

Piotr J. Slomka, PhD  
Cedars-Sinai Medical Center  
8700 Beverly Blvd, Suite Metro 203  
Los Angeles, CA 90048  
Email: [Piotr.Slomka@cshs.org](mailto:Piotr.Slomka@cshs.org)

### First Author:

Robert JH Miller MD  
University of Calgary, Department of Cardiac Sciences  
GAA08 HRIC, 3230 Hospital Drive NW  
Calgary, AB, T2N 4Z6  
Email: [Robert.miller@ahs.ca](mailto:Robert.miller@ahs.ca)

## **ABSTRACT**

### **Background**

Low-dose ungated CT attenuation correction (CTAC) scans are commonly obtained with SPECT/CT myocardial perfusion imaging. Despite characteristically low image quality of CTAC, deep learning (DL) can potentially quantify coronary artery calcium (CAC) from these scans in an automatic manner. We evaluated CAC quantification derived with a DL model including correlation with expert annotations and associations with major adverse cardiovascular events (MACE).

### **Methods**

We trained a convolutional long short-term memory DL model to automatically quantify CAC on CTAC scans using 6608 studies (2 centers) and evaluated the model in an external cohort of patients without known coronary artery disease (n=2271) obtained in a separate center. We assessed agreement between DL and expert annotated CAC scores. We also assessed associations between MACE (death, revascularization, myocardial infarction, or unstable angina) and CAC categories (0; 1-100; 101-400; >400) for scores manually derived by experienced readers and scores obtained fully automatically by DL using multivariable Cox models (adjusted for age, sex, past medical history, perfusion, and ejection fraction) and net reclassification index (NRI).

### **Results**

In the external testing population, DL CAC was 0 in 908(40.0%), 1-100 in 596(26.2%), 100-400 in 354(15.6%), and >400 in 413(18.2%) patients. Agreement in CAC category by DL CTAC and expert annotation was excellent (linear weighted Kappa 0.80), but DL CAC was obtained automatically in <2 seconds compared to ~2.5-minutes for expert CAC. DL CAC category was an independent risk for MACE with hazard ratios in comparison to CAC of zero: CAC 1-100 (2.20, 95% CI 1.54 – 3.14, p<0.001), CAC 101-400 (4.58, 95% CI 3.23– 6.48, p<0.001), and CAC > 400 (5.92, 95% CI 4.27 – 8.22, p<0.001). Overall NRI was 0.494 for DL CAC, which was similar to expert annotated CAC (0.503).

### **Conclusion**

DL CAC from SPECT/CT attenuation maps has good agreement with expert CAC annotations and provides similar risk stratification but can be obtained automatically. DL CAC scores improved classification of a significant proportion of patients as compared to SPECT myocardial perfusion alone.

## INTRODUCTION

SPECT myocardial perfusion imaging (MPI) is a well-established and widely utilized non-invasive imaging modality for the diagnosis and prognostication of coronary artery disease(CAD)(1,2). SPECT MPI is frequently obtained with an ungated, non-contrast computed tomography(CT) for attenuation correction(CTAC). SPECT/CT scanners use a common bed to move the patient sequentially through both scanners(3), with some models incorporating solid-state detector arrays. CTAC allows correction for soft-tissue attenuation artifacts, leading to improved diagnostic accuracy of SPECT MPI(4).

However, CTAC scans can also potentially be used to provide an anatomic assessment including evaluation of coronary artery calcium(CAC)(5). CAC scores are a well-established marker of the extent of coronary atherosclerosis(6-8). Integrating CAC scores from dedicated, gated CT scans with assessments of myocardial perfusion can be used to improve the diagnostic accuracy of SPECT(7) and PET MPI(9). Additionally, CAC from dedicated ECG-gated scans can provide incremental risk stratification when combined with SPECT MPI perfusion(10,11). However, CTAC scans are typically acquired with lower radiation doses and without cardiac gating, leading to worse image quality and often thicker slices compared to dedicated CAC scans, which may influence CAC scores(12). While it is possible to quantify CAC manually from CTAC scans, this can be time consuming and is not commonly performed. It is also possible to visually estimate CAC(13,14), but visual estimation is inherently subjective and requires experience to be performed accurately. Deep learning(DL) has been applied to image segmentation, including models for automated measures of CAC primarily from dedicated CAC scans. We developed a novel convolutional long short-term(ConvLSTM) DL model which integrates adjacent image slices, mimicking the clinical approach of scrolling between slices, to quantify CAC more

efficiently. We evaluated the correlation between DL and expert annotated CAC scores in patients undergoing SPECT/CT MPI. We then evaluated the prognostic significance of DL and expert annotated CAC scores for major adverse cardiovascular events (MACE), including incremental risk stratification over traditional SPECT MPI parameters, in an external population imaged with solid-state SPECT/CT MPI.

## **METHODS**

### **Study Population**

Patients who underwent SPECT/CT MPI with CTAC at one of two centers (Yale and Cardiovascular Imaging Technologies) were used to train the ConvLSTM model. Patients who underwent SPECT/CT MPI from a third center (University of Calgary) were used as an external testing cohort. Patients without CTAC were excluded. For external testing, patients with a history of CAD (n=673), defined as previous myocardial infarction or revascularization with either percutaneous coronary intervention or coronary artery bypass grafting (15), were excluded. Details of clinical data acquisition are available in the supplement. The study protocol complied with the Declaration of Helsinki. The study was approved by the institutional review board at all sites. To the extent allowed by data sharing agreements and institutional review board protocols, data and codes used in this manuscript will be shared upon written request.

### **Image Acquisition and Interpretation**

Details of MPI and CTAC image acquisition and interpretation are available in the supplement(16). Additional details of the training population are in **Supplemental Table 1**.

### Annotation of datasets for training, internal testing, and external testing

Two separate cohorts (each comprising of 10% of the total number available scans, n=661) were sampled out of the initial training cohort with equal number of cases in each CAC score category. One of those cohorts was held out as a validation set during training, and the model parameters were tuned to this set, while the second was held out for internal testing.

All the training, internal validation, internal testing, and external testing cases have been annotated on-site by two expert readers with at least 5 years of experience in CAC scoring using dedicated quantitative software (QPS Suite, Cedars Sinai Medical Center). DL annotations were processed using a custom-developed pipeline and both expert and DL annotated CAC scores were calculated according to the standard clinical algorithm(6), with additional details in the supplement.

The DL and expert annotated cases were categorized based on the CAC score (category 1: CAC score = 0, category 2: CAC score 1-100, category 3: CAC score 101-400, category 4: CAC score >400).

### Model Architecture

The model architecture is outlined in [Figure 1](#). The model was built using PyTorch version 3.7.4. We automatically segmented CAC from CTAC using a cascaded system of convLSTM(17). This system consists of two networks, first of which is trained for segmentation of the heart silhouette and the second network was trained to segment the CAC. The heart convLSTM was trained on a subset of training data with expert reader annotations from QFAT software(18). A supervised learning regime was used for both segmentation networks. The heart mask was applied to final CAC prediction to reduce any spurious bone overcalling or calcification in non-cardiac regions. To imitate physician approach of aggregating information from adjacent slices, three

slices were provided to both networks as input(19). This network architecture was shown previously to have significantly reduced memory consumption and almost 2x faster inference times, with similar accuracy, compared to U-Net on a typical CPU (17). Case examples with expert and DL annotations are shown in Figure 2.

### Statistical Analysis

Details in supplement (20,21). The Proposed Requirements for Cardiovascular Imaging-Related Machine Learning Evaluation (PRIME)(22) checklist is shown in Supplemental Table 2. Improvement in likelihood ratio chi-square and area under the receiver operating characteristic curve were also assessed.

## RESULTS

### Population Characteristics

In total, 6608 patients were included in the training population. The external testing population included 2271 patients, with population characteristics by DL CAC category shown in Table 1. Based on the DL model results, CAC was 0 in 908 (40.0%) patients, CAC 1-100 in 596 (26.2%) patients, CAC 100-400 in 354 (15.6%) patients, and >400 in 413 (18.2%) patients.

### DL vs Expert Annotated CAC

DL-CTAC were obtained fully automatically in less than 2 seconds per scan (time required to load the study, select slices, and annotate lesions for the entire CTAC volume). This compares to approximately 5 minutes for expert annotations which includes time required to load the study, review all slices, and annotate lesions on selected slices. Figure 3 outlines concordance between DL CAC and expert annotation CAC categories. The category-wise agreement (Figure 3) between DL CAC and expert CAC was excellent (linear weighted Kappa 0.80). There was also good pair-

wise correlation between DL CAC and expert annotated CAC as continuous measures ( $r^2 = 0.693$ ,  $p < 0.001$ , [Supplemental Figure 1](#)). Summary of categorization by visual CAC estimation compared to DL and expert annotated CAC is shown in Supplemental Table 3. Review of discrepant cases are shown in the supplemental results and [Supplemental Figure 2](#).

#### Associations with MACE

During median follow-up of 2.8 (IQR 1.7 – 4.1) years, 320 patients experienced at least one MACE. [Supplemental Table 4](#) outlines patient characteristics in patients who experienced MACE compared to those who did not. Patients who experienced MACE had higher median CAC (178 vs. 11,  $p < 0.001$ ) and were more likely to have CAC > 400 (35.9% vs. 15.3%,  $p < 0.001$ ). Patients who experienced MACE were also older (median 70.7 vs 66.1,  $p < 0.001$ ) and more likely to have a history of diabetes (31.6% vs 22.1%,  $p < 0.001$ ) in addition to higher rates of other cardiovascular risk factors.

Increasing DL CAC and expert CAC category were associated with an increased risk of MACE ([Figure 4](#)). Compared to patients with DL CAC of 0, patients with scores 1-100 (unadjusted hazard ratio [HR] 2.20, 95% CI 1.54 – 3.14), CAC 101-400 (unadjusted HR 4.58, 95% CI 3.23 – 6.48), and CAC > 400 (5.92, 95% CI 4.27 – 8.22) were at significantly increased risk of MACE. The risk was similar across categories of expert annotated CAC categories. Kaplan-Meier survival curves stratified by visually estimated CAC are shown in [Supplemental Figure 3](#).

Associations with MACE in the multivariable model are outlined in [Table 2](#). DL CAC category continued to be associated with increased risk of MACE in adjusted analyses for patients with CAC 1-100 (adjusted HR 1.90, 95% CI 1.32 – 2.73,  $p < 0.001$ ), CAC 101-400 (adjusted HR 3.32, 95% CI 2.29 – 4.81,  $p < 0.001$ ), and CAC > 400 (adjusted HR 3.58, 95% CI 2.47 – 5.19,  $p < 0.001$ ) compared to CAC 0. This risk stratification was similar to the risk associated with mild

stress perfusion abnormalities (stress total perfusion deficit [TPD] 5-10%, adjusted HR 1.70, 95% CI 1.19 – 2.44,  $p=0.004$ ) and moderate to severe stress perfusion abnormalities (stress TPD > 10%, adjusted HR 4.73, 95% CI 3.02 – 7.46,  $p<0.001$ ). The risk associated with expert annotated CAC categories was similar to DL categories (CAC 1-100 adjusted HR 2.20, 95% CI 1.52 – 3.19,  $p<0.001$ ; CAC 101-400 adjusted HR 3.57, 95% CI 2.45 – 5.20,  $p<0.001$ ; and CAC > 400 adjusted HR 4.05, 95% CI 2.78 – 5.90,  $p<0.001$ ).

Associations with primary outcome were similar if patients who underwent early revascularization were excluded (DL CAC 1-100: adjusted HR 2.00, 95% CI 1.34 – 2.94,  $p=0.001$ ; DL CAC 101-400: adjusted HR 2.98, 95% CI 1.97 – 4.49,  $p<0.001$ ; DL CAC > 400: adjusted HR 3.07, 95% CI 2.03 – 4.66,  $p<0.001$ ). Results were also similar for associations with death or myocardial infarction as well as associations with death ([Supplemental Table 5](#)).

#### Net Reclassification

We assessed patient reclassification with the addition of CAC categories to all other components of the multivariable model outlined in [Table 2](#). The results of the NRI analysis are shown in [Figure 5](#). Both DL CAC and expert annotated CAC significantly improved model fit and AUC (all  $p<0.01$ ), with additional details in [Supplemental Table 6](#). DL CAC categories improved risk-classification of patients with events (event NRI 0.230, 95% CI 0.142 to 0.314), patients without events (non-event NRI 0.264, 95% CI 0.204 to 0.309), and overall patient classification (overall NRI 0.494, 95% CI 0.363 to 0.607). Event, non-event, and overall NRI was similar for both measures as shown in [Supplemental Table 6](#). Additionally, overall NRI was lower for visually estimated CAC (overall NRI 0.409, 95% CI 0.278 to 0.537).



## DISCUSSION

We demonstrated that DL derived CAC scores from CTAC imaging could be used to stratify risk of MACE, with scores derived rapidly (~2 seconds) in a completely automated manner. There was good agreement between CAC score categorization by DL and expert annotations as evaluated in a large external population with different characteristics. Lastly, we demonstrated that DL CAC categories provided additional prognostic information over clinical information and quantitative assessment of perfusion and ventricular function, with improved classification of a quarter of patients who experienced MACE and a quarter of patients who did not experience MACE. DL CAC scores from CTAC could be utilized clinically to significantly improve risk stratification in patients undergoing SPECT/CT MPI without the need for physician or technician time for manual annotation.

We demonstrated that the convLSTM model was able to quantify CAC from CTAC imaging, with excellent agreement with and similar risk stratification to expert annotated CAC. Importantly, the model was trained with data from two sites which have different CTAC imaging protocols compared to the external testing site. This has not been commonly performed in the existing literature and provides evidence that the convLSTM and associated DL CAC scores should be generalizable to a variety of acquisition protocols. We also previously demonstrated that this approach has faster inference times compared to a U-net model, and therefore should not negatively impact clinical workflow. Another major strength of the current study is the large number of expert annotations performed on CTAC scans, which are not typically performed clinically. This allowed us to evaluate agreement with expert CAC scores more precisely and robustly compare the risk stratification provided by the two measures, including their improvements for risk prediction of traditional SPECT/CT variables.

Several other approaches to CAC scoring with artificial intelligence have been applied previously (23-27). The agreement between CAC categories in our study (Cohen's kappa 0.80) is similar to agreement demonstrated using dedicated ECG-gated scans with other DL approaches(25). Isgum et al. developed a convolutional neural network which quantified CAC from low-dose CT scans obtained for lung cancer screening(26). When the same model was applied in patients undergoing PET MPI, the agreement between manual and automated scoring in CTAC was lower compared to the present study (linear weighted kappa 0.70 to 0.74) and was tested in a much smaller patient population (n=133)(28). Sartoretti et al. also demonstrated good agreement between expert annotated and DL CAC scores in a cohort of 56 patients undergoing SPECT/CT MPI(29). Importantly, these methods demonstrate similar rates of agreement as would be expected between two expert readers scoring CAC from low-dose CT scans(30). High noise levels and partial volume effects impact the appearance of CAC lesions(12), leading to frequent false negative physician interpretations as evidenced by our finding that physician interpretation of the presence or absence of calcium was discrepant in ~10% of patients. Additionally, we identified cases where DL annotations differed with expert annotations for calcium in coronary ostia vs. adjacent aorta and for valvular calcification vs. adjacent coronary arteries.

While agreement between DL and expert CAC categories is important by itself, we demonstrated that significant improvements in risk stratification are possible with DL annotated CAC scores. We demonstrated that increasing DL CAC category was associated with an increased risk of MACE, similar to recent findings from Zeleznik et al. in both symptomatic and asymptomatic populations(27). However, in the present study we demonstrated that the risk associated with each category was similar to the corresponding category of expert annotated

CAC. Additionally, both DL and expert reader CAC categories were significantly associated with MACE after correcting for relevant confounders including age, sex, medical history, and SPECT MPI results. Lastly, we demonstrated that improvement in patient risk classification with DL CAC was similar to that achieved by expert annotated categories of CAC, with both being higher than is possible with subjective expert visual estimates. Importantly, visually estimated CAC was performed at the time of clinical reporting and was informed by clinical history and perfusion findings. Improved classification compared to expert visual estimate is particularly relevant since nuclear cardiology laboratories more frequently rely on this method for CAC classification given the time required for expert annotation. While Dekker et al. found that DL CAC scores had a NRI of 0.13 in patients undergoing PET MPI (31), in our study ~1 in 4 patients who experienced MACE would have their risk correctly reclassified, with a similar proportion of patients who did not experience MACE correctly re-classified. Therefore, this approach could be applied to automatically improve risk classification in a substantial proportion of patients.

Our work adds to a growing body of literature supporting integration of CAC scores when interpreting MPI. Chang et al. demonstrated that quantitative CAC combined with SPECT-MPI findings provided independent and complementary prognostic information among a cohort of 1,126 patients without prior CAD(11). Engbers et al. evaluated combined Agatston CAC score and SPECT-MPI in 4,897 symptomatic patients without prior CAD(10), demonstrating a stepwise increase in MACE with increasing CAC score among patients with both normal and abnormal perfusion. Visually estimated CAC(13) can also provide risk stratification in patients undergoing SPECT/CT MPI (14). However, in the present work we demonstrate that the

improvement in risk classification is higher with DL CAC, which are rapidly and automatically derived from SPECT/CT attenuation maps.

Our study has a few important limitations. CT attenuation imaging was utilized clinically to visually assess coronary calcification, and knowledge of CAC can influence patient management(32). However, results were similar for associations with hard outcomes, and this would be expected to, if anything, decrease the associations between CAC and hard outcomes. Additionally, it is unknown whether associations with outcomes would differ significantly between CAC from dedicated gated studies compared to CAC; however, previous studies have demonstrated close agreement between the measures(33). We trained the ConvLSTM model using scans with different acquisition parameters compared to the external testing population. More precise quantification of CAC may be possible if the model was trained with similar data, but this also suggests that the model should be broadly generalizable. The model was trained to differentiate coronary from non-coronary calcifications using expert annotations. However, some lesions are challenging for expert readers to annotate (such as ostial calcium compared to adjacent aortic calcifications) and the DL model would also be expected to have difficulties with these areas. While the DL method provides fully automated results, they will still need to be verified by a physician. The training population included patients with previous revascularization; however, we excluded patients with known CAD from the external testing population and dedicated studies are needed to evaluate the model's ability to differentiate CAC from coronary stents. Lastly, we are not able to ascertain cardiovascular mortality in this large, retrospective population.

## **CONCLUSION**

DL CAC derived from SPECT/CT attenuation maps has good agreement with expert CAC annotations. DL and expert annotated CAC are associated with MACE, but DL scores can be obtained automatically in a few seconds. DL CAC scores can be quantified automatically following SPECT/CT MPI, without impeding clinical workflow, in order to improve classification of a significant proportion of patients.

## **SOURCE OF FUNDING**

This research was supported in part by grant R01HL089765 from the National Heart, Lung, and Blood Institute/ National Institutes of Health (NHLBI/NIH) (PI: Piotr Slomka). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

## **DISCLOSURES**

Dr. Robert Miller has received consulting fees and research support from Pfizer. Drs. Berman and Slomka and Mr. Kavanagh participate in software royalties for QPS software at Cedars-Sinai Medical Center. Dr. Slomka has received research grant support from Siemens Medical Systems. Dr. Berman, and Dr. Edward Miller have served as consultants for GE Healthcare. Remaining authors have no relevant disclosures.

## **KEY POINTS**

**QUESTION:** Do coronary artery calcium scores quantified automatically with a deep learning model provide similar risk stratification to expert annotated scores?

**PERTINENT FINDINGS:** In this retrospective multicenter study, with dedicated training and external testing populations, DL CAC scores good agreement with expert annotated scores. DL and expert annotated CAC are associated with MACE, but DL scores can be obtained automatically in a few seconds.

**IMPLICATIONS FOR PATIENT CARE:** DL CAC scores could be used to improve risk prediction of a significant proportion of patients, without impeding clinical workflow.

## REFERENCES

1. Berman DS, Hachamovitch R, Kiat H, et al. Incremental value of prognostic testing in patients with known or suspected ischemic heart disease. *J Am Coll Cardiol*. 1995;26:639-647.
2. Fihn SD, Gardin JM, Abrams J, et al. 2012 ACCF/AHA/ACP/AATS/PCNA/SCAI/STS Guideline for the diagnosis and management of patients with stable ischemic heart disease. *J Am Coll Cardiol*. 2012;60:e44-e164.
3. Dorbala S, Di Carli MF, Delbeke D, et al. SNMMI/ASNC/SCCT Guideline for cardiac SPECT/CT and PET/CT 1.0. *J Nucl Med*. 2013;54:1485-1507.
4. Huang JY, Huang CK, Yen RF, et al. Diagnostic performance of attenuation-corrected myocardial perfusion imaging for coronary artery disease. *J Nucl Med*. 2016;57:1893-1898.
5. Patchett ND, Pawar S, Miller EJ. Visual identification of coronary calcifications on attenuation correction CT improves diagnostic accuracy of SPECT/CT myocardial perfusion imaging. *J Nucl Cardiol*. 2017;24:711-720.
6. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R. Quantification of coronary artery calcium using ultrafast computed tomography. *J Am Coll Cardiol* 1990;15:827-832.
7. Schepis T, Gaemperli O, Koepfli P, et al. Added value of coronary artery calcium score as an adjunct to gated SPECT for the evaluation of coronary artery disease in an intermediate-risk population. *J Nucl Med*. 2007;48:1424-1430.
8. Blaha MJ, Blankstein R, Nasir K. Coronary artery calcium scores of zero and establishing the concept of negative risk factors. *J Am Coll Cardiol*. 2019;74:12-14.
9. Brodov Y, Gransar H, Dey D, et al. Combined quantitative assessment of myocardial perfusion and coronary artery calcium score by hybrid 82Rb PET/CT improves detection of coronary artery disease. *J Nucl Med*. 2015;56:1345-1350.
10. Engbers EM, Timmer JR, Ottervanger JP, Mouden M, Knollema S, Jager PL. Prognostic value of coronary artery calcium scoring in addition to in symptomatic patients. *Circ Cardiovasc Imaging*. 2016;9:e003966.
11. 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;54:1872-1882.
12. Muhlenbruch G, Thomas C, Wildberger JE, et al. Effect of varying slice thickness on coronary calcium scoring with multislice computed tomography in vitro and in vivo. *Invest Radiol*. 2005;40:695-699.
13. 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;56:1914-1921.
14. Trpkov C, Savtchenko A, Liang Z, et al. Visually estimated coronary artery calcium score improves SPECT-MPI risk stratification. *Int J Cardiol Heart Vasc*. 2021;35:100827.

15. Miller RJH, Klein E, Gransar H, et al. Prognostic significance of previous myocardial infarction and previous revascularization in patients undergoing SPECT MPI. *Int J Cardiol.* 2020;313:9-15.
16. Slomka PJ, Nishina H, Berman DS, et al. Automated quantification of myocardial perfusion SPECT using simplified normal limits. *J Nucl Cardiol.* 2005;12:66-77.
17. Pieszko K, Shanbhag A, Killekar A, et al. Calcium scoring in low-dose ungated chest CT scans using convolutional long-short term memory networks. *SPIE Med Imaging.* 2022;120323:1-12.
18. Eisenberg E, McElhinney PA, Commandeur F, et al. Deep learning-based quantification of epicardial adipose tissue volume and attenuation predicts major adverse cardiovascular events in asymptomatic subjects. *Circ Cardiovasc Imaging.* 2020;13:e009829.
19. Lin TY, Goyal P, Girshick R, He K, Dollar P. Focal loss for dense object detection. *IEEE Trans Pattern Anal Mach Intell.* 2020;42:318-327.
20. Miller RJH, Bonow RO, Gransar H, et al. Percutaneous or surgical revascularization is associated with survival benefit in stable coronary artery disease. *Eur Heart J Cardiovasc Imaging.* 2020;21:961-970.
21. Azadani PN, Miller RJH, Sharir T, et al. Impact of early revascularization on major adverse cardiovascular events in relation to automatically quantified ischemia. *JACC Cardiovasc Imaging.* 2021;14:644-653.
22. Sengupta PP, Shrestha S, Berthon B, et al. Proposed requirements for cardiovascular imaging-related machine learning evaluation (PRIME). *JACC Cardiovasc Imaging.* 2020;13:2017-2035.
23. Kurkure U, Chittajallu DR, Brunner G, Le YH, Kakadiaris IA. A supervised classification-based method for coronary calcium detection in non-contrast CT. *Int J Cardiovasc Imaging.* 2010;26:817-828.
24. Isgum I, Rutten A, Prokop M, van Ginneken B. Detection of coronary calcifications from computed tomography scans for automated risk assessment of coronary artery disease. *Med Phys.* 2007;34:1450-1461.
25. 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.
26. Isgum I, Prokop M, Niemeijer M, Viergever MA, van Ginneken B. Automatic coronary calcium scoring in low-dose chest computed tomography. *IEEE Trans Med Imaging.* 2012;31:2322-2334.
27. Zeleznik R, Foldyna B, Eslami P, et al. Deep convolutional neural networks to predict cardiovascular risk from computed tomography. *Nat Commun.* 2021;12:715.
28. Isgum 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;25:2133-2142.
29. Sartoretti T, Gennari AG, Sartoretti E, et al. Fully automated deep learning powered calcium scoring in patients undergoing MPI. *J Nucl Cardiol.* 2022:Epub ahead of print.
30. Jacobs PC, Isgum I, Gondrie MJ, et al. Coronary artery calcification scoring in low-dose ungated CT screening for lung cancer *Am J Roentgenol.* 2010;194:1244-1249.



- 31.** Dekker M, Waissi F, Bank IEM, et al. The prognostic value of automated coronary calcium derived by a deep learning approach on non-ECG gated CT images from (82)Rb-PET/CT MPI. *Int J Cardiol.* 2021;329:9-15.
- 32.** Rozanski A, Gransar H, Shaw Leslee J, et al. Impact of coronary artery calcium scanning on coronary risk factors and downstream testing. *J Am Coll Cardiol.* 2011;57:1622-1632.
- 33.** Pieszko K, Shanbhag AD, Lemley M, et al. Reproducibility of quantitative coronary calcium scoring from PET/CT attenuation maps: comparison to ECG-gated CT scans. *Eur J Nucl Med Mol Imaging.* 2022:Epub ahead of print.

## FIGURES

### Model Architecture

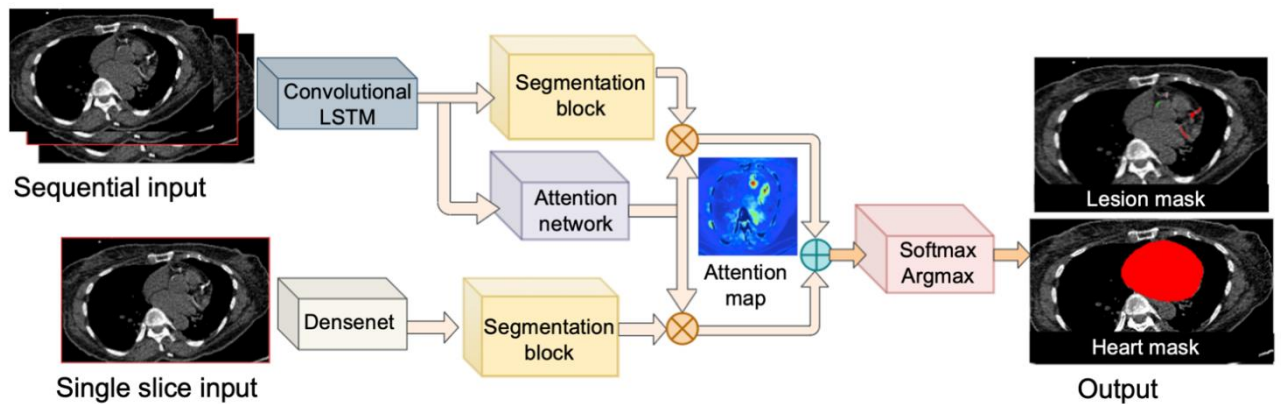


Figure 1: Outline of model architecture. The convolutional long short-term memory (ConvLSTM) includes a network trained to segment coronary artery calcium (CAC) and a second network for segmentation of the heart which limits CAC scoring. The softmax argmax function normalizes output of the network to expected probabilities. The model identifies coronary calcium (red) and non-coronary calcium (green) within the heart mask.

## Deep Learning Coronary Artery Calcium Scores for Attenuation Imaging

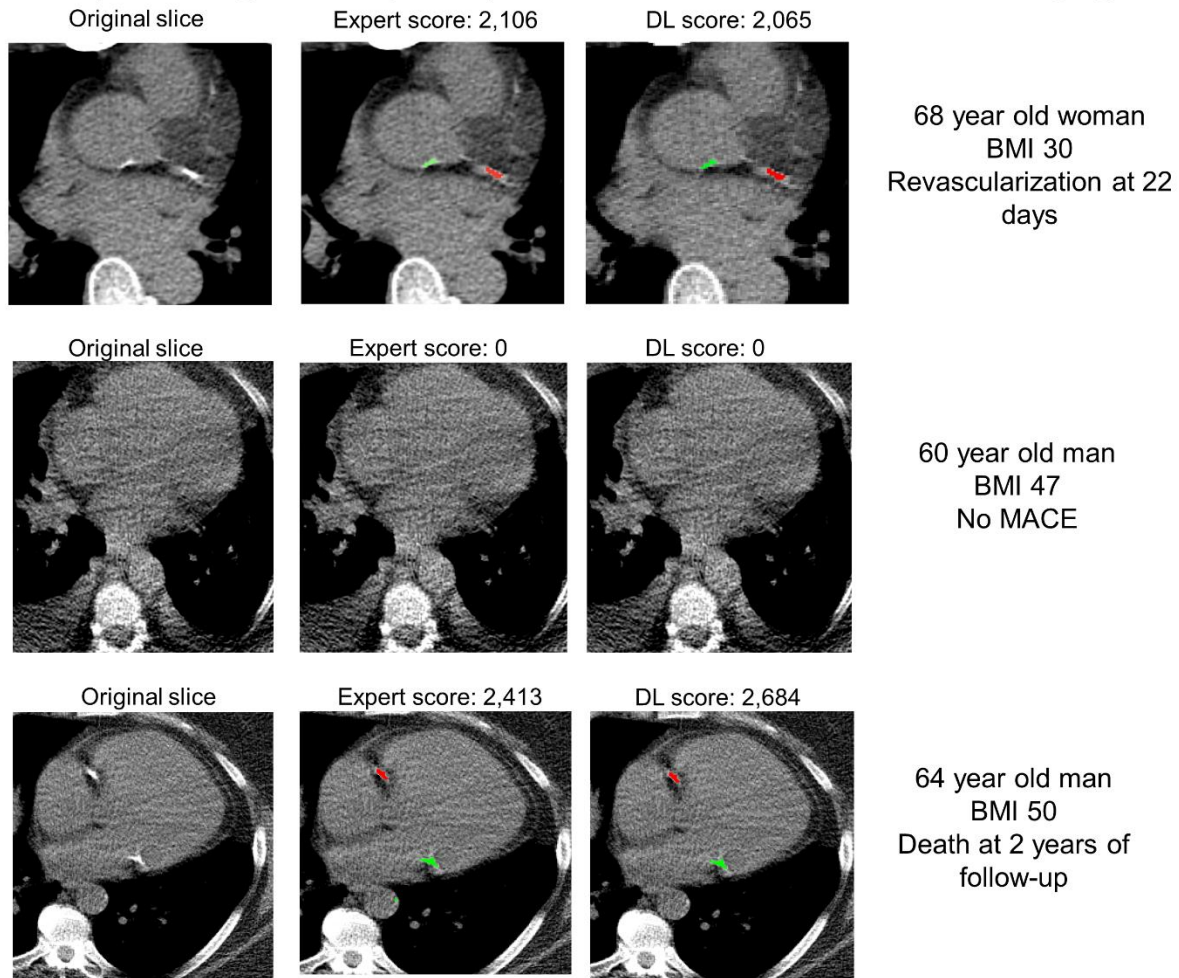


Figure 2: Examples of expert compared to deep learning (DL) coronary artery calcium (CAC) scores. The model identifies coronary calcium (red) as well as non-coronary calcium. In case 1, expert and DL annotations identified similar left circumflex CAC (red) as well as ascending aorta calcium (green). No CAC was identified by either expert or DL scoring in case 2. In case 3, expert and DL annotations identified similar right coronary artery CAC (red) as well as mitral annular calcification (green). BMI – body mass index, MACE – major adverse cardiovascular events.

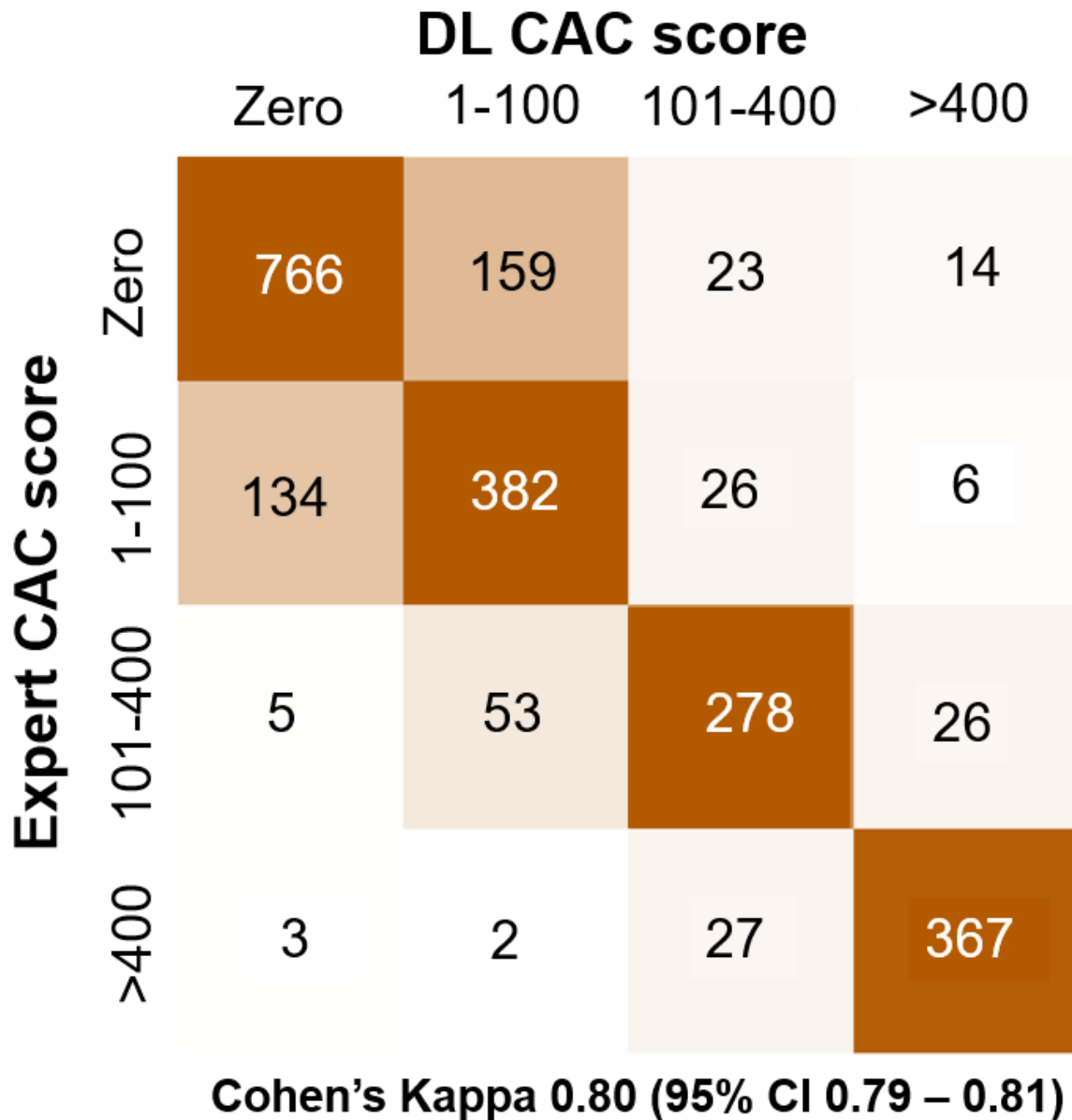


Figure 3: Concordance matrix between deep learning (DL) and expert coronary artery calcium (CAC) categories in the external testing population. CI - confidence interval.

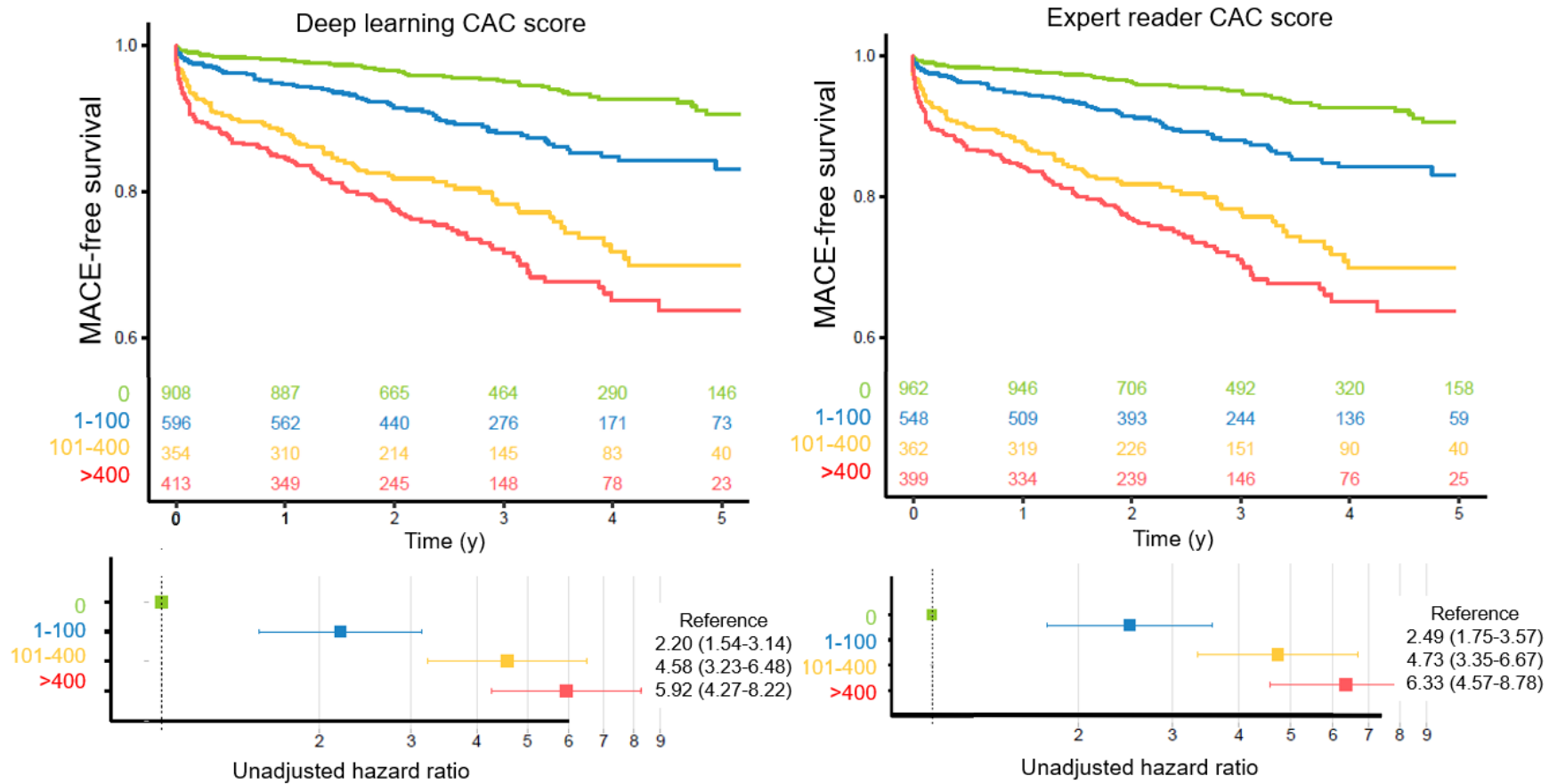


Figure 4. Kaplan-Meier survival curves for major adverse cardiovascular events (MACE). Increasing coronary artery calcium (CAC) category was associated with increasing risk of MACE for deep learning and expert annotated CAC scores on SPECT/CT attenuation maps.

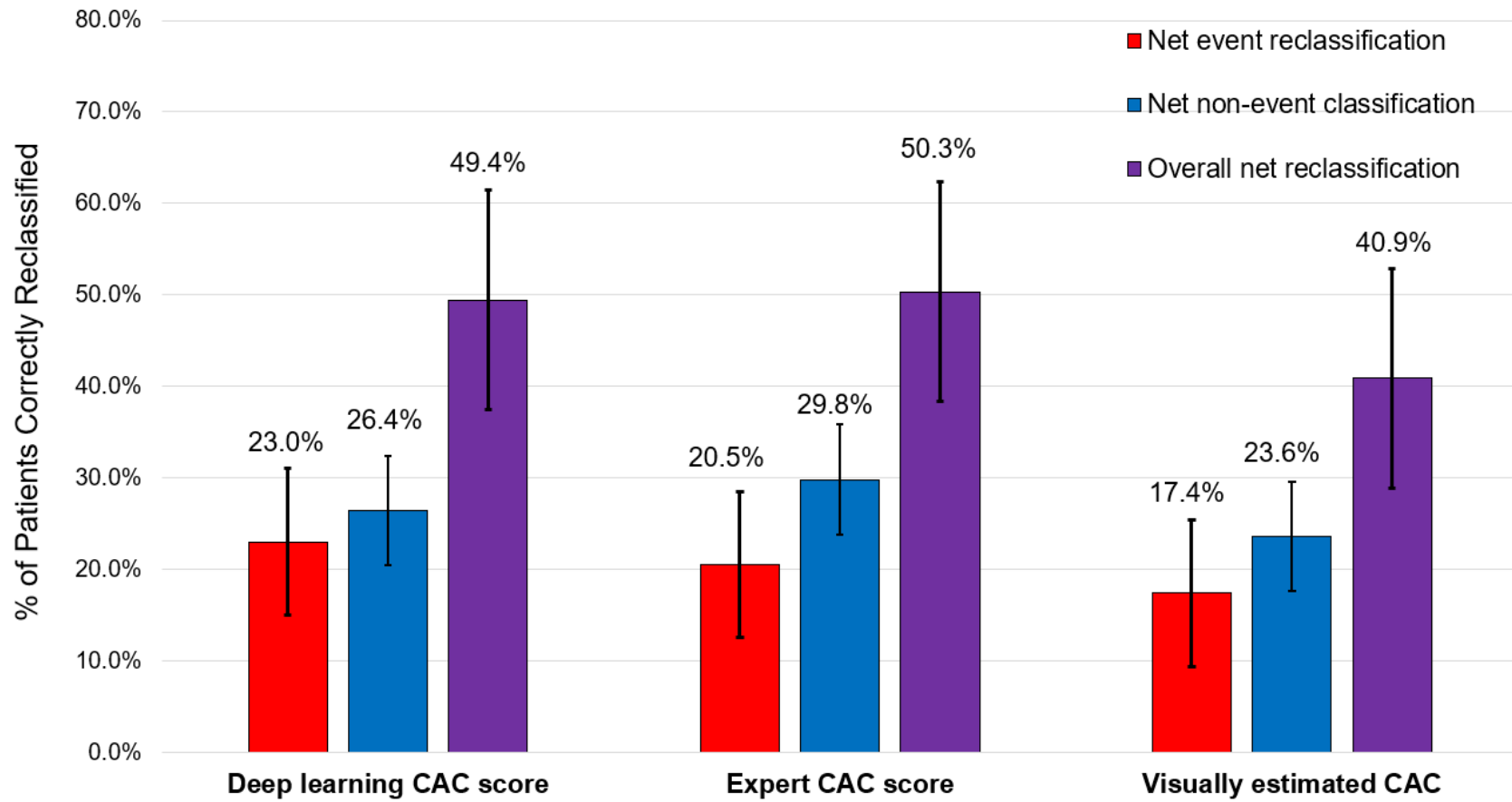


Figure 5: Results of the net-reclassification analysis. We assessed the addition of coronary artery calcium (CAC) categories to the full multivariable model outlined in Table 2.

## TABLES

	CAC < 1 n = 908, 40.0%	CAC 1-100 n = 596, 26.2%	CAC 100 – 400 n = 354, 15.6%	CAC > 400 n = 413, 18.2%	p-value
Age, median (IQR)	61.9(55.1, 69.3)	66.4(57.3, 74.2)	70.8(65.3, 77.3)	72.3(66.3, 77.9)	<0.001
Male, n(%)	368(40.5%)	293(49.2%)	200(56.5%)	286(69.2%)	<0.001
BMI, median (IQR)	29.3(25.1, 32.6)	30(25.8, 34.4)	29.3(25.4, 32.9)	29.4(25.2, 32.4)	0.048
Past Medical History, n(%)					
Hypertension	423(46.6%)	355(59.6%)	240(67.8%)	268(64.9%)	<0.001
Diabetes	136(15.0%)	146(24.5%)	111(31.4%)	140(33.9%)	<0.001
Dyslipidemia	334(36.8%)	246(41.3%)	187(52.8%)	236(57.1%)	<0.001
Family history	453(49.9%)	305(51.2%)	155(43.8%)	205(49.6%)	0.20
Smoking	67(7.4%)	35(5.9%)	21(5.9%)	27(6.5%)	0.67

Table 1. Patient characteristics according to coronary artery calcification (CAC) category determined by the deep-learning model. BMI – body mass index, IQR – interquartile range.

	Unadjusted HR (95% CI)	p-value	Adjusted HR (95% CI)	p-value
DL CAC Categories				
CAC <1	Reference	--	Reference	--
CAC 1 – 100	2.20(1.54–3.14)	<0.001	1.90(1.32–2.73)	<0.001
CAC 101 – 400	4.58(3.23–6.48)	<0.001	3.32(2.29–4.81)	<0.001
CAC > 400	5.92(4.27–8.22)	<0.001	3.58(2.47–5.19)	<0.001
Age (per 10 years)	1.37(1.24–1.52)	<0.001	1.12(1.00–1.26)	0.046
Male	1.75(1.39–2.19)	<0.001	1.11(0.86–1.43)	0.418
BMI (per kg/m <sup>2</sup> )	0.98(0.96–1.00)	0.021	0.99(0.97–1.01)	0.157
Hypertension	1.22(0.98–1.53)	0.079	0.98(0.77–1.25)	0.862
Diabetes	1.60(1.26–2.02)	<0.001	1.28(0.99–1.64)	0.060
Dyslipidemia	1.34(1.08–1.67)	0.008	1.00(0.78–1.27)	0.997
Family history	0.82(0.65–1.02)	0.071	0.90(0.72–1.13)	0.353
Smoking	1.18(0.81–1.72)	0.389	1.18(0.80–1.74)	0.415
Stress AC TPD Category				
TPD < 1%	Reference	--	Reference	--
TPD 1 - < 5%	1.28(0.96–1.71)	0.097	1.22(0.90–1.65)	0.200
TPD 5 - <10%	2.06(1.46–2.90)	<0.001	1.70(1.19–2.44)	0.004
TPD ≥ 10%	7.52(5.43–10.4)	<0.001	4.73(3.02–7.46)	<0.001
Rest AC TPD	1.07(1.05–1.08)	<0.001	1.00(0.97–1.03)	0.836
Stress LVEF	0.97(0.97–0.98)	<0.001	0.99(0.98–1.00)	0.293

Table 2: Associations with major adverse cardiovascular events. AC – attenuation correction, BMI

– body mass index, CAC- coronary artery calcium, DL – deep learning, HR – hazard ratio, LVEF

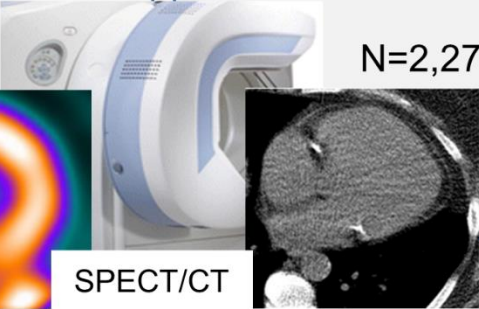
– left ventricular ejection fraction, TPD – total perfusion deficit.



# Graphical Abstract

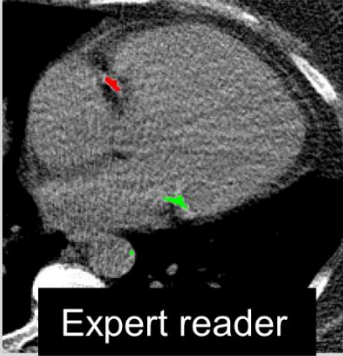
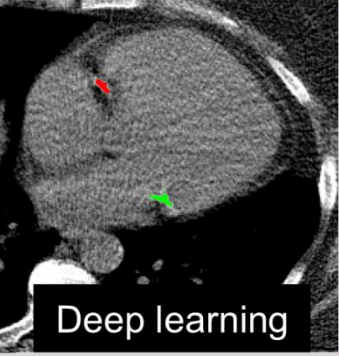
**Training population**  
2 sites  
Attenuation correction  
CT scans  
N=6,608

**External testing**  
SPECT perfusion and  
attenuation correction  
CT scans  
N=2,271



SPECT/CT

Coronary artery calcium scoring on  
attenuation correction CT scans

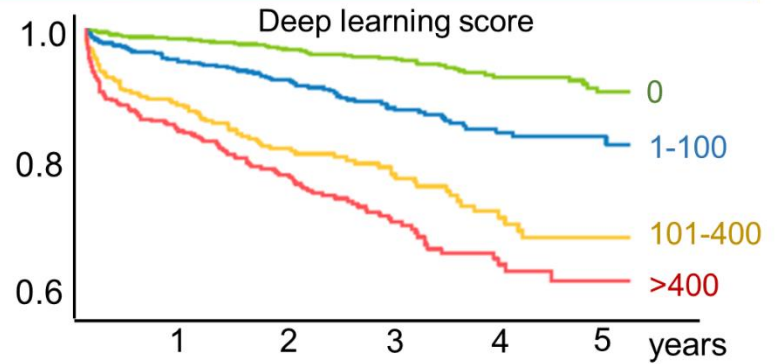
Expert reader

Deep learning

## Agreement with experts

		<b>Deep learning score</b>			
		0	1-100	101-400	>400
Expert score	0	766	159	23	14
	1-100	134	382	26	6
	101-400	5	53	278	26
	>400	3	2	27	367

## Risk stratification for major adverse cardiovascular events



## **SUPPLEMENTAL MATERIAL**

Supplemental Methods:

### Clinical Data:

Past medical history and family history were prospectively collected in the Alberta Provincial Project for Outcome Assessment in Coronary Heart Disease (APPROACH) database. Follow up for MACE in the external testing population was obtained through the Discharge Abstracts/National Ambulatory Care Reporting system and Alberta Vital Statistics. MACE was defined as revascularization, non-fatal MI, admission for unstable angina, or all-cause mortality.

### Myocardial Perfusion Image Analysis

Quality control for MPI was performed by experienced core laboratory technologists without knowledge of the clinical data. Stress and rest images were analyzed by Quantitative Perfusion SPECT (QPS) software (Cedars-Sinai Medical Center, Los Angeles, CA) as previously described to quantify total perfusion deficit (TPD) (16). TPD is a continuous measure which incorporates both the extent and severity of perfusion defects (16). Attenuation-corrected (AC) TPD was used for all analyses. Left ventricular ejection fraction (LVEF) was calculated from post-stress gated images.

### CTAC Image Acquisition and Interpretation

At the University of Calgary, CTAC was performed using a built in CT scanner (Lightspeed VCT 64, GE, Boston, USA). CTAC imaging was performed after the rest acquisition during end-expiratory breath hold with no ECG-gating, in helical mode with a slice thickness of 5-mm, tube voltage of 120 kVp and 20 mA, using a 512x512 matrix. CTAC images were reviewed at the time of SPECT/CT MPI reporting and coronary calcium was graded visually as: absent, equivocal, present or extensive. Extensive calcification was defined as visually estimated CAC greater than 400. For comparisons with expert and DL annotated CAC scores, equivocal was combined with

present due to the low number of patients with equivocal visual CAC (n=28). Importantly, the expert visual estimates were informed by clinical information and perfusion findings. Details for the training population are available in [Supplemental Table 1](#).

### Expert CAC Annotations

The Expert reader annotation process included pixel by pixel calcification assignment into coronary calcification or non-coronary calcification using Cardiac Suite (Cedars Sinai Medical Center, Los Angeles, CA). Coronary calcium was annotated according to the involved vessel as LAD, LCX, LM and RCA. Non coronary calcification included calcification in the mitral valve, ascending aorta, descending aorta, aortic arch, aortic valve, tricuspid valve, pulmonary valve and pericardium. The DL model used these 2 categories to distinguish between coronary and non-coronary calcifications. CAC was quantified as previously described using the weighted sum of lesions with a density above 130 Hounsfield units, and multiplying the area of calcium by a factor related attenuation(6).

### Model architecture

The model was built using PyTorch. We automatically segmented CAC from CTAC using a cascaded system of convLSTM(17). This system consists of two networks, first of which is trained for segmentation of the heart silhouette and the second network was trained to segment the CAC. The heart convLSTM was trained on a subset of training data with expert reader annotations for QFAT software(18). A supervised learning regime was used for both segmentation networks.

The input to the network consists of a CT slice(single slice input) along with the previous and next slices (Sequential input). This process was completely automated and there were no exclusions. The convolutional LSTM block takes in the sequential input to imitate the radiologist approach of sliding across various slices after looking at a single slice of interest. The output of the network

consists of an attention weighted combination of the results from the sequential input and the single slice of interest (17). Softmax function is applied on the output to classify each pixel as background, coronary calcification or non-coronary calcification.

The heart mask was applied to the final CAC prediction to reduce any spurious bone overcalling or calcification in non-cardiac regions. In order to imitate the radiologist approach of aggregating information from adjacent slices, multiple slices were provided to both the networks as input and an attention map was generated by the convLSTM. The segmentation uses the attention weighted combination of the results from the sequential input and the single slice of interest which is later passed on to the softmax layer for final lesion mask creation. To counter the large class imbalance between CAC and background, we used subset sampling of the majority class as well as focal loss(19) as cost function between the ground truth expert reader annotation and network generated mask. The network was shown previously to have significantly reduced memory consumption for training and almost 2x faster inference times on a typical CPU(17).

### Statistical Analysis

Continuous variables were summarized as mean (standard deviation [SD]) if normally distributed and compared using a Student's t-test. Continuous variables that were not normally distributed were summarized as median (interquartile range [IQR]) and compared using a Mann-Whitney U-test. Associations with MACE were assessed with univariable and multivariable Cox proportional hazards analyses. Net reclassification index (NRI) was used to assess the additive prognostic utility of DL and expert annotated CAC. NRI was calculated when added to all other components of the multivariable model including: age, sex, past medical history, stress and rest AC TPD, and LVEF. Improvement in likelihood ratio chi-square (as a measure of model fit) and improvement in area under the receiver operating characteristic curve were also assessed. We also performed a

sensitivity analysis evaluating associations with hard adverse outcomes (death or MI) as well as an analysis in which patients who underwent early revascularization (revascularization within 90 days of SPECT/CT MPI) were excluded (n=52) since this may alter long-term outcomes (20,21). All analyses were performed using Stata/IC version 13.1 (StataCorp, College Station, Texas, USA) and R (version 4.1.2).

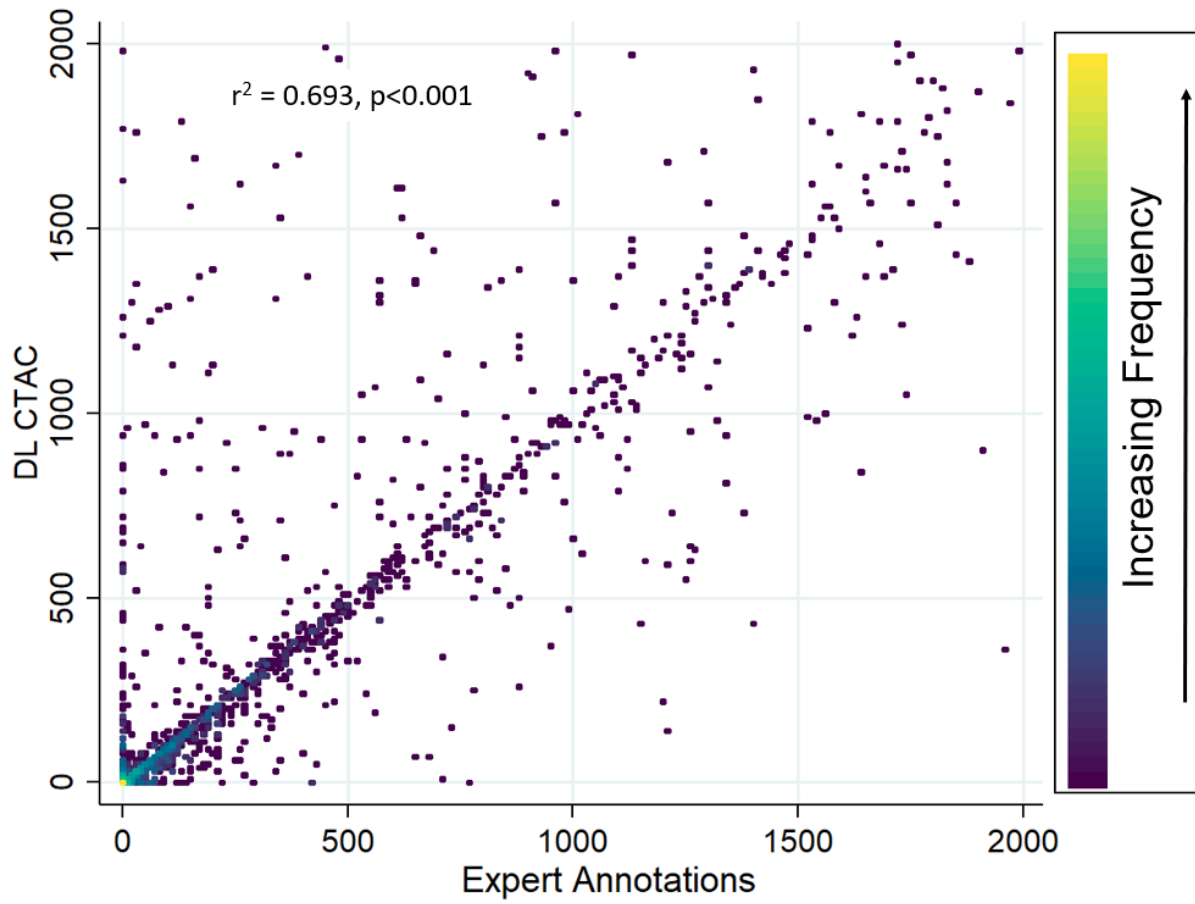
## **SUPPLEMENTAL RESULTS**

### Review of Discrepant Cases

There were 142 cases with DL CAC of 0 with expert annotated CAC >0, which were reviewed manually to determine the most likely cause. This review identified image noise (n=125, 88%), confusion with valve-related calcium (n=13, 9%) or confusion with aortic calcification (n=10, 7%) as contributing to these cases. There were 196 cases with DL CAC >0 with expert annotated CAC of 0, which were reviewed manually to determine the most likely cause. This review identified image noise (n=159, 81%), confusion from pacing devices (n=19, 10%), confusion with valve-related calcium (n=9, 5%), confusion with aortic calcification (n=9, 5%), one case of calcified pericardium and one case of calcified lymph nodes as contributing to these cases.

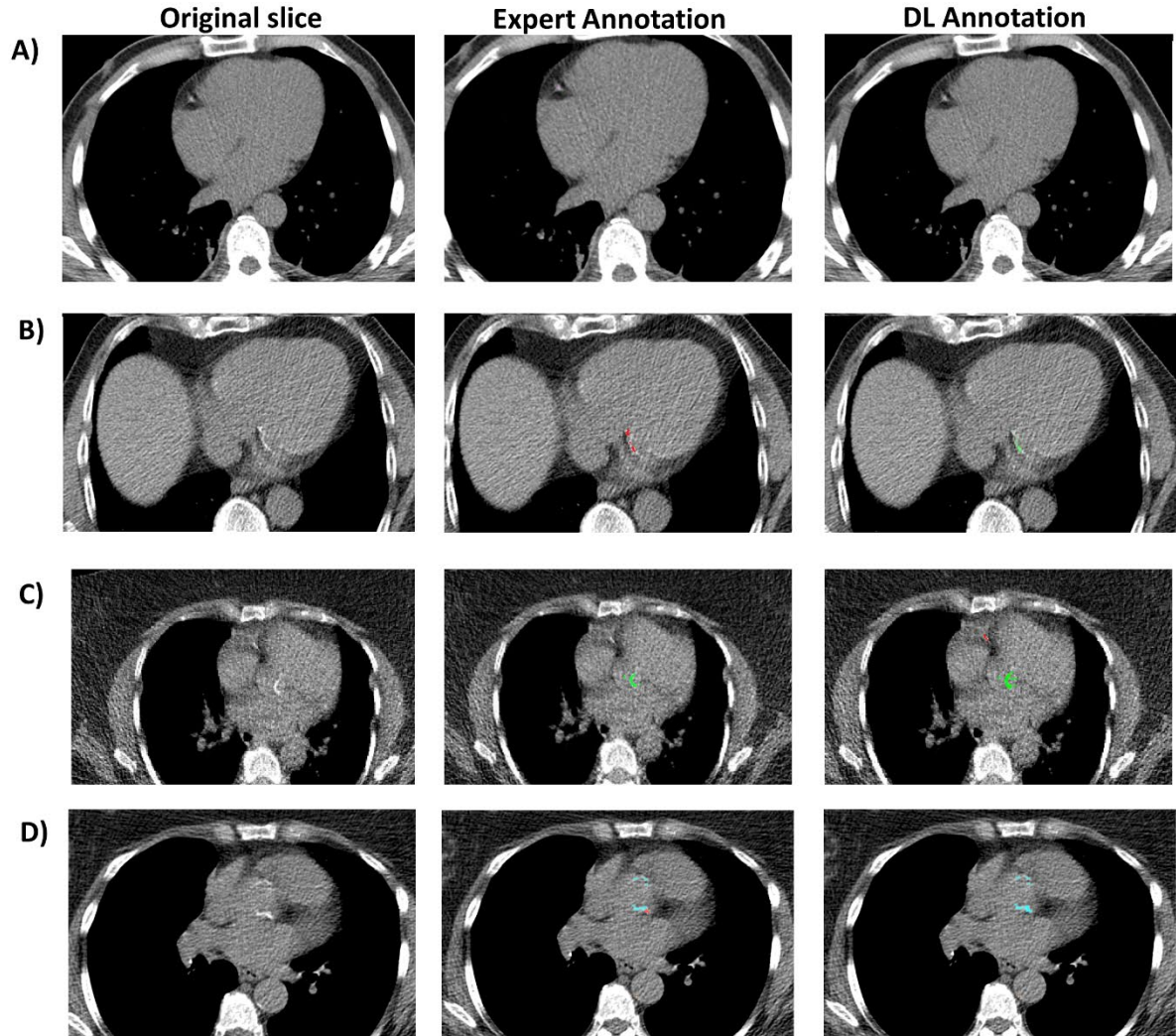
## SUPPLEMENTAL FIGURES

Supplemental Figure 1:



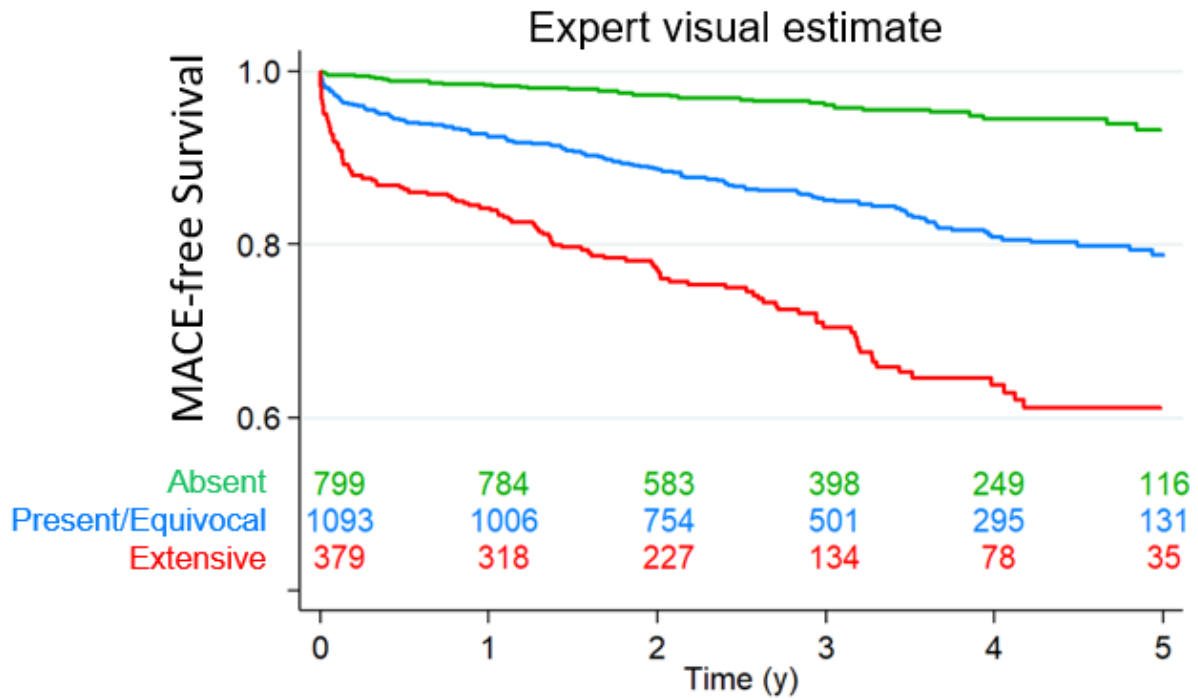
Supplemental Figure 1: Pair-wise correlation between DL CTAC and expert annotations.

Supplemental Figure 2



Supplemental Figure 2: Case examples of discrepant expert and deep learning (DL) annotations. In panel A, there is a small area of calcification identified by the expert in the right coronary artery (red) which was not identified by DL. In panel B, there is calcification that was annotated as belonging to the left circumflex by the expert reader (red) but attributed to mitral annular calcification by DL (green). In Panel C, there is aortic valve calcification noted by expert reader and DL (green), but also a small area of calcification noted in the RCA (red) which was annotated only by DL. In panel D, there is an area of calcification attributed to the left main artery by the expert reader (red) and the ascending aorta by DL (light blue).

Supplemental Figure 3



Supplemental Figure 3. Kaplan-Meier survival curves for major adverse cardiovascular events (MACE). Increasing visual coronary artery calcium estimate was associated with increasing risk of MACE. Equivocal and present were considered as a single category due to the low number of patients with equivocal visually estimated coronary artery calcium (n=28).



Supplemental Table 1:

	Training Population N = 6608	Site 1 n = 1827	Site 2 n = 4781	p-value	External Population N=2271
Age, Median (IQR)	63 (55, 72)	61 (53, 69)	64 (56, 73)	<0.001	69 (58, 74)
Male	3447 (52%)	796 (44%)	2651 (55%)	<0.001	1147 (51%)
Hypertension	4397 (68%)	1266 (70%)	3131 (67%)	0.038	1286 (57%)
Diabetes	1757 (27%)	487 (27%)	1270 (27%)	0.7	533 (23%)
Previous PCI	550 (8.5%)	4 (0.2%)	546 (12%)	<0.001	0 (0%)
Previous CABG	273 (4.2%)	0 (0%)	273 (5.8%)	<0.001	0 (0%)
Expert CAC, median (IQR)	52 (0, 480)	13 (0, 179)	88 (0, 602)	<0.001	15 (0, 208)
Expert CAC Category				<0.001	
CAC = 0	2196 (33%)	666 (36%)	1530 (32%)		962 (42%)
CAC 1-100	1508 (23%)	585 (32%)	923 (19%)		548 (24%)
CAC 101-400	1096 (17%)	281 (15%)	815 (17%)		362 (16%)
CAC >400	1808 (27%)	295 (16%)	1513 (32%)		399 (18%)
Slice Thickness (mm)				<0.001	
2.5	4511 (68%)	0 (0%)	4511 (94%)		0 (0%)
3	1827 (28%)	1827 (100%)	0 (0%)		0 (0%)
5	270 (4.1%)	0 (0%)	270 (5.6%)		2271 (100%)
Kilovolt potential				<0.001	
110	1 (<0.1%)	1 (<0.1%)	0 (0%)		0 (0%)
120	6401 (97%)	1620 (89%)	4781 (100%)		2271 (100%)
130	201 (3.0%)	201 (11%)	0 (0%)		0 (0%)
Unknown	5	5	0		
Tube current, median (IQR)	60 (60, 246)	416 (331, 568)	60 (60, 60)	<0.001	20 (20, 20)

Supplemental Table 1: Population characteristics for the training populations. CABG – coronary artery bypass grafting, CAC – coronary artery calcium, IQR – interquartile range, PCI – percutaneous coronary intervention.

Supplementary Table 2.

Section	Checklist item	Location in the manuscript
<b>1</b>	<b>Designing the study plan</b>	
1.1	Describe the need for the application of machine learning to the dataset	Page 3, par 2
1.2	Describe the objectives of the machine learning analysis	Page 4, par 1
1.3	Define the study plan	Pages 6 and 7
1.4	Describe the summary statistics of baseline data	Page 7 and 8, Table 1
1.5	Describe the overall steps of machine learning workflow	Figure 1 and page 6 and 7
<b>2</b>	<b>Data standardization, feature engineering, and learning</b>	
2.1	Describe how the data were processed in order to make it clean, uniform, and consistent	Figure 1 and page 6 and 7
2.2	Describe whether variables were normalized and if so, how this was done	N/A
2.3	Provide details on the fraction of missing values (if any) and imputation methods	No missing values
2.4	Perform and describe feature selection process	N/A
2.5	Identify and describe the process to handle outliers if any	N/A
2.6	Describe whether class imbalance existed, and which method was applied to deal with it	Page 6 and 7
<b>3</b>	<b>Selection of Machine Learning Model</b>	
3.1	Explicitly define the goal of the analysis e.g., regression, classification, clustering	Page 6 and 7
3.2	Identify the proper learning method used (e.g., supervised, reinforcement learning etc.) to address the problem	Page 6 and 7
3.3	Provide explicit details on the use of simpler, complex, or ensemble models	N/A
3.4	Provide the comparison of complex models against simpler models if possible	N/A
3.5	Define ensemble methods, if used	N/A
3.6	Provide details on whether the model is interpretable	Page 6 and 7

<b>4</b>	<b>Model Assessment</b>	
<b>4.1</b>	Provide a clear description of data used for training, validation, and testing	Figure 1, page 6
<b>4.2</b>	Describe how the model parameters were optimized (e.g., optimization technique, number of model parameters etc.)	N/A
<b>5</b>	<b>Model Evaluation</b>	
<b>5.1</b>	Provide the metric(s) used to evaluate the performance of the model	Pages 10-12
<b>5.2</b>	Define the prevalence of disease and the choice of the scoring rule used	Page 6 and 8
<b>5.3</b>	Report any methods used to balance the numbers of subjects in each class	Page 6 and 7
<b>5.4</b>	Discuss the risk associated to misclassification	Page 9 and 10
<b>6</b>	<b>Best Practices for Model Replicability</b>	
<b>6.1</b>	Consider sharing code or scripts on public repository with appropriate copyright protection steps for further development and non-commercial use	Page 4
<b>6.2</b>	Release data dictionary with appropriate explanation of the variables	N/A
<b>6.3</b>	Document version of all software and external libraries	Page 7
<b>7</b>	<b>Reporting limitations, biases and alternatives</b>	
<b>7.1</b>	Identify and report the relevant model assumptions and findings	Page 12
<b>7.2</b>	If well-performing models were tested on a hold-out validation dataset, detail the data of that validation set with the same rigor as that of the training dataset (see section 2 above)	N/A

Supplementary Table 2. Proposed Requirements for Cardiovascular Imaging-Related Machine Learning Evaluation (PRIME) Checklist

Supplemental Table 3:

	<b>Deep Learning Coronary Artery Calcium Score</b>			
Visual Estimate	0	1-100	101-400	>400
Absent	623 (27.4%)	145 (6.4%)	17 (0.7%)	14 (0.6%)
Equivocal/Present	281 (12.4%)	429 (18.9%)	260 (11.4%)	123 (5.4%)
Extensive	4 (0.2%)	22 (1%)	77 (3.4%)	276 (12.2%)
	<b>Expert Coronary Artery Calcium Score</b>			
Visual Estimate	0	1-100	101-400	>400
Absent	740 (32.6%)	53 (2.3%)	2 (0.1%)	4 (0.2%)
Equivocal/Present	219 (9.6%)	478 (21%)	279 (12.3%)	117 (5.2%)
Extensive	3 (0.1%)	17 (0.7%)	81 (3.6%)	278 (12.2%)

Supplemental Table 3: Classification by visually estimated coronary artery calcification compared to deep-learning or expert annotated coronary artery calcium score. Equivocal and present were considered as a single category due to the low number of patients with equivocal visually estimated coronary artery calcium (n=28).

Supplemental Table 4

	<b>No MACE n=1951</b>	<b>MACE n=320</b>	<b>p-value</b>
DL CAC, median (IQR)	11.1 (0, 151.0)	178.0 (25.4, 851.9)	<0.001
DL CAC Categories			
CAC <1	856 (43.9%)	52 (16.3%)	<0.001
CAC 1 – 100	524 (26.9%)	72 (22.5%)	<0.001
CAC 101 – 400	273 (14.0%)	81 (25.3%)	<0.001
CAC > 400	298 (15.3%)	115 (35.9%)	<0.001
Age, median (IQR)	66.1 (58.1, 73.6)	70.7 (61.4, 77.2)	<0.001
Male, n(%)	946 (48.5%)	201 (62.8%)	<0.001
BMI, median (IQR)	29.8 (25.5, 33.1)	29.1 (24.9, 32.5)	0.069
Past Medical History, n(%)			
Hypertension	1092 (56.0%)	194 (60.6%)	0.12
Diabetes	432 (22.1%)	101 (31.6%)	<0.001
Dyslipidemia	844 (43.3%)	161 (50.3%)	0.019
Family history	978 (50.1%)	140 (43.8%)	0.029
Smoking	120 (6.2%)	30 (9.4%)	0.025
Stress AC TPD, median (IQR)	2.4 (0.8, 5.2)	6.2 (2.5, 13.9)	<0.001
Stress AC TPD Category			
Stress AC TPD < 1%	555 (28.4%)	36 (11.3%)	<0.001
Stress AC TPD 1 - < 5%	889 (45.6%)	109 (34.1%)	<0.001
Stress AC TPD 5 - <10%	337 (17.3%)	65 (20.3%)	<0.001
Stress AC TPD ≥ 10%	170 (8.7%)	110 (34.4%)	<0.001
Rest AC TPD, median (IQR)	0 (0, 0.6)	0.4 (0, 3.0)	<0.001
Stress LVEF, median (IQR)	67 (59, 74)	61 (51, 71)	<0.001

Supplemental Table 4. Patient characteristics in patients who experienced major adverse cardiovascular events (MACE) compared to those who did not. AC – attenuation correction, BMI – body mass index, CAC – coronary artery calcium, DL – deep learning, IQR – interquartile range, LVEF – left ventricular ejection fraction, TPD – total perfusion deficit.

Supplemental Table 5:

	Death or myocardial infarction		Death	
	Adjusted HR (95% CI)	p-value	Adjusted HR (95% CI)	p-value
DL Categories				
CAC <1	Reference	Reference	Reference	Reference
CAC 1 – 100	1.67 (1.09 – 2.55)	0.018	1.53 (0.95 – 2.45)	0.078
CAC 101 – 400	2.38 (1.53 – 3.71)	<0.001	2.22 (1.37 – 3.60)	0.001
CAC > 400	2.55 (1.64 – 3.97)	<0.001	2.30 (1.42 – 3.74)	0.001
Expert Categories				
CAC <1	Reference	Reference	Reference	Reference
CAC 1 – 100	1.52 (0.98 – 2.35)	0.059	1.43 (0.88 – 2.31)	0.147
CAC 101 – 400	2.19 (1.42 – 3.39)	<0.001	2.06 (1.28 – 3.30)	0.003
CAC > 400	2.55 (1.64 – 3.92)	<0.001	2.15 (1.33 – 3.47)	0.002

Supplemental Table 5: Associations between coronary artery calcium (CAC) categories and secondary clinical outcomes. CI – confidence interval, HR – hazard ratio.

Supplemental Table 6:

Deep Learning CAC	Increase in LR chi-square	50.4
	Increase in AUC	0.028 (0.010 to 0.046)
	Event NRI (95% CI)	0.230 (0.142 to 0.314)
	Non-event NRI (95% CI)	0.264 (0.204 to 0.309)
	Overall NRI (95% CI)	0.494 (0.363 to 0.607)
Expert Reader CAC	Increase in LR chi-square	55.6
	Increase in AUC	0.033 (0.014 to 0.052)
	Event NRI (95% CI)	0.205 (0.120 to 0.294)
	Non-event NRI (95% CI)	0.298 (0.239 to 0.346)
	Overall NRI (95% CI)	0.503 (0.376 to 0.623)
Visually Estimated CAC	Increase in LR chi-square	32.0
	Increase in AUC	0.020 (0.007 to 0.033)
	Event NRI (95% CI)	0.174 (0.083 to 0.264)
	Non-event NRI (95% CI)	0.236 (0.177 to 0.298)
	Overall NRI (95% CI)	0.409 (0.278 to 0.537)

Supplemental Table 6: Net-reclassification analysis for the addition of coronary artery calcium (CAC) category. The reference model included all other components of the multivariable analysis outlined in Table 3. CI – confidence interval, NRI – net reclassification index.